Towards Blockchain-based GDPR-compliant Online Social Networks: Challenges, Opportunities and Way Forward

Javed Ahmed[1,2], Sule Yildirim[1], Mariusz Nowostawski[1], Mohamed Abomhara[1], Raghavendra Ramachandra[1], and Ogerta Elezaj[1]

[1] Norwegian University of Science and Technology, Norway
[2] Sukkur IBA University, Pakistan

Abstract. Online Social Networks (OSNs) are very popular and widely adopted by the vast majority of Internet users across the globe. Recent scandals on the abuse of users' personal information via these platforms have raised serious concerns about the trustworthiness of OSN service providers. The unprecedented collection of personal data by OSN service providers poses one of the greatest threats to users privacy and their right to be left alone. The recent approval of the GDPR (General Data Protection Regulation) presents OSN service providers with great compliance challenges. A set of new data protection requirements are imposed on data controllers (OSN service providers) by GDPR that offer greater control to data subjects (OSN users) over their personal data. This position paper investigates the link between GDPR provisions and the use of blockchain technology for solving the consent management problem in online social networks. We also describe challenges and opportunities in designing a GDPR-compliant consent management mechanism for online social networks. Key characteristics of blockchain technology that facilitate regulatory compliance were identified. The legal and technological state of play of the blockchain-GDPR relationship is reviewed and possible ways to reconcile blockchain technology with the GDPR requirements are demonstrated. This paper opens up new research directions on the use of the disruptive innovation of blockchain to achieve regulatory compliance in the application domain of online social networks.

Keywords: Privacy · Data Protection · GDPR · Blockchain · Online Social Network · Transparency · Subject Rights.

1 Introduction

The recent Cambridge Analytica scandal on the abuse of personal data by social media for leveraged political influence has raised serious concerns regarding technical, commercial, political and ethical aspects of personal data collection. The scandal brought up the fact that service providers, such as Facebook and Google, in particular, exhibit enormous social influence that can shake or derail the democratic foundation of western societies. Since its inception, Facebook has

collected 300 petabytes of personal data which is increasing at the speed of 4 new petabytes of data per day [27]. In the current Big Data era, data is an immensely valuable asset in an economy. It is commonly considered as the oil of the 21st century, which is not only fueling the success of the tech giants (i.e. Facebook, Google, Apple, Amazon) but also driving innovation and economic growth. The current situation is that the benefits of a data-driven society are reaped by a few multi-national organizations that make the majority of their profit through services offered to users who pay for them with their personal data. Users have little or no control over how their personal data is stored and used. In recent years, mainstream media has been recurrently addressing controversial incidents related to privacy breaches [18].

The GDPR [28, 30] came into force across Europe on 25^{th} May 2018. It aims to give back control over personal data to the data subjects by imposing new data protection requirements on data controllers and processors. GDPR presents various challenges for Online Social Networks (OSNs), not only from a legal perspective but also from a technical view, mainly in the areas of data management and automation. OSNs are not fully prepared to comply. And when they attempt to comply, there are often major gaps [4]. GDPR recognizes data subjects' consent as a legitimate ground for data processing. The main aim of promoting the notion of consent is to provide data subjects control over their personal data. At present, consent management mechanisms in OSNs are either non-existent or not GDPR compliant [11]. It is not an easy task for OSNs to attain consent compliance. A consent compliance mechanism must have certain characteristics that are acceptable to both data subjects and data controllers. These issues uncover new research directions and pose interesting research challenges.

In this paper, we identify challenges imposed by the new GDPR data protection requirements for OSN service providers. These requirements aim to offer OSN users more control over their personal data, while at the same time enabling transparency in data processing and sharing activities carried out by the service providers and third parties. We also identify some of the opportunities offered by the disruptive innovation of blockchain that facilitates regulatory compliance by maintaining tamper-evident audit logs for information accountability. We also explore whether it is possible to reconcile blockchain with new requirements of GDPR. This research paves the way for designing a block-based GDPR compliant consent management model for personal data processing and sharing in online social networks. The key characteristics of blockchain technology including transparency and decentralization add value to the consent management model for regulatory compliance.

An explanation for the use of blockchain technology for GDPR compliance is still necessary, as the current research literature suggests that two initiatives (GDPR and Blockchain) are at odds [20]. They appear to be at odds with each other until the underlying principles of GDPR and Blockchain are observed. Both share common principles of data privacy and give data subjects more control over their digital private data. Both GDPR and blockchain aspire to increase integrity, trust, and transparency in a generally unsafe environment. GDPR does

so by imposing responsibilities upon data controllers and processors. GDPR assumes to the extent that data controllers and processors are centralized actors with control over the system. GDPR compliance approaches based on a centralized architecture result in limited transparency and a lack of trust. On the other hand, blockchain ensures trust and transparency by utilizing the computational power of the masses and by sharing the register with all peers in the P2P network. The unprecedented transparency provided by blockchain technology sits uneasily with GDPR obligations related to privacy and information confidentiality. The dilemma with adopting blockchain for consent management is in finding the trade-off between transparency and information confidentiality. One of the solutions is to use private blockchain that allows only permitted parties to have access to all transactions. However, private blockchain thus loses the primary advantage of decentralization. Moreover, a dishonest central authority is capable of tampering the transaction history for personal gain. Wang et al. [32] proposed a framework that preserves information confidentiality without compromising transparency using zero-knowledge proof (ZKP). We conclude that prominent features of the blockchain technology can be effectively utilized to manage personal data full compliance with the GDPR legislation.

The rest of the paper is organized as follows. In section 2, we provide the theoretical background of GDPR, Blockchain and Online social networks and how they interplay to achieve regulatory compliance. This section also presents a comparative analysis of existing literature in the domain. In section 3, we identify challenges imposed by the new regulatory requirements of GDPR for OSN service providers. To better understand legal obligations, we also identify various stake-holders involved in data processing and sharing activities carried by social web systems. In section 4, we discuss some of the opportunities offered by inherent blockchain technology features for designing GDPR compliant online social networks. Finally, we conclude the paper with future research directions and open research questions in section 5.

2 Theoretical Foundation

This section describes the basic building blocks of online social networks, blockchain technology and general data protection regulation (GDPR). We also discuss their interplay that facilitate regulatory compliance.

2.1 Online Social Networks

Online social networks have undergone exponential growth in the last decade. Topmost visited sites by internet surfer are online social networks [3]. Surfing social media is the fourth most popular activity on the internet nowadays [4]. Socializing with friends and family across the globe via online social networks is

[3] Alexa http:/www.alexa.com/topsites
[4] Nielsen http://www.nielsen.com/

a cost-effective mechanism for the masses. A large proportion of the success of these platforms can be attributed to the fact that they allow their users to create their own space and a great way to connect with like-minded people, learn and share knowledge. Online social networks promote the vision of a human-centric web that is a key breakthrough attributed to these platforms. The primary source of information in the human-centric web is users, their network and interests that reside entirely in social networking services [2]. A widely used definition given by Boyd et al. [8] captures all the key elements of OSNs. The authors define OSNs as follows:

> An online social network is a web-based service that allows individuals to construct a public or semi-public profile within the service; articulate a list of other users with whom they share a connection, and view and traverse their list of connections and those made by others within the system.

According to this definition, every OSN user can create his/her own profile. An OSN user maintains his digital persona using profile that contains a number of attributes related to user such as demographics information, interests, preferences and various types of user generated content. Connections is another important feature offered by online social networks. Connections refer to existing social relationships of the OSN users. Many online social networks label connections as a friend, however, it is problematic due to the reason that all connections in the network are treated equally. Whereas, connections established between many users whose relationship may be better described by a different label such as family, friends, colleagues, acquaintances, etc. Labeling connections as friends precludes differentiation for selective information sharing. According to the aforementioned definition, the third feature of OSNs is traversing connection. This feature allows OSN users to find each other and construct a networked community within which they can share information.

The most popular fora for self-representation and user interactions are online social networks these days. Internet surfers join social networks to present themselves and share huge amounts of personally identifiable information. This behavior of OSN users causes a serious privacy threat to them. Online social network usage gives rise to several privacy threats including privacy threats related to OSN users [15], third party applications [3] and OSN service providers [16]. Online social networks provide a multitude of privacy tools to mitigate privacy threats concerning to OSN users. Despite the array of privacy controls, current online social networks fail to provide an effective mechanism to manage access to uploaded user content. The main reason for this failure is the shortcoming of online social networks to represent diverse social relationships. This problem has been addressed extensively in existing literature [1]. In this paper, we are addressing the issue of privacy threats related to service providers and third parties.

Privacy threat related to a service provider involves the relationship between user and service provider that is based on trust. The service provider has full access to any user data because the OSN's underlying system is designed and

configured by the service provider. The important privacy threats concerning service providers are data retention, data selling, and targeted marketing. Recent scandals like Cambridge Analytica have shaken the trust that users have put in their service providers. Moving towards a decentralized architecture is one of the most straightforward solutions to mitigate service provider related privacy threats. Decentralization is among the distinct blockchain features that play a vital role in challenging the monopoly of financial institutions by introducing cryptocurrencies. The blockchain-based solution may provide decentralization to address service provider related privacy threats.

With the emergence of Web 2.0, online social networks are offering open platforms to enable external developers to build applications that provide seamless integration of profile data with third-party applications. Facebook [5] and Google's Open Social [6] are leading this effort. These platforms have opened doors for external developers to launch their applications for social networking sites. These third-party applications pose serious privacy risks for online social network users because installed applications receive privileges equal to those of profile owners and can access users' profile data. Online social network users are unaware of the amount of data being exposed to external developers because such information flow is hidden from, or not clear to users. The mainstream media also recognized this data breach [7]. Enabling transparent data flow between service providers and application developers can mitigate privacy threats related to third parties. Transparency is a distinct feature of blockchain that may have a vital role in consent-based privacy compliance.

2.2 Blockchain

Due to the hype and success of cryptocurrencies, blockchain attracted significant interest from governments, businesses, capital markets, and the research community. It is foreseen as the core backbone of many future technologies such as the internet of things, smart cities, etc. Originally, the term blockchain was coined by a person using the name Satoshi Nakamoto in 2008 [22]. Blockchain technology is still in its infancy stage of development. The first phase of its development is termed as Blockchain 1.0. This embryo phase of blockchain development deals with cryptocurrencies such as bitcoin, ether, etc. Blockchain 2.0 refers to the second phase of its development which introduces the concept of smart contracts. Blockchain technology employs smart contracts to deals with issues of mutual trust and identity among participants. The next generation of blockchain technology will become a powerful tool for Industry 4.0 [24]. Blockchain technology can be understood as an emerging distributed ledger technology (DLT) that enables applications to operate in a fully decentralized fashion without the

[5] Facebook for Developers, https://developers.facebook.com/

[6] Open Social, https://www.getopensocial.com/

[7] Millions of Facebook user records exposed in data breach, https://www.telegraph.co.uk/technology/2019/04/03/millions-facebook-user-records-exposed-data-breach/

need for any trusted central authority. Blockchain is secure and transparent by design and relies on well-known cryptographic tools and distributed consensus mechanisms to provide key characteristics such as anonymity, immutability, auditability, transparency and trust [34]. To better understand the core concepts of the blockchain, this section presents an overview of the technology.

Blockchain is a sequence of blocks and each block is cryptographically linked to the previous block after validation through a consensus decision. The genesis block of the blockchain has no previous block. All the blocks are linked together in the chronological order i.e. from the genesis block to the latest block. A block consists of a block body and block header. The block body is composed of a transaction counter and transactions. The block header includes various headers such as nonce, timestamp, parent block hash, Merkle tree root hash, etc. The concept of consensus mechanism is central to blockchain technology. What blocks are added to the ledger will be decided by an agreed-upon consensus mechanism. There exist a different kind of consensus mechanism such as proof of work, proof of stake, practical byzantine fault tolerance, etc. Instead of trusting a central authority, trust is placed in the algorithms underlying the consensus mechanism. This is the basis for the characterization of blockchain as a trust-less system. A digital signature scheme based on asymmetric cryptography is used in an untrustworthy environment to validate the authentication of transactions. Consensus mechanism and asymmetric cryptography are implemented to achieve ledger consistency and security.

Current blockchain systems are divided into three categories: public blockchain, private blockchain, and consortium blockchain [34]. Public blockchain is also known as permissionless due to its open nature. Any participant node can read, write and engage in the consensus process. Public blockchain is completely transparent and decentralized in nature. The typical example of a public blockchain is Bitcoin [31]. Private blockchain is also known as permissioned due to its closed nature. The private blockchain is limited to a specific organization. In private blockchain write permission is kept centralized and read permission may be public or restricted to specific nodes. Private blockchain is closed and centralized in nature. Consortium blockchain is also known as a hybrid. It is a partially decentralized blockchain. Pre-selected nodes will engage in the consensus process. In consortium blockchain read permission may be public or restricted to specific nodes. The typical example of consortium blockchain is Hyperledger.

With the emergence of blockchain technology, many decentralized services and applications are proposed that build on or use blockchain to achieve independence from centralized service providers' monopoly. Blockchain-based online social networking services can be designed to address the aforementioned privacy threats from third party and service providers. Blockchain technology enhances transparency. Each user has complete transparency over what data is being collected about him/her and how it is accessed. Blockchain asserts data ownership and user privacy by enabling transparency. At the same time, blockchain provides anonymity to its users by allowing them to create pseudo-anonymous transactions without the need for revealing personally identifiable information about

them. Blockchain-based solutions not only provide greater transparency to OSN users regarding their personal data but also offer advantages towards regulatory compliance. In the context of newly enforced GDPR, the consent of the user is the legitimate ground of data processing. At present, most of the OSNs lack GDPR compliant consent mechanism which means users' lack of control over their personal data. Before introducing a blockchain-based consent management for OSNs, we present a brief overview of the legal obligations of GDPR on service providers and third parties.

2.3 General Data Protection Regulation

Since digital technology has profoundly changed the way how personal data is collected, accessed and used, on 25th May 2018, the European Commission (EC) implemented a new legislative framework called the General Data Protection Regulation (GDPR) (EU) 2016/679 [28, 30]. The aim of this Regulation was to remedy the shortcomings of Directive 95/46/EC and to further harmonize the data protection rules within the EU as well as raise the privacy and data protection levels of affected individuals. GDPR has reshaped the way organizations approach data protection and data privacy. Organizations dealing with personal data of EU citizens must ensure they are compliant with the new GDPR requirements. The GDPR regulations are described in 99 articles that cover all aspects of personal data processing by organizations. GDPR extends the responsibility and accountability requirements of organizations involved in processing personal data of the EU citizens.

In the context of GDPR, three main roles are identified: data subject, data controller and data processor. A data subject is the owner of personal data. Personal data means any information pertaining to an identified or identifiable natural person. Data controller determines the purposes and means of processing personal data. The data controller is the point of accountability in GDPR. A data processor is responsible for processing personal data on behalf of a controller. GDPR sets out six core data processing principles that facilitate the protection of personal data processing. The first principle deals with the lawful, fair and transparent processing of personal data. The purpose limitation principle ensures the collection of data for specific purposes and prohibits processing for incompatible purposes. The data minimization principle discourages excessive collection of personal data and only adequate data collection is advised. The collection of accurate and up to date data is handled with the principle of accuracy. As per the storage limitation principle, data should not be stored for longer than necessary. Finally, the integrity and confidentiality principle deals with the secure processing of personal data. In addition to these data processing principles, GDPR recognizes the consent of a user as a legitimate ground for data processing. Data processing principles are used to derive a set of rights for data subjects. The most important rights of data subjects under GDPR are: right of rectification, right of access, right to erasure and rights pertaining to automate processing.

GDPR is applicable to all organizations that process personal data of EU citizens whether based within the jurisdictional boundaries of the European Union or any third country. Compliance with the regulations is enforced by public authorities. Apparently, the supervisory authority in each EU member state is responsible for monitoring GDPR compliance. Data processing organizations are required to demonstrate their compliance with GDPR only in cases of data subjects lodging complaints with the supervisory authority about the misuse of their personal data. The supervisory authority has also extensive rights to access personal data processing activities in cases of suspected violations by organizations. Failure to comply with GDPR results in huge fines up to 20 million euros or 4% of the total annual profit. Due to the irregular nature of GDPR compliance verification by a supervisory authority, each organization has to prove and document that it has been continuously abiding by GDPR requirements. Moreover, due to the lack of transparency, it is beyond the capability of data subjects to recognize whether the data controller fully complies with GDPR and effectively protects his/her personal data. Therefore, a prospective GDPR compliant mechanism must inherit features of transparency and auditability to enable data subjects to oversee what data is collected and how it is processed by the data controller or processor.

3 Challenges of Blockchain-based OSNs under GDPR

With the introduction of blockchain technology, a plethora of decentralized services have been proposed that achieve independence from centralized entities. Blockchain can also be used in online social networks. Decentralization, transparency and distributed consensus give blockchain the potential to address most of the prevalent privacy concerns in OSNs. One of the initial efforts in this direction is Steem [8]. Steem is a blockchain-based social media platform that supports community building and social interaction with cryptocurrency rewards. Such OSN can be further made self-healing by a blockchain-based reputation system. One such system was proposed by Qin et al. [25]. The authors presented a blockchain-based academic social network and proposed a new consensus algorithm named proof of reputation(PoRe). Chen et al. [10] proposed a blockchain-based trusted social network that ensures privacy by doing peer-to-peer information exchange. The proposed model uses blockchain for limiting large-scale rumors spreading via online social networks. Blockchain technology has a set of very attractive features for applications in this domain. However, such applications are required to comply with the GDPR. Blockchain-based applications pose several challenges to regulatory compliance in the light of new changes imposed by GDPR. The following subsections present a detailed description of such challenges.

[8] Steem, https://steem.com/SteemWhitePaper.pdf

3.1 Challenge of Informed Consent

The concept of consent originated in the field of medicine. Consent is aimed at providing the data controllers legitimate grounds for personal data processing. Consent has various forms, such as informed, explicit, unambiguous, broad, etc. Each of these forms is quite diverse in nature. GDPR promotes the notion of consent to provide data subjects full control of their personal data. The absence of consent means a lack of control for data subjects over their personal data. Consent ensures that the data subject has ownership of their personal data and they choose how to navigate in the digital data world. In this section, the aim is to present a set of guidelines to manage consent in online social networks taking into account the provisions of the GDPR.

Online social networks like Facebook have a long history of using the personal data of their users without obtaining explicit or direct consent. One such example is the US midterm election in 2010, where the company experimented with 61 million users and in attempts to influence their voting behavior [7]. GDPR came into force in May 2018 and is applicable to all companies (data controllers) that process personal data of EU citizens, whether operating in Europe or any third country. GDPR compliance is the only way forward for data controllers handling personally identifiable information of EU data subjects. Whereas, non-compliant data controllers are subject to hefty fines and may suffer loss of reputation among the increasingly privacy-conscious users. The current state of the art reveals that consent management mechanisms in online social networks are either non-existent or not GDPR compliant. Scientific literature also reveals the case of consent misuse in OSNs and one such example is the contagion study conducted by Facebook [19]. Facebook did not require any explicit user consent for the study on the grounds that users have already given broad consent when they signed up to use the social network. The study provoked extended criticism and Facebook publicly acknowledged and apologized for its fault. Being a leading OSN service provider, Facebook has been questioned by regulators over the years about its privacy practices [9]. It is yet to be seen how Facebook and other OSN service providers manage to comply with new changes imposed by GDPR.

A consent management mechanism constitutes a major step towards becoming compliant with GDPR. It is challenging to achieve consent compliance in current online social networks. A valid consent under GDPR must be freely given, specific, informed and unambiguous. Current state of the art reveals that consent given by users to online social networks lacks granularity because OSN service providers seek consent for several purposes bundled together and users do not have the freedom to give or deny consent for each purpose. One such example is Facebook's facial recognition feature that requires users to accept all purposes even if the user finds only one of them acceptable. If a user consents in such a setting, then it is not freely given consent. Another important feature of valid consent under GDPR is specificity that promotes transparency. As per the

[9] Mark Zuckerberg Testimony: Senators Question Facebooks Commitment to Privacy, https://www.nytimes.com/2018/04/10/us/politics/mark-zuckerberg-testimony.html

aforementioned example of Facebook's facial recognition feature, single consent is sought for multiple purposes and the OSN service provider does not allow users to give specific consent for each purpose. Informed consent also promotes transparency by enabling data subjects to understand the nature of processing and the data collected and used in an intelligible format. However, the current terms and conditions of service providers do not represent a minimal set of information in plain and clear language. Data subjects should take clear affirmative action to give consent. Inactivity, silence or pre-ticked boxes do not make valid consent under GDPR.

Recently, several tools that focus on GDPR compliance have been released. [10] [11] These tools are designed to target self-assessment after completing standard questionnaire. The research community should focus on designing a consent management model that automatically checks if existing data processing and sharing activities are GDPR-compliant. Transparency with respect to the collection, processing, and sharing of personal data is a key enabler for OSN service providers to achieve GDPR compliance. Blockchain can play a vital role in order to provide said transparency with respect to personal data processing and sharing.

3.2 Challenge of Data Erasure and Amendment

GDPR introduces the concept of the right to be forgotten which empowers data subjects to request the data controller for the erasure of their personal data completely. This right is introduced in article 17 of the GDPR, but, it is not an absolute right and only applies in certain circumstances. Data subjects may request their data erasure only when one of the six legislative grounds outlined in the article applies. Moreover, the data controller can retain personal data for archiving purposes in the public interest under the legal protection of the GDPR. However, a data controller must be able to prove when one of the exceptions to the right to erasure applies. Therefore, the data controller must be able to erase the data subject's personal data from their records. The blockchain-based solution for OSNs brings many benefits, but it introduces a feature of immutability which guarantees that stored data is tamper-proof. This feature of immutability prohibits the straightforward application of the right to be forgotten. However, the regulation only applies to personal data stored in the blockchain. If the data is rendered completely anonymous, it falls out of the scope of the GDPR legal framework. Nevertheless, encryption is considered a pseudonymization technique and pseudonymized data continues to qualify as personal data and falls under the scope EU data protection regime. This section aims at setting guidelines on how to mitigate the challenge of data erasure and amendment taking into account the provisions of GDPR.

Recently, several solutions are being developed to achieve the goal of designing GDPR-compliant blockchain use cases. The straight forward solution may be

[10] GDPR Compliance Toolkit, https://info.nymity.com/gdpr-compliance-toolkit
[11] Microsoft GDPR Detailed Assessment, https://aka.ms/gdrpdetailedassessment

storing personal data off-chain where the blockchain merely holds proof that the data is valid [21]. Off-chain storage facilitates the right to erasure and amendment. Personal data could be restricted to a private permissioned blockchain instead of a public permissionless blockchain. Private permissioned blockchain will find it easier to apply the letter of the GDPR than public permissionless blockchain. Current state of art suggests many data obfuscation, encryption and aggregation techniques that can be used to turn personal data into digital signatures that are cryptographically linked to original data without actually revealing that data [13]. When applying such techniques to process personal data should take into consideration reversal and linkability risks. Farshid et al. [12] purposed design for a forgetting blockchain. The authors implemented a proof concept prototype that maintains most of the key features of blockchain technology but facilitates data erasure. The technique is applied to permissioned blockchains and it is evaluated with the help of domain experts. Bayle et al. [6] proposed a modular architecture that ensures GDPR compliance by providing the means to enforce the right to be forgotten. The mechanism to implement the right to be forgotten relies on the centralized infrastructure of the data controller. Geelkerken et a. [29] identified two ways in which the blockchain technology could be utilized to store personal data in compliance with the requirements of the GDPR. According to the authors the opaque blockchain instead of the transparent blockchain ideal for erasing and altering personal data, but it would require a trusted third party to comply with legal requirements.

From the GDPR perspective, public keys also constitute personal data. The pertinent question is why public keys cannot qualify as anonymous data. To qualify as anonymous data, the public key must irreversibly prevent the identification of a specific data subject. Blockchain history reveals that despite asymmetric encryption identification remains possible by connecting public keys with additional information that means public keys are pseudonymous data and fall under the scope of GDPR. It is not straight forward to design GDPR compliant solutions for public keys. The keys are an essential component of blockchain technology and constitute part of the transaction's metadata that is required for validation. Therefore, public keys cannot be moved to off-chain like transactional data. Despite the difficult nature of the problem, the research community suggested various techniques which include mixing services, ring signature, and zero-knowledge proof. A comprehensive overview of these techniques is presented in the research survey by Feng et al. [13]. These techniques are capable of anonymizing public keys for GDPR compliance that is evident from existing solutions based on these techniques such as CryptoNote, Monero, Zerocoin, and Zerocash.

3.3 Challenge of Identifying the Data Controller

The main roles identified in the GDPR context are the data subject, data controller and data processor. The data controller has a key role in determining the purposes and means of processing personal data in accordance with the constraints imposed by GDPR. The data controller is the point of accountability in

the new legal framework. Therefore, it must be possible to identify a data controller, which is not always easy in the context of blockchain technology. From a technical perspective, a blockchain is a network consisting of nodes that send and receive messages in decentralized fashion without a central point of control, which makes it difficult to assign roles. These roles are designed to fit the traditional centralized client-server scenario whereby a single entity offers some services to data subjects pertaining to the collection and processing of personal data. In this section, the aim is to evaluate who qualifies as a data controller in different variants of blockchain technology.

To identifying a data controller public permissionless blockchain is an object of debate and less straight-forward. Public permissionless blockchains are distributed and decentralized peer-to-peer networks, where each node can read, write and engage in the consensus process. The basic idea is to replace the traditional client-server model with one based on the collective processing of data via shared protocols. In such a setting, either no node qualifies as the data controller or every node qualifies as a data controller. Another interesting question related to data controller identification is whether all participating nodes can qualify as potential joint controllers. In principle, to qualify as a joint controller under the GDPR legal framework, the nodes jointly determine the purposes and means of processing. Whereas, permissionless blockchain is shaped by the nodes' behavior. They do not determine the modalities for data processing of other nodes. Therefore, they do not qualify as joint data controllers. The data subject adds personal data to the blockchain and controls its data using asymmetric cryptography. This prompts the question of whether a data subject himself/herself can qualify as a data controller. As a matter of fact, a data subject may be able to qualify as a data controller in some situations where he/she is adding personal data to a blockchain [14]. In the case of online social networks, the data subject (OSN user) is a content producer and manager rather than a content consumer [2], thus qualifying as a data controller. Fortunately, the situation is quite clear in private permissioned blockchains and there are two possible scenarios. In the first scenario, the community decides together with the validation rules that the blockchain implements, in which case all nodes fall under the definition of joint data controllers and share the responsibility of compliance. In the second scenario, the blockchain accepts the contribution of validators that do not participate in defining the validation rules, and these nodes fall under the definition of data processors [17].

4 Opportunities of Blockchain-based OSN under GDPR

Despite the multitude of privacy controls, the current online social networks fail to provide an effective mechanism to manage access to the uploaded content of users. The issue of privacy has received significant attention in both the research literature and the mainstream media. OSNs users are also becoming conscious of their online presence and expect to have more control, traceability, accountability, and ownership of their data. The emergence of blockchain technology

has led the computing domain towards a decentralization, transparency, and autonomy. Most of the prevalent privacy concerns in OSNs can be addressed by the disruptive innovation of blockchain technology using the aforementioned features. Blockchain-based solutions for OSNs enable users to control, trace and claim ownership of every piece of content they share. The following subsections describe in detail some of the opportunities offered by blockchain-based privacy protection mechanisms for online social networks.

4.1 Decentralized Protection of Personal Data

The most widely used online social networks are centralized services that are controlled and managed by single-large corporation. This allows a single entity to collect and control an unprecedented massive amount of personal and sensitive data of users from across the globe. Such an unprecedented collection of personal data constitutes a major threat to users' privacy and to their rights to be left alone. Moreover, OSNs users have lost control of what happens with their data afterward and they cannot withdraw permission in the current privacy setting. To address the issue of privacy, the decentralized online social network architectures were proposed [5, 35]. Decentralization has been considered as the panacea to privacy issues, especially in the realms of online social networks [5]. Major privacy concerns related to the centralized model are addressed by properly achieving decentralization. However, research reveals that decentralization architecture brings new technical challenges such as control and coordination, reliability and authenticity, etc. With the advent of blockchain technology, most of these issues can be addressed with inherent features of the blockchain such as immutability, transparency, and peer-2-peer (P2P) consensus.

Blockchain is designed to operate without the need for a central authority. The validation of transactions is done through peer-based consensus which is suitable for authentication of ownership rights as the history of all transactions is validated. Some of the advantages of using blockchain-based solution for OSNs are control over the ledger is distributed across many mining nodes, therefore, no one can monopolize the network and reduces the dependence on centralized service providers. P2P consensus mechanism ensures data integrity and manipulation of any information in the distributed ledger is rendered practically impossible. Monitoring and surveillance are much harder to achieve in blockchain-based solutions for OSNs because social communication is peer to peer without involving a third party controlling process. Blockchain-based online social networks intend to promote individual privacy and data sovereignty if properly implemented. Using blockchain-based solutions users remain in full control of their personal data. Blockchain offers the potential to provide a truly open and free service architecture for social communication services where users are not locked into any distributed service maintained in a centralized fashion. Adaption of decentralized OSNs was slow due to their federated social communication architecture which locked the users into their service platform that hampered a truly open and heterogeneous ecosystem for social communication.

Some of the earliest blockchain-based solutions for decentralized online social networks are Ushare and Tawki. Ushare [9] is a blockchain-based solution for user-controlled social media. It is a user-centric blockchain supported social media network that enables users to control, trace and claim ownership of every piece of content they share. The proposed solution consists of: a hash table with encrypted content shared by a user, a system for controlling the maximum number of shares performed by the users circle members, a local personal certificate authority (PCA) that manages the users circles and the Blockchain. However, this study presents only conceptual design and technical details on how to develop the platform are missing. Tawki [33] is a decentralized service architecture for social communication proposed by Westerkamp et al. Tawki that allows users full control of their personal data which is stored and managed by personal data storage. Tawki data storage is implemented using common REST-based API which facilitates users to send and request data to and from other users' personal data storage. Tawki uses the Ethereum blockchain to manage user identities.

4.2 Enabling privacy through Transparency

Online social networks collect a huge amount of personally identifiable information of the users. There is implicit trust by the users that OSNs will not misuse their sensitive personal information. The process of information usage by these web systems is fairly complex and users are not completely aware of what is happening with their data. We have previously described a recent scandal about the misuse of users' data from the social web. This raises the serious need for information accountability where appropriate use of the personally identifiable information can be determined after reviewing the usage pattern from audit logs. Furthermore, these audit logs can be used to check compliance with user's usage restrictions that assert no unauthorized data usage has taken place and also enable OSN service providers to be transparent with regards to data usage. Thus, we stress on implementing transparency to achieve information accountability with provenance mechanisms. Blockchain technology is regarded as a tamper-evident database that comprises a log of all transactions with transparency as an inherent feature. The transparency of blockchain provides a greater degree of control to end-users, who no longer need to trust OSN service providers with opaque data processing and usage mechanisms. According to [33], the blockchain-based solution provides data sovereignty by enhancing user control over personal data. Enabling transparency in the social web systems is a necessity to assert data ownership and privacy of users. Nissenbaum [23] introduces the notion of contextual integrity as a new benchmark for privacy. Transparency plays a pivotal role to achieve better contextual integrity. Seneviratne [26] also argues that transparency is a key component in achieving privacy and compliance.

4.3 Tamper Evident Ledger based Regulatory Compliance

With the enforcement of GDPR, OSN providers are obliged to comply with this new regulation. At present, they are not ready to comply or even had major gaps in GDPR-compliance. Non-compliance with this regulation imposes hefty fines apart from other legal issues while operating in the EU or processing the personal data of EU citizens. Thus, GPDR compliance is the only way forward to operate in the EU or process the data of EU citizens. GDPR imposes great challenges for the OSN providers concerning their current business model. Such an obligation is the recognition of users' consent as a legitimate ground for their personal data processing. Current OSN providers lack the GDPR compliant consent management model which means OSN users have a lack of control over their personal data. GDPR promotes the notion of consent to provide users more control over their personal data. Achieving consent compliance is not an easy task in online social networks. However, blockchain technology offers features that can add value to the consent management model for the processing of personal data in the context of OSNs. It aims to facilitate users to assert their rights and get bigger control over their personal data.

5 Conclusion and Future Work

Most of the online social networks lack an effective consent management mechanism that is compliant with the newly enforced GDPR. In the absence of such a mechanism, data subjects lose control over their personal data that poses a serious privacy problem. GDPR compliant consent management is an important step towards protecting user privacy in online social networks. In this paper, we presented some of the opportunities for using blockchain technology to address this issue. Blockchain technology provides certain features that offer OSNs users fine-grained control over their personal data. We also took into consideration the challenges imposed by new EU regulations (GDPR) for OSN providers and its apparent conflicts with blockchain technology. The paper opens up new research directions to be explored. In the future, we intend to develop a proof of concept prototype for blockchain-based GDPR compliant consent management model for online social networks.

Acknowledgement

This work was carried out at the department of information security and communication technology, Norwegian University of Science and Technology, Gjovik, Norway during the tenure of an ERCIM Alain Bensoussan Fellowship Programme.

References

1. Javed Ahmed. Privacy in online social networks: An ontological model for self-presentation. In *International Conference on Knowledge Engineering and the Semantic Web*, pages 56–70. Springer, 2016.
2. Javed Ahmed, Guido Governatori, Leendert WN van der Torre, and Serena Villata. Social interaction based audience segregation for online social networks. In *ECSI*, pages 186–197, 2014.
3. Javed Ahmed and Zubair Ahmed Shaikh. Privacy issues in social networking platforms: comparative study of facebook developers platform and opensocial. In *International Conference on Computer Networks and Information Technology*, pages 179–183. IEEE, 2011.
4. Fatemeh Alizadeh, Timo Jakobi, Jens Boldt, and Gunnar Stevens. Gdpr-reality check on the right to access data: Claiming and investigating personally identifiable data from companies. In *Proceedings of Mensch und Computer 2019*, pages 811–814. ACM, 2019.
5. Leila Bahri, Barbara Carminati, and Elena Ferrari. Decentralized privacy preserving services for online social networks. *Online Social Networks and Media*, 6:18–25, 2018.
6. Aurelie Bayle, Mirko Koscina, David Manset, and Octavio Perez-Kempner. When blockchain meets the right to be forgotten: Technology versus law in the healthcare industry. In *2018 IEEE/WIC/ACM International Conference on Web Intelligence (WI)*, pages 788–792. IEEE, 2018.
7. Robert M Bond, Christopher J Fariss, Jason J Jones, Adam DI Kramer, Cameron Marlow, Jaime E Settle, and James H Fowler. A 61-million-person experiment in social influence and political mobilization. *Nature*, 489(7415):295, 2012.
8. Danah M Boyd and Nicole B Ellison. Social network sites: Definition, history, and scholarship. *Journal of computer-mediated Communication*, 13(1):210–230, 2007.
9. Antorweep Chakravorty and Chunming Rong. Ushare: user controlled social media based on blockchain. In *Proceedings of the 11th international conference on ubiquitous information management and communication*, page 99. ACM, 2017.
10. Yize Chen, Quanlai Li, and Hao Wang. Towards trusted social networks with blockchain technology. *arXiv preprint arXiv:1801.02796*, 2018.
11. Sourya Joyee De and Abdessamad Imine. On consent in online social networks: Privacy impacts and research directions (short paper). In *International Conference on Risks and Security of Internet and Systems*, pages 128–135. Springer, 2018.
12. Simon Farshid, Andreas Reitz, and Peter Roßbach. Design of a forgetting blockchain: A possible way to accomplish gdpr compatibility. In *Proceedings of the 52nd Hawaii International Conference on System Sciences*, 2019.
13. Qi Feng, Debiao He, Sherali Zeadally, Muhammad Khurram Khan, and Neeraj Kumar. A survey on privacy protection in blockchain system. *Journal of Network and Computer Applications*, 2018.
14. Michèle Finck. Blockchains and data protection in the european union. *Eur. Data Prot. L. Rev.*, 4:17, 2018.
15. Ralph Gross and Alessandro Acquisti. Information revelation and privacy in online social networks. In *Proceedings of the 2005 ACM workshop on Privacy in the electronic society*, pages 71–80. ACM, 2005.
16. Ai Ho, Abdou Maiga, and Esma Aïmeur. Privacy protection issues in social networking sites. In *2009 IEEE/ACS International Conference on Computer Systems and Applications*, pages 271–278. IEEE, 2009.

17. Luis-Daniel Ibáñez, Kieron O'Hara, and Elena Simperl. On blockchains and the general data protection regulation. 2018.
18. Mike Isaac. Facebook security breach exposes accounts of 50 million users. 2018.
19. Adam DI Kramer, Jamie E Guillory, and Jeffrey T Hancock. Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences*, 111(24):8788–8790, 2014.
20. Christopher Millard. Blockchain and law: Incompatible codes? *Computer Law & Security Review*, 34(4):843–846, 2018.
21. Carlos Molina-Jimenez, Ioannis Sfyrakis, Ellis Solaiman, Irene Ng, Meng Weng Wong, Alexis Chun, and Jon Crowcroft. Implementation of smart contracts using hybrid architectures with on and off–blockchain components. In *2018 IEEE 8th International Symposium on Cloud and Service Computing (SC2)*, pages 83–90. IEEE, 2018.
22. Satoshi Nakamoto. Bitcoin: A peer-to-peer electronic cash system (white paper). *Retrieved from (28-05-2019): https://bitcoin. org/bitcoin. pdf*, 2008.
23. Helen Nissenbaum. Privacy as contextual integrity. *Wash. L. Rev.*, 79:119, 2004.
24. Md Mehedi Hassan ONIK, KIM Chul-Soo, and YANG Jinhong. Personal data privacy challenges of the fourth industrial revolution. In *2019 21st International Conference on Advanced Communication Technology (ICACT)*, pages 635–638. IEEE, 2019.
25. Dong Qin, Chenxu Wang, and Yiming Jiang. Rpchain: a blockchain-based academic social networking service for credible reputation building. In *International Conference on Blockchain*, pages 183–198. Springer, 2018.
26. Oshani Seneviratne and Lalana Kagal. Enabling privacy through transparency. In *2014 Twelfth Annual International Conference on Privacy, Security and Trust*, pages 121–128. IEEE, 2014.
27. Kit Smith. 53 incredible facebook statistics and facts. 2019.
28. Christina Tikkinen-Piri, Anna Rohunen, and Jouni Markkula. Eu general data protection regulation: Changes and implications for personal data collecting companies. *Computer Law & Security Review*, 34(1):134–153, 2018.
29. FWJ van Geelkerken and K Konings. Using blockchain to strengthen the rights granted through the gdpr. In *Litteris et Artibus*, pages 458–461, 2017.
30. Paul Voigt and Axel Von dem Bussche. The eu general data protection regulation (gdpr). *A Practical Guide, 1st Ed., Cham: Springer International Publishing*, 2017.
31. Dejan Vujicic, Dijana Jagodic, and Sinisa Randic. Blockchain technology, bitcoin, and ethereum: A brief overview. In *2018 17th International Symposium INFOTEH-JAHORINA (INFOTEH)*, pages 1–6. IEEE, 2018.
32. Yunsen Wang and Alexander Kogan. Designing confidentiality-preserving blockchain-based transaction processing systems. *International Journal of Accounting Information Systems*, 30:1–18, 2018.
33. Martin Westerkamp, Sebastian Göndör, and Axel Küpper. Tawki: Towards self-sovereign social communication.
34. Dylan Yaga, Peter Mell, Nik Roby, and Karen Scarfone. Blockchain technology overview. *arXiv preprint arXiv:1906.11078*, 2019.
35. Ching-man Au Yeung, Ilaria Liccardi, Kanghao Lu, Oshani Seneviratne, and Tim Berners-Lee. Decentralization: The future of online social networking. In *W3C Workshop on the Future of Social Networking Position Papers*, volume 2, pages 2–7, 2009.