

CYBER ETHICS OF BLOCKCHAIN, AI AND WORLDCHAIN

Dario Elias Félix de Oliveira Rodrigues
Instituto Politécnico de Santarém, Santarém, Portugal

ABSTRACT

This paper discusses the evolution of cyber ethics steered by blockchain technology (BT), artificial intelligence (AI), and, perhaps, Worldcoin, which is an ambitious global project to build a decentralized "proof-of-personhood" encrypted in an anonymous "World ID" for each human being. To investigate the digital conundrum and evaluate the ethical implications of BT, Explainable AI (XAI), and Worldcoin, the "Cyberethics-Mix" framework will be a compass oriented to four cardinal points of ethics in cyberspace: Privacy (data protection), Property (data ownership), Precision (data accuracy), and Pervasiveness (data access). The paper's main conclusion is the ethical desirability of an inclusive and decentralized policy in cyberspace.

KEYWORDS

Cyber Ethics, Blockchain Technology, Explainable Artificial Intelligence (XAI), Worldcoin, Proof-of-Personhood, Cyberethics-Mix Framework

1. INTRODUCTION

Technology poses ethical challenges in today's digital landscape, requiring new ethical frameworks. Rodrigues's "Cyberethics-Mix" (2011) proposed a template to navigate these ethical waters, foregrounding four pivotal elements encapsulated by the letter "P": data's property, precision, privacy, and pervasiveness. Blockchain promotes decentralized trust (Rodrigues, 2021a), while AI stresses the importance of explainability (Gunning & Aha, 2019). Worldcoin, combining cryptography and AI, aims to balance privacy and convenience, including a guarantee of a worldwide equitable coin distribution (Gent, 2023, p.42).

2. BLOCKCHAIN'S ETHICAL IMPLICATIONS AND TRUST DYNAMICS

Historically, trust expanded from tribal groups to the nation-state "Leviathan" (Hobbes, 1996). Blockchain technology (BT) extends the trust genesis through distributed consensus (Xiao et al., 2020, p. 1432), introducing the "Internet of Value" (Lacity, 2022, pp. 326, 327). Originating decentralized (Nakamoto, 2008), BT changes ethical dynamics and may drive innovation. The cybernetics-mix compass (Rodrigues, 2011, 2021) underscores the following ethical challenges.

2.1 Privacy

Blockchain's cryptographic security ensures transparency and quasi-anonymity in transactions, though identities could be traced raising ethical concerns. Despite this, blockchain enhances accountability, deterring fraud, and strengthens privacy by safeguarding sensitive data (Zyskind & Nathan, 2015). This focus on privacy can be advantageous, as consumers prefer companies that manage data responsibly (Bleier et al., 2020).

2.2 Property

The internet's early phase favored new intermediaries, leading to predatory digital business models exacerbating agency costs (Loureiro et al., 2021). As blockchain enhances property rights and eases asset transfers (Tan & Saraniemi, 2022), cryptocurrencies, NFTs, and smart contracts streamline intellectual

property monetization (De Filippi & Wright, 2018), reshaping business models and prompt an ethical reevaluation of ownership rights dynamics that may include monetary systems, suggesting an ethical shift to bridge individual goals with sustainable collective action (Rodrigues, 2021b).

2.3 Precision

Digital content accuracy is crucial in cyber ethics. Blockchain ensures data authenticity and integrity via collective validation, utilizing digital signatures and hash functions (Nakamoto, 2008; Pierro, 2017). This fosters trust in sectors like logistics and healthcare. Blockchain's principles align with the early internet's democratic ideals, reflecting Surowiecki's "crowd wisdom" (2004) and modern tools like "Twitter Community Notes" (Kim, 2023), emphasizing decentralized truth validation.

2.4 Pervasiveness

Blockchain technology (BT) brings forth transformative ethical implications, especially evident in decentralized finance (DeFi) sectors, which aim for broader financial inclusivity by bypassing traditional banking constraints (Abdulkhakeem & Hu, 2021). BT's inherent features champion transparency, diversity, and fairness, echoing the principles of democratic systems (Gupta et al., 2023, p. 556). Rodrigues underscores the ethical importance of securely sharing information, highlighting that such openness fosters innovation (2011, p. 335). With advancements like the "Orb" eye-scanner (Worldcoin, 2023) combined with Privacy Preserving Cryptographic Technologies (PPCT, Gupta et al., 2023), there is an amplification in digital accessibility, further solidifying democratic ethical underpinnings. However, the promise of BT is two-sided. Its transformative potential largely depends on political will and societal understanding. Barriers such as the digital divide, technological illiteracy, and ambiguous regulations might limit its reach.

3. THE ETHICAL LANDSCAPE OF EXPLAINABLE AI (XAI)

Explainable AI (XAI) represents a paradigm shift in artificial intelligence. There were no trust concerns while humans operated with early hardware and software. With straightforward algorithms, these systems executed the tasks they were programmed to do. In a sense, they represented the purest form of trust. However, ethical dilemmas arose with machine learning evolution, especially with "deep learning," the architecture that powers advanced Generative AI and Natural Language Processing models (like the GPT series).

A machine incapable of autonomy was not a threat and could be easily trusted, and people could easily believe more in a machine's output than human output. However, the advent of Machine Learning (ML) and the successful implementation of the concept of making machines think and perform tasks autonomously has resulted in a breach of the trust that existed, as machines can now operate and act independently. (Chamola et al., 2023)

XAI seeks to ethically restore trust in mathematical models eroded by business algorithms' pragmatism by ensuring AI's transparency. This endeavor is a considerable challenge, and the cyberethics-mix framework will now be used for investigating its ethical implications.

3.1 Privacy

XAI aims for transparency but can expose personal data used in AI models, potentially risking re-identification. While useful for audits, such clarity might overshare data details. "Differential privacy" counters this by adding controlled noise to data, preserving individual privacy while allowing overall analysis (Utaliyeva et al., 2023, p. 3). "Privacy-preserving machine learning" (PPML) further enables confidential model training (Hulsen, 2023, p. 657). EU's GDPR strengthens the "data minimization" principle, ensuring only necessary data processing (Oprescu et al., 2022, p. 59). Additionally, "layered explanations" vary AI detail based on user access (Guesmi et al., 2022, p. 177), and "synthetic data" provides computer-generated data for secure AI training (Oblizanov, 2023, p. 5).

3.2 Property

XAI's transparency can expose AI models to potential replication and reverse-engineering, posing intellectual property challenges. However, XAI can also identify possible infringements within copyrights or patents, allowing for rectifications. While the focus remains on AI's outputs, the intellectual property aspects of AI inputs, like training data, are often overlooked. Datasets may include copyrighted works, raising legal questions. Since AI relies on public or copyrighted data representing societal knowledge, should its creations be communal? Moreover, AI's inventiveness stirs debate about recognizing it as a creative entity.

It must be kept in mind that there are also many AI systems that appear to have independent hypothesis generation and testing capabilities, which in patent parlance might lead to independent invention by the intelligent system, i.e., without any direct action by humans. The question here is whether the owner, user, or programmer of the robot or an AI system is entitled to patent protection for an invention developed directly by AI. (Laukyte & Lucchi, 2022, p.177)

As AI assumes new societal roles, its legal identity becomes an issue. If AI undertakes human tasks, should it receive similar legal rights? Intellectual property laws might expand beyond human-only contributions. The European Parliament's 2017 resolution suggests an "electronic personhood" for advanced robots (Pagallo, 2018, p. 2), and more research is needed to align machines' property rights with ethical data usage.

3.3 Precision

XAI aims for clearer AI decisions, aligning with demands for transparency (Gerlings & Constantiou, 2020, p. 1284). However, "post-hoc" explanations might not reflect true model logic (Vale et al., 2022, p. 826), and there's a risk of favoring simpler models (Rudin, 2019) or crafting explanations to fit regulations (Wachter et al., 2017). Thorough research and oversight are vital to ensure XAI's authenticity.

3.4 Pervasiveness

XAI promotes AI understanding, which invites ethical input. However, simplifying complex AI decisions might misrepresent their intricacy (Doshi-Velez & Kim, 2017). Geoffrey Hinton suggests valuing performance over explainability (Ridley, 2022a), but although unexplained practical algorithms can have their merits, a thorough evaluation of AI systems is still vital (Arrieta et al., 2020, p. 83).

4. WORLDROID AT THE ETHICAL CONUNDRUM

Led by OpenAI's CEO, Sam Altman, Worldroid aims to revolutionize identity verification by combining a unique World ID with cryptocurrency rewards (Worldroid, 2023). The cyberethics-mix lens will now focus on that.

4.1 Privacy

Differentiating between humans and AI is vital, making it crucial to confirm humanness online without exposing personal details, balancing online ease with privacy. WLD, the currency behind World ID, demonstrates such unique utility, which proves cryptocurrency's multidimensional value (Rodrigues, 2021b), emphasizing transparency and user autonomy. Worldroid's authentication process is based on an eye scan that determines human uniqueness and then, allegedly, immediately erases the users' data, producing only an encrypted code called a zero-knowledge proof (ZKP). This method hides ledger input data while allowing validation (Morais et al., 2019), offering "proof of personhood" by giving each person a unique "World ID." Rather than verifying identity, this ID asserts its holder's digital uniqueness. Hence, users can skip password

memorization using their World ID to reaffirm their unique status, simplifying online interactions. The cyberethics-mix compass will be used to discern Worldcoin's ethical implications.

4.2 Property

World ID, a nuclear element of the Worldcoin project, redefines digital identity by granting human users a unique digital stamp, affirming their individuality. WLD tokens act as a currency and symbolize users' stake and involvement in the Worldcoin project. This dual approach amplifies the users' sense of ownership and belonging, bridging the gap between individual and collective rights. Cryptocurrency may encourage an organic collective action rooted in a principle of individual freedom and unique contribution, as stated by Rodrigues (2021a) and Worldcoin (2023). As put by George and colleagues, Worldcoin's ambition is more than to create a new currency; it is a pioneering endeavor to mitigate economic disparities and champion financial inclusion globally, regardless of an individual's location or socio-economic standing (2023, p. 136).

4.3 Precision

Reliance on AI underscores the need for accurate information and genuine human input verification. World ID counters AI-driven misinformation by confirming human authors using the iris's unique biometric traits (Sehrawat & Sankhyan, 2016, p. 437), ensuring unparalleled precision and preventing fraud.

4.4 Pervasiveness

Worldcoin aims to democratize online verification and finance access, bridging the digital divide. However, diverse cultural and regional trust challenges arise. Verifying individual uniqueness promotes equal digital access and deters abuse, emphasizing the potential for Universal Basic Income via World ID.

5. CONCLUSION

Distributed technologies may foster transparent, secure ecosystems to improve the world. For ethical engagement, user education is crucial. Worldcoin and World ID embody this digital evolution's opportunities and concerns. Once upholding transparency and biometric data privacy, as promised in Worldcoin's white paper, this worldwide initiative may result in a safe global decentralization. Thus, continuous scrutiny is essential. While blockchain and AI redefine cyber ethics, and Worldcoin addresses users' convenience and security challenges, well-informed political decisions are paramount for a promising future.

REFERENCES

- Abdulhakeem, S.A. and Hu, Q. (2021). Powered by Blockchain technology, DeFi (Decentralized Finance) strives to increase financial inclusion of the unbanked by reshaping the world financial system. *Modern Economy*, 12(01), pp. 1.
- Arrieta, A.B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A. and Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, 58, pp. 82-115.
- De Filippi, P. & Wright, A. (2018). *Blockchain and the law: The rule of code*. Harvard University Press.
- Doshi-Velez, F. & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. *arXiv preprint*, [online] Available at: <https://arxiv.org/abs/1702.08608>.
- European Parliament (2017). Recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)). *Official Journal of the European Union*, C 252, pp. 239-257.
- George, A. S., George, A. H. & Baskar, T. (2023). Worldcoin: A Decentralized Currency for a Unified Global Economy. *Partners Universal International Research Journal*, 2(2), pp. 136-155.
- Gerlings, J., Shollo, A. & Constantiou, I. (2020). Reviewing the need for explainable artificial intelligence (xAI). *arXiv preprint arXiv:2012.01007*.

- Guesmi, M. et al. (2022). Explaining user models with different levels of detail for transparent recommendation: A user study. In *Proceedings of the 30th ACM Conference on User Modeling, Adaptation and Personalization*, pp. 175-183.
- Gupta, R. et al. (2023). A Voting System Using Block Chain A Deep Survey. In *2023 3rd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*. IEEE, pp. 556-561.
- Hobbes, T. (1996) (1651). *Leviathan*. Oxford University Press.
- Hulsen, T. (2023). Explainable Artificial Intelligence (XAI): Concepts and Challenges in Healthcare. *AI*, 4(3), p. 652.
- Lacity, M.C., 2022. Blockchain: From Bitcoin to the Internet of Value and beyond. *Journal of Information Technology*, 37(4), pp. 326-340.
- Kim, J. (2023). When Partisans Fly: Twitter Community Notes and the Political Economy of Social Media Disinformation.
- Laukyte, M. & Lucchi, N. (2022). Creative AI: The Complex Relationship between Human Inventiveness and Intellectual Property. *BioLaw Journal-Rivista di BioDiritto*, (3), pp. 169-183.
- Loureiro, M., Pêgo, A. and Raposo, I. G. (2021). The Real Blockchain Game Changer: Protocols and DAOs for Coordinating Work to Provide Goods and Services. In: D. Rodrigues, ed., *Political and Economic Implications of Blockchain Technology in Business and Healthcare*. IGI Global, pp. 160-172.
- Morais, E., Koens, T., Van Wijk, C. & Koren, A. (2019). 'A survey on zero knowledge range proofs and applications', *SN Applied Sciences*, 1(1), pp. 1-17.
- Oblizanov, A. et al. (2023). Evaluation Metrics Research for Explainable Artificial Intelligence Global Methods Using Synthetic Data. *Applied System Innovation*, 6(1), p. 26.
- Oprescu, A.M. et al. (2022). Towards a data collection methodology for Responsible Artificial Intelligence in health: A prospective and qualitative study in pregnancy. *Information Fusion*, 83(84), pp. 53-78.
- Pagallo, U. (2018). Apples, oranges, robots: four misunderstandings in today's debate on the legal status of AI systems. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 376(2133), p. 20180168.
- Di Pierro, M. (2017). What is the blockchain?. *Computing in Science & Engineering*, 19(5), pp. 92-95.
- Rodrigues, D. (2011). Cyberethics of E-Business Social Networking. In M.M. Cruz-Cunha et al. eds. *Handbook of Research on Business Social Networking: Organizational, Managerial, and Technological Dimensions*. IGI Global – Business Science Reference Publishers, pp. 314-338.
- Rodrigues, D.D. (2021b). Blockchanging Money: Reengineering the Free World Incentive System. In D.D. Rodrigues, ed. *Political and Economic Implications of Blockchain Technology in Business and Healthcare*. IGI Global, pp. 69-117.
- Rodrigues, D. D. and Lopes, P. S. (2021a). Blockchanging Politics: Opening a Trustworthy but Hazardous Reforming Era. In: D. Rodrigues, ed., *Political and Economic Implications of Blockchain Technology in Business and Healthcare*. IGI Global, pp. 118-159.
- Rodrigues, D. D. (2021). Blockchanging Trust: Ethical Metamorphosis in Business and Healthcare. In D.D. Rodrigues, ed. *Political and Economic Implications of Blockchain Technology in Business and Healthcare*. IGI Global, pp. 1-41.
- Ribeiro, M.T., Singh, S. & Guestrin, C. (2016). "Why should I trust you?" Explaining the predictions of any classifier. *Proceedings of the 22nd ACM SIGKDD conference on knowledge discovery & data mining*, pp. 1135-1144.
- Ridley, M. (2022a). Explainable Artificial Intelligence (XAI). *INFORMATION TECHNOLOGY AND LIBRARIES*.
- Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, 1(5), pp. 206-215.
- Selbst, A. & Barocas, S. (2018). The intuitive appeal of explainable machines. *Fordham Law Review*, 87(3), pp. 1085-1139.
- Sehrawat, J.S. and Sankhyam, D. (2016). Iris Patterns as A Biometric Tool For Forensic Identifications: A Review. *Brazilian Journal of Forensic Sciences, Medical Law and Bioethics*, 5(4), pp. 431-440.
- Surowiecki, J. (2004). *The Wisdom of Crowds*. Lisboa: Lua de Papel.
- Utaliyeva, A., Shin, J. & Choi, Y. H. (2023). Task-Specific Adaptive Differential Privacy Method for Structured Data. *Sensors*, 23(4), p. 1980.
- Vale, D., El-Sharif, A. & Ali, M. (2022). Explainable artificial intelligence (XAI) post-hoc explainability methods: Risks and limitations in non-discrimination law. *AI and Ethics*, 2(4), pp. 815-826.
- Xiao, Y., Zhang, N., Lou, W. and Hou, Y. T. (2020). A survey of distributed consensus protocols for blockchain networks. *IEEE Communications Surveys & Tutorials*, 22(2), pp.1432-1465.
- Wachter, S., Mittelstadt, B. & Russell, C. (2017). Counterfactual explanations without opening the black box: Automated decisions and the GDPR. *Harvard Journal of Law & Technology*, 31(2), pp.841-887.
- Worldcoin (2023). Worldcoin Whitepaper. [online] Available at: <https://whitepaper.worldcoin.org/>.
- Worldcoin.org (2023b). Worldcoin open source and public goods. [online] Available at: <https://worldcoin.org/open-source>.
- Worldcoin (2023a). Worldcoin tokenomics. Available at: <https://whitepaper.worldcoin.org/tokenomics>.