Putting Data Protection by Design on the Blockchain

Alexandra Giannopoulou*

The principle of data protection by design, as it is enshrined in article 25 of the GDPR, is difficult to apply in blockchains. This article will assess how the reliance on asymmetric encryption and other privacy enhancing technological architectures -necessary in a blockchain-based system- approach both user control and data protection by design compliance from the single scope of anonymization and unlinkability. Data subjects' rights, accountability, and the potential shortcomings of applied technological constraints are thus sidelined. Ultimately, this limited understanding of technological privacy, acts as a misguiding set of principles for technological co-regulation through standardisation in blockchains. The standardization of these choices without a holistic analysis of data protection by design imperatives could ultimately weaken the position of data subjects, whose trust in the technological protections of personal data might prove to be relatively misplaced.

Keywords: Anonymity | Blockchain | Data Protection by Design | Encryption | EU General Data Protection Regulation | Privacy

I. Introduction

In words from history, let us speak no more of faith in man, but bind him down from mischief by the chains of cryptography

Edward Snowden[1]

In November 2018, a report published by the European Parliament and entitled 'Report on Blockchain: A forward-looking trade policy' pointed out that 'blockchain technology can provide solutions for the data protection by design provision in the GDPR implementation on the basis of their common principles of ensuring secured and self-governed data'[2]. This position appears to be aligned with existing conclusions acknowledging the alliance of objectives between the technological architecture of blockchains and the principle of data protection by design[3].

Compliance of blockchain applications follows this premise: user control of personal data is both a fundamental principle of the GDPR and a feature of blockchain architecture. As observed by Finck[4], when

DOI: 10.21552/edpl/2021/3/7

* Postdoctoral Researcher, Faculty of Law, University of Amsterdam. For Correspondence : <a.giannopoulou@uva.nl>.
Acknowledgements : The Lab has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme under grant agreement No 759681. An earlier version of this Article was presented at the International Conference Tilting Perspectives 2019, organised by Tilburg University on 17th May 2019, and at the 4th European Privacy Law Scholars Conference (PLSC Europe) in Amsterdam on 24 October 2019. The paper was also discussed at the Journées du Centre Internet et Société 2020. I would like to thank Professor Tara Whalen, Professor Francesca Musiani, and dr. Ksenia Ermoshina, for their insightful comments and suggestions.

1 From Glenn Greenwald's book, *No Place to Hide: Edward Snowden, the NSA, and the U.S. Surveillance State*, (Metropolitan Books, 2014). This quote is paraphrasing Thomas Jefferson's famous quote: 'In questions of power then, let no more be heard of confidence in man but bind him down from mischief by the chains of the Constitution'. Interestingly, from the juxtaposition of the two quotes, emerges the decade-old conundrum of trusting the system and/or trusting code.

2 European Parliament, 'Report on Blockchain: a Forward-Looking Trade Policy (AB-0407/2018)' (2018) para 14.

3 As highlighted by Finck, 'blockchains, if adequately designed, and the GDPR can share a common objective: giving a data subject more control over her data'. Michèle Finck, 'Blockchain and the General Data Protection Regulation: Can distributed ledgers be squared with European data protection law?' (2019) Study for the European Parliament.

4 ibid.

control is understood as individual agency over personal data, blockchains are marketed as an ideal technology to achieve this objective. Notably, the cryptographic elements introduced to blockchains, fundamental to its functions, coincide with 'technical measures' that are used to ensure data protection. Thus, a risk assessment of this technological design would tend to be regarded favourably[5] with regard to compliance, due to these privacy-enhancing technological features.

This article challenges this initial premise by further investigating the interplay of the objectives that guide both the GDPR and blockchain technologies. It highlights and examines the conflicts of data protection by design, as implemented by blockchain systems, with the privacy guarantees and protections provided by EU law. The orthogonal relationship between transparency and privacy is overarching in blockchain-based systems, which use transparency as a 'beacon' that ensures consensus over the recorded data. However, transparency is a fundamental principle in data protection regulation. In that regard, it is not considered as a countervailing force to (state) surveillance, but an empowering tool in the hands of data subjects. Blockchain-based systems and their corresponding communities have assimilated data protection as a framework of tools depicting the concepts of data minimalization, unlinkability, and confidentiality. As explained in the following analysis, data protection by design is only partially conveyed by these concepts as it requires the complementarity of data subjects' rights, transparency, and accountability.

Historically, distributed ledgers largely consist of a combination of pre-existing tools that was first put together by Satoshi Nakamoto. Although the Bitcoin white paper[6] does not make any reference to blockchains[7], their potential has been largely discussed and, rather often, hyped. While there is no consensus on a comprehensive definition of a blockchain[8], it can be broadly described as a distributed database that is shared between a network of computers, and which uses a consensus mechanism to validate updates. Specifically, it is considered ideal for reaching consensus among a network of peers that do not necessarily trust each other[9].

The explosion of blockchain-based business models across Europe coincided with General Data Protection Regulation's[10] coming into force. Meanwhile, the data protection mechanisms embodied in it, introduced a number of pressing questions in relation to emerging technologies. Namely, while the technological neutrality principle of the instrument promotes the applicability of the GDPR across the technological spectrum, multiple points of friction have been identified with blockchain technology[11]. For instance, the distributed architecture of the data processing, the transnational nature of such processing, the append-only feature of the database, and the (relative) anonymity of the actors are all significant points of friction/incompatibility between the technology and the legal instrument.

At the same time, blockchain-based systems are also regarded as a regulatory tool that could be used to potentially achieve GDPR objectives[12] and digital sovereignty as it is expressed through individual self determination claims, emphasizing "the autonomy of citizens in their roles as employees, consumers, and users of digital technologies and services"[13]. For example, the design and the technological components of blockchains are purportedly enabling the creation of (personal and non-personal) data sharing models[14]. Whether blockchain-based systems can enable the sharing of data between actors that do not

5 For instance, consider the Data Protection Impact Assessment performed by a data controller deploying a blockchain-based technology product: would it be accurate to consider this product 'low-risk' given the *a priori* fundamental reliance of blockchains on encryption?

6 Satoshi Nakamoto, 'Bitcoin: A peer to peer electronic cash system' (2009) <https://bitcoin.org/bitcoin.pdf> accessed 15 April 2021.

7 The terms will be used interchangeably in the article.

8 Toni Caradonna, 'Blockchain and society' (2020) 43 Informatik Spektrum, 40–52; See María-Cruz Valiente and Florian Tschorsch, 'Blockchain-based technologies' (2021) Internet Policy Review <https://policyreview.info/open-abstracts/blockchain-based-technologies> accessed 15 April 2021.

9 Deborah Ginsberg, 'The building blocks of the blockchain' (2019) 20 North Carolina Journal of Law & Technology 4, 471-491

10 Hereinafter GDPR.

11 See Finck (n 3).

12 There are also claims that blockchains should supersede the GDPR as a superior data protection mechanism. See Andrea Tinianow, 'GDPR isn't the Answer, but Blockchain is', Forbes (4 June 2018) <https://www.forbes.com/sites/andreatinianow/2018/06/04/gdpr-isnt-the-answer-but-blockchain-is/> accessed 15 April 2021.
For a critical approach to the empowering nature of blockchains, see Robert Herian, Blockchain, 'GDPR, and fantasies of data sovereignty, Law, Innovation and Technology' (2020) Law, Innovation and Technology 12, 156-174.

13 Jürgen Pohle and Thorsten Thiel, 'Digital Sovereignty' (2020) 9 Internet Policy Review 4.

14 European Commission, 'Commission Staff Working Document on the free flow of data and emerging issues of the European Data Economy', (2017) SWD, 2 final 13.

trust each other, do not want or need third-party intermediaries, and/or do not necessarily want to provide full access or control for their respective databases[15] has yet to be proven.

The *a priori* inclusion of encryption methods within the architecture of the technology is a significant milestone signalling the integration of data protection mechanisms in the design of the system[16]. At the same time, the prioritization of these methods at the -oft occurring- detriment of other data-protection-enhancing technological and organisational models in blockchain-based systems creates an axiomatic claim of privacy that fails to convince of its attachment to its ideals of data protection, privacy, and even transparency.

II. Setting the Stage: Data Protection by Design Meets Blockchain-Based Systems

When techno-legal regulatory provisions, such as the data protection by design principle, meet a technological tool, such as blockchain-based systems, carrying a rather charged ideological background which prioritizes —in its design— resistance to any centralized control, including that of the state, compliance becomes burdensome. Article 25 GDPR encompasses both data subjects' rights and accountable actors' obligations. It has been succinctly pointed out that 'the weight of the entire Regulation was put on the shoulders of Article 25'[17], because this overarching principle —addressed to accountable actors— aims to enforce all GDPR rationales and obligations in the technological architecture. Thus, risk-based compliance exercises should take into account the full spectrum of rationales that are used to prioritize design choices over other, and that guide the final technological architecture.

1. Data Protection by Design in Article 25(1) GDPR

The GDPR guarantees a high-level personal data protection in an increasingly complex datafied society. It offers individuals transparency, mechanisms to control the processing of their data, rights pertained to their data— all while imposing a range of obligations and responsibilities on entities and actors determining the purposes and means of the processing of personal data. The accountability principle becomes foundational for the Regulation, since it defines the responsible actors that would be obligated to justify their decision-making, and face consequences for not complying with these legal imperatives. Article 25(1) GDPR addresses the concept of accountability by mandating that responsible actors put in place a system wherein all GDPR principles, rights, and obligations are (demonstrably) reflected.

Data controllers are key in translating data protection rules and facilitating rights in practice. For example, according to article 25 GDPR, data controllers would have to ensure the implementation of data protection principles (Article 5 GDPR), and are responsible to ensure that data subjects have the possibility to exercise their rights, as they are enumerated in the GDPR. As highlighted by the European Data Protection Board (EDPB),

> 'the requirement is for controllers to have data protection designed into and as a default setting in the processing of personal data. [This] means that controllers must be able to demonstrate that they have in place the appropriate measures and safeguards in the processing to ensure that the data protection principles and the rights and freedoms of data subjects are effective'[18].

The concept "privacy by design[19] refers to the approach that aims at addressing privacy and data protection issues by embedding legal rules, values, and principles in Information and Communication Tech-

15 Currently there are multiple initiatives that want to benefit and expand from this feature to create new data sharing markets in the European Union: See for example the Ocean Protocol. Similarly, projects related to digital identity and to health data are increasingly relying on DLTs.

16 Gerald Spindler and Phillip Schmechel, 'Personal Data and Encryption in the European General Data Protection Regulation' (2016) 7 JIPITEC 2, 163-177.

17 Lina Jasmontaite, Irene Kamara, Gabriela Zanfir-Fortuna and Stefano Leucci, 'Data protection by design and by default: Framing guiding principles into legal obligations in the GDPR' (2018) 4 European Data Protection Law Review 2, 168-189.

18 European Data Protection Board, Guidelines 4/2019 on Article 25 Data Protection by Design and by Default, 13 November 2019.

19 According to one of the first advocates of the concept, Ann Cavoukian, the seven guiding principles of privacy by design are: 1. Proactive not reactive, Preventative, not Remedial, 2. Privacy as the default, 3. Privacy Embedded into Design, 4. Full functionality - Positive Sum not Zero Sum, 5. End-to-end security - Lifecyle Protection, 6. Visibility and Transparency, 7. Respect for User Privacy.

nologies (ICTs) requirements and in business policies and practices"[20]. This concept does not include guidelines nor a checklist of obligations as a guiding vector for data controllers, but rather a set of guiding principles aimed at being translated in different data processing contexts[21]. In the GDPR, the same open-ended objective remains. The data protection by design principle, as implemented in article 25(1) GDPR, is a somewhat lengthy text with a high-level description of data controllers' obligations, coupled with guidelines and opinions from responsible bodies —these also abstain from providing a closed set of rules.

The data controllers' foundational obligation would appear as a

> 'qualified duty [...] to put in place technical and organizational measures that are designed to implement data protection principles effectively and to integrate necessary safeguards into the processing of personal data so that such processing will meet the Regulation's requirements and otherwise ensure protection of data subjects' rights'[22].

These 'technical and organizational measures' that they are required to take, are far from a set data protection methodological checklist. On the contrary, they represent the multifaceted obligations that the accountability principle entails for data controllers[23], which can be embodied through standard-setting, certification mechanisms, and sector/technology-specific codes of conduct (article 25(3) GDPR).

Data protection by design —as enshrined in the GDPR— requires that these measures are taken towards protecting both privacy and personal data in a proportionate, balanced manner based on the risks identified by the responsible actors. However, when the priorities in the technical and structural measures taken focus only on data minimization and anonymization, it is *privacy as confidentiality* that becomes significantly highlighted within the process.

While these practices are substantial in ensuring data protection, the absence of a more holistic approach oftentimes leads to data subjects' rights being compromised[24]. This risk has been pointed out in the context of both privacy engineering and regulation[25]. How this approach affects blockchain data protection by design implementations will now be showcased.

By now, it is almost truism to repeat how translating data protection obligations in a decentralized technological environment lacking structured centralized governance dynamics and carrying a rather charged principle-based framework, is presented with significant compliance risks. For this reason, it is becoming increasingly important to understand the sets of principles guiding the design choices of accountable actors before assessing how data protection by design fits into them.

2. Privacy Ideals in Blockchain-Based Systems

The overarching ideological considerations that guided the design of the first blockchain application, the cryptocurrency bitcoin, remain relevant in current blockchain projects in development as they highlight persisting conflicts between different conceptions of privacy and its relationship to transparency.

Firstly, the implementation of decentralization logics[26] created a unique *raison d'être* for blockchains. Decentralization is a design choice aiming to improve *privacy— and control—as—in censorship resistance*[27]. Censorship resistance through

20 Giorgia Bincoletto, 'EDPB Guidelines 4/2019 on Data Protection by Design and by Default' (2020) 6 EDPL 4, 574.

21 Seda Gurses et al, 'Engineering Privacy by Design' (2011) Computers, Privacy & Data Protection, 25.

22 Lee Bygrave, 'Data Protection by Design and by Default: Deciphering the EU's Legislative Requirements' (2017) 2 Oslo Law Review 4.

23 ibid.

24 It has been pointed out that 'anonymization can be used to disempower data subjects': Jeff Ausloos et al, 'Getting Data Subject Rights Right: A Submission to the European Data Protection Board from International Data Rights Academics, to Inform Regulatory Guidance' (2019) 10 JIPITEC 283.

25 (n 21); Michael Veale et al, 'When data protection by design and data subject rights clash' (2018) 8 International Data Privacy Law 2, 105-123.

26 Balázs Bodó and Alexandra Giannopoulou, 'The Logics of Technology Decentralization: the Case of Distributed Ledger Technologies' in Massimo Ragnedda and Giuseppe Destefanis (Eds.), *Blockchain and Web 3.0: Social, Economic, and Technological Challenges* (Routledge, 2019).

27 The preservation of anonymity relied in great part to decentralizing the database that holds the information risking the reidentification of individuals. 'A landmark raid on an anonymizing Usenet system in Finland, anon.penet.fi, was conducted by Interpol at the behest of the Church of Scientology] seeking the identity of a particular leaker. The Finnish Internet technologist Johan Helsingius, who ran the remailer, warned at the outset of his project: 'Well, if the police or the local Secret Service comes knocking at my door, with a court order to hand over the database, I might comply'. But what was the alternative?'; Finn Brunton, *Digital Cash* (Princeton University Press, 2019) 94.

decentralized design complements the anonymity features that are embedded in the system.

Secondly, blockchains have incorporated the foundational principles that guided the lineage of innovation envisaging the creation of digital cash. The early 'code-rebels'[28] brought forward the idea that encryption could be popularized beyond government-issued programs of the time, for all types of communication, as a tool for individual empowerment with the potential to challenge existing power structures. These 'crypto-rebels' viewed digital privacy, as 'the ability of people to communicate without fear of the government, and what they wanted, once they started to think hard about the problem, was a means of encrypting private communication'[29].

The history of cryptography (for example the evolution from symmetrical to asymmetrical encryption tools[30]), is paved by efforts to minimize vulnerability points and to make it too costly or too time-consuming for an adversary to invade the private communication between the parties involved[31]. Thus, confidentiality of communications meant minimizing the risk of deciphering encryption mechanisms. This constitutes a *technical* approach to privacy as confidentiality. Narayanan argues that the 'crypto dream'[32] can be further divided between *crypto for security* and *crypto for privacy*. While the first refers to the security of transactions, the second points to the security of communication free of government surveillance and control.

The cypherpunk movement emerged from among the first 'crypto-rebels'. The objective of this movement was to create a world free from corporate and governmental control, using the opportunity created by the Internet to enable direct un-censorable communications. They came together during the '80s in order to explore how 'the nexus between cryptography and politics' could be used to protect 'individual autonomy threatened by power'[33], and to develop technical solutions that would embody these values. This movement saw 'the Internet as proof that even the heavy hand of the state had to give way to the laws of mathematics underlying cryptography and the software engineering underlying packet-switched data networks'[34].

The vision of constructing a *secure* society –free from state surveillance, government control, and corporate constraints- was a major factor in researching independent currencies, and economic models as an alternative to state-powered ones. Electronic money was seen as a 'control apparatus'[35] serving the goals of centralized power structures. Thus, a new type of currency would have to be created, a form of digital cash that remained private and secure like physical money[36].

From a technological perspective, the design of such currency used cryptographical tools to ensure both anonymity in transactions and the trustworthiness of the system against adversaries. David Chaum was the first cryptographer to explore applying cryptographic features to cash, arguing that 'computerization is robbing individuals of the ability to monitor and control the ways information about them is used'[37]. He succeeded in designing a system that would permit anonymous transactions. Tim May[38] and the Extropian group[39] also envisioned anonymity in digital cash and transactions.

28 Steven Levy, *Crypto: How the Code Rebels Beat the Government, Saving Privacy in the Digital Age* (2002, Penguin Putnam). According to Narayanan, these 'code rebels' were 'a loose coalition of academics, hobbyists, and civil-liberties organizations; Arvind Narayanan, 'What happened to the crypto dream? Part 1' (2013) 11 IN IEEE Security and Privacy Magazine 2, 75–76.

29 Thomas Powers, 'Notes from the Underground' (2001) 48 New York Review of Books 10, 51-54.

30 The symmetry in the original encryption tools added multiple vulnerability points between the sender and the receiver during the communication.

31 According to the famous story of the development of asymmetric (public key) encryption, the cryptographer Whitfield Diffie visited researchers across the country to get their point of view on two questions: how to reliably verify ourselves and our machines and how to communicate with provable secrecy. A similar breakthrough -although called "non-secret encryption"- was made by James Ellis, Clifford Cocks, and Malcolm Williamson in the UK.

32 (n 29).

33 Philip Rogaway, 'The moral character of cryptographic work' (2015) Cryptology ePrint Archive, Report 2015/1162.

34 Kevin Werbach, *The Blockchain and the New Architecture of Trust* (MIT Press, 2019), 30.

35 (n 27) 51.

36 James Bridle in Jaya Klara Brekke, *The White Paper by Satoshi Nakamoto with a guide by Jaya Klara Brekke*, (Ignota, 2019); Quinn DuPont, *Cryptocurrencies and Blockchains* (Polity Press, 2019).

37 David Chaum, 'Security without identification: transaction systems to make big brother obsolete' (1985) 28 Communications of the ACM 10, 1030-1044.

38 Tim May, 'Untraceable Digital Cash, Information Markets, and BlackNet' (1997), <http://osaka.law.miami.edu/~froomkin/articles/tcmay.htm> accessed on 15 April 2021.

39 Extropians wanted to accelerate progress through technology as much as possible with the goal of extending human life. They envisioned the current financial model a hurdle to that progress and wanted thus, to reinvent it.

Anonymity in cryptocurrencies has been a contentious issue, since it has been simultaneously over- and under- valued. Principally, 'given the history of privacy as a primary motivation for the adoption of digital cash, we surmise that many of the cryptocurrency adopters (other than speculators) are primarily seeking privacy, whether to circumvent capital controls or just to avoid the pastoral gaze of state or corporate surveillance'[40]. However, anonymity as a property of the cryptocurrency protocol, is frequently distinguished from real-world anonymity[41]. For example, transactions using third party actors, coupled with added regulatory compliance obligations, such as Know your customer rules, can link real-world identities with cryptocurrency transacting addresses or accounts.

Within the same technological protocol, the computational processes deployed for money creation[42] can be considered in the context of anonymity and data protection[43]. The 'crypto' in cryptocurrency refers not to 'encryption' as a secure and confidential feature, but to the cryptographic methods applied in making the system function the way it does[44].

Overall, the implementation of anonymity and transparency in the technological design protocol often highlights the conflation between legal anonymity and technical anonymity. Firstly, in technical terms, anonymity refers to pseudonymity together with unlinkability. Narayanan explains that

> '[u]nlinkability is a property that is defined with respect to the capabilities of a specific adversary. Intuitively, unlinkability means that if a user interacts with the system repeatedly, these different interactions should not be able to be tied to each other from the point of view of the adversary in consideration'[45].

Secondly, in legal terms and according to Recital 26 GDPR, anonymity refers to information that cannot be related to a natural person, or that is no longer *reasonably likely* to be attributed to a natural person. Thus, the legal concept of anonymity, as is the case for personal data, is dynamic and context-dependent[46]. The regulator adopted the versatility found in the concept of personal data to fit the concept of anonymisation. For example, the outcome of legal anonymisation is likely to change over time as the context and circumstances of the identification efforts required might shift.

Based on these two approaches, Nakamoto's promise for transactional anonymity is upheld because the only requirement for transacting with bitcoin is a newly generated cryptographic public key or its hash. In that sense, bitcoin transacting accounts could remain technically anonymous. Nevertheless, de-anonymization processes have been achieved with the help of available transactional data on the blockchain[47] even without the intermediation of third-party actors operating under special regulatory identification obligations. For instance, transactions appear susceptible to reidentification through data mining on public ledger transactional data, or side channel attacks, such as triangulating anonymous transactions with IP addresses, transaction times, and other less anonymous data. This is where the early discourse on the conflicts between privacy and public permissionless blockchains[48] focused, namely, on the privacy versus transparency paradigm. So, on the one hand, the application of various encryption techniques ensures the confidentiality of the transacting parties and on the other, the decentralized architecture of the public network is what ensures transparency.

These transactions are integrated in the public record of the chain of prior transactions. Brunton observes that 'humans may try to conceal themselves,

40 Geoff Goodell and Tomaso Aste, 'Can Cryptocurrencies Preserve Privacy and Comply with Regulations?' (2018) Frontiers in Blockchain.

41 Daniel Genkin, Damitrios Papadopoulos and Charalampos Papamanthou, 'Privacy in Decentralized Cryptocurrencies' (2018) 61 Communications of the ACM 6, 78-88.

42 These computational processes include ensuring fault tolerance through proof-of-work (for Bitcoin) and the byzantine fault tolerance to design the decentralized architecture. The famous double spending problem was the most burdensome hurdle that the cypherpunks and researchers had to address during the conception of new digital currencies.

43 'Participants can be anonymous', writes Satoshi Nakamoto at the Bitcoin white paper.

44 (n 27) 159; (n 37) 41.

45 Arvind Narayanan et al, *Bitcoin and Cryptocurrency Technologies: A Comprehensive Introduction*, (Princeton University Press, 2016) 166.

46 Michèle Finck and Frank Pallas, 'They who must not be identified—distinguishing personal from non-personal data under the GDPR' (2020) 10 International Data Privacy Law 1, 11–36.

47 (n 37).

48 Primavera de Filippi, 'The interplay between decentralization and privacy: the case of blockchain technologies' (2016) Journal of Peer Production, 7.

but their money has an identity, and it never forgets'[49]. The system is set up in such a way that while the asset is not a priori secret, the identity of the transacting parties remains hidden. Thus, as argued by Bridle, the greatest outcome from the deployment of the bitcoin blockchain is the establishment of a persistent but yet private identity. 'The real 'product' of Bitcoin [is] a decentralized, deniable identity'[50] that resembles the cypherpunk concept of *nyms*[51].

III. Standardising Compliance: Data Protection-Aware Blockchains

While the concept of privacy as confidentiality is central in the portrayal of decentralization as a tool against surveillance from the state, the concept privacy as control is regularly viewed as the most substantial value among decentralized networks enthusiasts and crypto-communities. However, the control envisaged in this concept of privacy differs from the control as a core principle of the GDPR[52]. The former describes control as 'designed to enable people to more actively decide when and with whom to share their own personal information'[53] while the latter translates data subjects' control in a set of rights and accountability measures for liable actors[54].

With effectiveness being 'at the heart of the data protection by design'[55] rule, data controllers are required to implement all GDPR principles and to accommodate data subjects' rights. Thus, 'accountability provides further means to check what happens on the side of the controller when the data has been released and therefore to move from blind trust to proven trust'[56].

While the a priori control, enforced through privacy enhancing technologies, is usually found in blockchains, the ex post *proven* trust finds little conceptual compatibility within the current state of the technology[57]. This technological development attempts to guide the techno-legal framing of the data protection by design compliance for blockchains in general. It risks to lead to the application of a single aspect of data protection, a 'monoculture', that promotes and highlights a priori attempts in deanonymization but ignores the rather substantial questions related to the accountability of actors and individual empowerment through data subjects' rights.

1. Technology to the Rescue or Privacy Enhancing Technologies

The wording of article 25 GDPR, requiring that data controllers take 'appropriate technical and organizational measures' to ensure that data protection rules[58], leaves ample rule for the necessary flexibility in the specific choices that they are making towards compliance.

The contextual and dynamic nature of the principle[59] creates the conditions for the designers of the system to identify the relevant risks involved in the data processing at hand, and to introduce the appropriate safeguards. Ultimately, the rule remains true to the objective of technological neutrality.

Ensuring data protection by design in blockchains, involves both the essential building block of the technology –cryptography- and further privacy-enhancing technologies (PETs) built on top of the core service ledgers as added elements. The use of encryption techniques as central features to the design of blockchains, would make them appear in compliance with part of the data protection by design obligations, since encryption is particularly underlined in article 25(1) GDPR. However, this represents a rather limited view of the concept, since adopted appropriate measures should also 'integrate necessary safeguards' to protect the rights of the data subjects. For

49 (n 27) 163.

50 (n 37).

51 Nick Szabo introduced the concept in order to underline the vulnerabilities to which our identities are exposed when too much unrelated information is linked together. 'As in magic, knowing a true name can confer tremendous power to one's enemies'. Nick Szabo, Smart Contracts Glossary (1995) <http://www.fon.hum.uva.nl/rob/Courses/InformationInSpeech/CDROM/Literature/LOTwinterschool2006/szabo.best.vwh.net/smart_contracts_glossary.html> accessed 15 April 2021.

52 Robert Herian, *Regulating blockchain: Critical perspectives in law and technology* (Routledge, 2018).

53 (n 48).

54 Jef Ausloos, *The Right to Erasure in EU Data Protection Law: From Individual Rights to Effective Protection* (Oxford University Press, 2020).

55 (n 18).

56 Denis Butin et al, 'Strong Accountability: Beyond Vague Promises' in Serge Gutwirth et al (eds.) *Reloading Data Protection: Multidisciplinary Insights and Contemporary Challenges* (Springer, 2014) 343-369.

57 (n 35).

58 See supra.

59 Mireille Hildebrandt and Laura Tielemans, 'Data protection by design and technology neutral law' (2013) 29 Computer, Law & Security Review 5, 509-521.

example, while responsible intermediaries (ie online wallets, exchanges etc) could employ best practices as accountable data controllers, data protection by design compliance would also require making it virtually impossible for users to act differently, and thus implementing safeguards at the protocol level[60].

According to the risk-based approach adopted by the GDPR, controllers have to apply 'state of the art' technological developments. Namely, they could exercise 'discretion to choose which available measures in the market are the best suited (appropriate to the risk level) for their particular case of personal data processing'.

Historically, PETs were first developed as anonymity tools[61]. While these tools were fashioned strongly towards preserving unlinkability of data, concepts such as transparency enhancing or intervenability enhancing technologies[62] were also introduced, although not achieving similar popularity. Thus, in the high-risk environment of blockchains, the already developed field of PETs became the ideal domain on which GDPR compliance would be founded.

Privacy engineering has been increasingly involved in various blockchain projects. While decentralized architectures are historically considered to be beneficial to privacy, they do not necessarily provide inherent privacy guarantees. Blockchains are deployed with some fundamental trade-offs in terms of maintaining privacy, transparency, and honesty between different parties[63]. The interest spike in blockchains led to the continuous development of PETs for the purpose of better serving different decentralized blockchain projects[64]. The aims of these projects tend to align with PETs, whose primary focus is information disclosure within an adversary-led threat model[65]. Some PETs have been integral to the development of specific blockchain projects or cryptocurrencies[66]. Overall, their use signals an ideological adherence to privacy as confidentiality protection, one that underlines anonymization as a data protection dominant 'monoculture'.

The adequacy of the above-stated mechanism to ensure data protection on a blockchain-based protocol layer requires two distinct forms of approval: one that stems from the decentralised network of decision-making and change-integrating actors and developers and another, from authorised communities creating GDPR-compliant standards.

2. Technological Co-Regulation: The Promises of Standardisation

The shift from command-and-control to co-regulatory approaches towards achieving a more efficient protection of personal data is best illustrated through the GDPR self-regulatory tools, such as codes of conduct, certification mechanisms, and standards. Frequently orchestrated by supervisory authorities and steered towards specific industries, these tools are often reflecting the forefront discussions and practices of techno-legal data protection[67].

The GDPR provides a general framework for certification mechanisms in articles 42 and 43. Conversely, there are additional certification mechanisms covering specific obligations in the GDPR; article 25(3) GDPR falls within that category as it introduces "approved certification mechanism(s) by which a controller can demonstrate compliance with Art. 25". So, compliance with data protection by design obligations would be demonstrable through the rolling out of certification mechanisms and standard setting.

It is truism to repeat how the creation of standards is challenging in a fast-paced technological environment. This has already been highlighted especially in the security technological field, where the rapid technological advancements means that "changes to standardisation are too slow to keep up"[68]. The end-

60 (n 37) 162.

61 John Borking and Charles Raab, 'Laws, PETs and other Technologies for Privacy Protection' (2001) 1 Journal of Information Law and Technology, 1.

62 George Danezis et al, 'Privacy and Data protection by Design-from policy to engineering' (2014), ENISA Report <https://www.enisa.europa.eu/publications/privacy-and-data-protection-by-design> accessed 15 April 2021.

63 Carmela Troncoso et al, 'Systematizing Decentralization and Privacy: Lessons from 15 Years of Research and Deployments' (2017) Proceedings on Privacy Enhancing Technologies, 4:307–329. The authors go on to point out that the tradeoffs should not be 'preordained', referring to PET applications on decentralized environments.

64 For an overview of privacy-related research on bitcoin see <https://en.bitcoin.it/wiki/Privacy> accessed 15 April 2020.

65 (n 26).

66 See for example cryptocurrencies such as Monero and Zcash.

67 See for instance, Irene Kamara et al, 'Data Protection Certification Mechanisms' Study on Articles 42 and 43 of the Regulation (EU)2016/679, (2019), Report, <https://ec.europa.eu/info/sites/info/files/data_protection_certification_mechanisms_study_final.pdf> accessed 15 April 2021.

68 Irene Kamara et al, 'Raising trust in security products and systems through standardisation and certification: The crisp approach' (2015) ITU Kaleidoscope: Trust in the Information Society (K-2015).

less game of *whack-a-mole* between standards and innovation is not a new criticism to standardisation efforts, nor to the overall co-regulatory approach towards new technologies.

Overall, in the fast-developing technological environment of blockchains, where many of the technical elements are in constant flux, compliance uncertainty is a fact. With little guidance on the efficiency of recently developed PETs, accountable data controllers (whose qualification in a decentralized network remains unclear) would look for technological standardization as a helpful GDPR toolbox. Recently, these processes have been evoked in a study conducted for the European Parliament[69].

From the developing sector of PETs applied on blockchains, two examples stand out because of their upcoming popularity in ensuring minimization of data disclosure. Firstly, zero knowledge proof applications and secondly, Schnorr signatures applications aim to safeguard transparency and anonymity while preserving trust.

a. Zero Knowledge Proof

The most prominent example of a privacy-enhancing technology in blockchains is zero knowledge proof. It is rather the only encryption method that has been specifically mentioned in a European Parliament report, as a means for blockchain projects to be able to comply with the data protection by design requirement. Namely, after pointing out the need to comply with data protection by design imperatives, the report explains that 'future blockchain applications should implement mechanisms that protect personal data and the privacy of users and ensure that data can be fully anonymous' by funding research on 'new blockchain technologies that are compatible with the GDPR and based on the principle of data protection by design, such as zk-SNARK (zero-knowledge succinct non-interactive arguments of knowledge)'[70].

This encryption method, first developed at the end of the 1980s, 'allows you to prove possession of a secret without actually revealing it. Moreover, the verifier of such a proof cannot convince anybody else of this fact'[71]. In essence, zero knowledge proofs could achieve both proportionate data minimization for transaction data put on chain, and verifiability[72].

While the implementation of this cryptographical method to distributed ledger technology is becoming progressively more commonplace, it is still far from being established as a technological standard that would be implemented in the bitcoin protocol[73]. Lately, a process has begun to technological zero knowledge proof standardization[74], but this has not been subject to any officialised standardisation yet–one that would render zero knowledge proofs, data protection by design compliant.

b. Schnorr Signatures

The second component that has been subject to community-wide formal approval is Schnorr signatures. Digital signatures are essential because they constitute a verification condition of the transactions recorded on the distributed network of nodes. However, due to the rapid growth of the bitcoin blockchain transaction history, the balancing between transparency and unlinkability is continuously under revision. Maintaining the concept of privacy as confidentiality as a guiding principle, the need for privacy-enhancing digital signatures has been highlighted from the community. For instance, the preservation of data disclosure minimisation is brought froward by a protocol amendment proposal[75] to replace existing digital signatures with Schnorr signatures.

These signatures are considered to be cryptographically more secure, and they have been intrinsically tied with multiple privacy preserving proposals. For example, the term 'scriptless scripts' has been used to denote a transparency and privacy preserving mechanism of executing transactions[76]. Ring signatures are, as Vitalik Buterin highlights, signatures proving 'that the signer has a private key correspond-

69 (n 3).

70 European Parliament (27 November 2018) Report on Blockchain: a Forward-Looking Trade Policy (AB-0407/2018) para 21.

71 (n 63).

72 (n 3).

73 The protocol use of zero knowledge proofs exists in other cryptocurrencies such as Zcash.

74 For instance, the ZKP community is leading a standardization process for the technology: <https://zkproof.org/>

75 BIP 340- Schnorr Signatures for secp256k1 <https://github.com/bitcoin/bips/blob/master/bip-0340.mediawiki> accessed 15 April 2021.

76 The term was first proposed and coined by mathematician Andrew Poelstra.

ing to one of a specific set of public keys, without revealing which one'[77].

These example applications have not yet been implemented in the bitcoin protocol. They represent a subset of anonymisation techniques which provide privacy while also minimizing technical or organizational trade-offs. Since they are not yet subject to standardisation processes neither by the informal governance of developer communities that would implement these changes on a network protocol level nor by formal standardisation bodies, the legal uncertainty as to whether these applications reach the GDPR-required anonymization threshold remains.

IV. Trust in Blockchain Technology

Blockchain design developments aim to inspire trust in continuous efforts to improve technological constraints that would effectively prevent reidentification. The objective does not appear to be the enhancement of individual control over personal data based on the meaning that GDPR instils on the concept but, rather, the creation of technologically-mediated trust[78] that would prioritize user security and 'anonymity'.

It is through this displacement of trust and the misconceptualization of data subject control that blockchains depart from the data protection by design principle. While both the technological environment of blockchains and the one envisioned by article 25 GDPR rely on technical measures (such as encryption techniques), the latter creates guarantees founded on the accountability of responsible actors[79], in contrast with blockchains, where the guarantee would lie in the robustness of the applied PET within its applicable context[80].

A salient feature of both blockchains and the GDPR is the concept of controlling data, as a means to protect oneself from bad faith actors or adversaries. According to this aspect of control, information disclosure is a prioritized risk in data processing. Thus, it could take priority over other personal data breaches or legal abuses, that would justify the data subjects' intervenability. For example, data anonymization cannot constitute a justification for refusing to respond to an access request, nor is it equivalent to the right to erasure[81]. While article 11 GDPR specifically exempts controllers from the obligation to conform to data rights (ie articles 15-20 GDPR), it also specifies that data subjects can provide additional information to the data controllers effectively facilitating their identification. This way, they can subsequently request the respect of any of the GDPR data rights. While considerable effort is put into highlighting the efforts to prevent reidentification, there is significantly less certainty in enforcement rules for keeping responsible actors accountable. Nourishing this privacy 'monoculture' could lead to a distortion of fundamental data protection principles such as that of transparency.

GDPR makes it clear in both articles 5 and 25 that the design of a data processing technology shall not only focus on information disclosure, but that it shall incorporate all data protection principles. However, as it has been already shown, multiple examples highlight the prioritization of de-identification techniques as part of the privacy-as-confidentiality concept[82]. Thus, the principle of anonymization is prioritized through the promotion of PETs. The popularity that this principle has gained in blockchains, because of anonymity being enshrined as an essential feature, creates a competition-like environment where de/re-identification technologies take priority over other transparency-enhancing measures that could accommodate data subjects' rights[83].

77 Vitalik Buterin, 'Privacy on the blockchain' <https://blog.ethereum.org/2016/01/15/privacy-on-the-blockchain/> accessed 15 April 2021.

78 Balázs Bodó, 'Mediated Trust – A Theoretical Framework to Address the Trustworthiness of Technological Trust Mediators' (2020) New Media and Society.

79 According to Quelle, 'the risk-based approach provides a way to carry out the shift to accountability that underlies much of the data protection reform, using the notion of risk as a reference point in light of which we can assess whether the organisational and technical measures taken by the controller offer a sufficient level of protection.'; Claudia Quelle, 'Enhancing Compliance under the General Data Protection Regulation: The Risky Upshot of the Accountability- and Risk-based Approach' (2018) 9 European Journal of Risk Regulation 3, 502-526.

80 There are also vulnerabilities born from the lack of privacy-awareness of the transacting parties (i.e. reuse of the same address in multiple transactions in the case of bitcoin).

81 This has been pointed out specifically for the early days of public blockchains, which tended to inscribe a considerable amount of personal data on the public ledger. Also, the existence of illegal data such as child pornography or revenge porn data showcases the relative nature of anonymization and the salient feature of actor accountability as a feature of data subject control.

82 (n 26).

83 The popularity of zero-knowledge proofs and the reliance of blockchain-based system actors on its promise to deliver data protection-compliant blockchains is a recent example of this competition.

The regulation-by-constraint model that inspires trust in the encryption technologies used in blockchains, are putting the weight of data protection in the choice of technological tools. Consequently, this progressively erodes control and the power of intervenability from data subjects, in favour of a secure, private, and trust-inspiring system.

The orthogonal relationship between trust and transparency is perfectly conveyed through the pervasive juxtaposition of privacy and transparency in the discourse surrounding the creation of decentralized blockchains. As argued by de Filippi, 'the more we shift towards a decentralized infrastructure, the less we need to rely on trust and the more we rely on transparency instead'[84]. The author showcases that the trust in entities is replaced not only by transparency aiming to keep all actors 'honest' but also by trust in the technological means of encryption involved in making the information private.

Nissenbaum has highlighted the paradoxical relationship between security and trust. She points out that, 'where people are guaranteed safety (...) trust is redundant; it is unnecessary. What we have is certainty, security, and safety- not trust'[85]. Admittedly, the 'trustless' feature of the blockchain tends to be inherently linked with the transparency feature, and not with the applied encryption methods.

As argued by Werbach, both encryption and data protection by design constitute in that sense 'mechanisms of trust'[86]. Namely, they tend to be restricting mechanisms that regulate people's actions towards the achievement of a specified objective.

As an accountability structure and enforcement mechanism, part of the *ratio legis* in the data protection by design principle is ensuring that the entity equipped with the most decision-making power during the data processing, (ie the data controller), takes effective data protection measures. The reliance and trust in the responsible data processing by data controllers is only created through the underlying liability regime. Overall, this is how the law would constrain these responsible actors into compliance.

Cryptography's constraints for maintaining trust among people is different. The obligation to apply 'appropriate technological and organizational measures' inscribed in the GDPR is founded in the accountability principle that the law instils in data controllers. Yet, the obligation to apply 'trusted' technological constraints for the effective protection of data subjects and personal data is inherent in this measure. It is through this design that the conceptual affinity becomes more apparent, since data subjects trust the data processing technological safeguards implemented by an undefined group of developers on one hand, and by the data controllers on the other.

V. Conclusion

This article showed that, fundamental technological tools employed in DLTs have been implemented both as a means of reaching an ideologically guided objective and of addressing the technical obstacles posed when trying to achieve it. Ultimately, the assessment of compliance with data protection by design cannot depend on 'check lists'. Rather, it should take into consideration a combination of factors that include what each tool's intended use within the architecture is, and what the ideological components that comprise it are.

While in data protection regulation the accountability of responsible actors is what produces trust, decentralized networks rely on technological design to inspire trust. The confidence that the transactions and personal data processed by DLTs are private (as in confidential), is produced by the technology at hand.

Well-intentioned as it may be, the noble goal of creating a system that relies on or even enables individual control of data, is not fulfilling its intended purpose. By instilling trust in the technology to ensure the protection of individuals, it fails to empower them with the armoury of rights that the data protection regulatory framework includes. Admittedly, 'hard-edged, cryptographically secured code, can never fully encompass human intentions'[87].

Naturally, reliance on encryption and PETs is a necessary predicament for data protection. However, the tools developed so far in line with this objective, do

84 (n 48).

85 Helen Nissenbaum, 'Will Security Enhance Trust Online, or Supplant It' in Roderick Kramer and Karen Cook (eds), *Trust And Distrust In Organizations: Dilemmas And Approaches* (Russell Sage, 2004) 173.

86 The same author points out how the trust-inspiring encryption technology is the foundation of all Internet communications and transaction. Werbach ties this 'cryptoregulation' model to the productions of encryption standards. See (n 35).

87 (n 35) 222.

not appear to translate the full spectrum of individual empowerment prerogatives from law into technological design. This failure is further accentuated by the lack of adoption of related trust-producing technological tools, that could finally align with the GDPR objectives.

PETs have been intrinsically linked to article 25 GDPR compliance. Nevertheless, these practices do not have fundamental independent value, nor are they intended to operate in a vacuum. Their implementation is dependent on a greater structure that permits compliance, one that includes accountability measures, liable actors, and governance safeguards for applying the relevant obligations.