

# AML/CFT/CPF endeavors in the crypto space: from blockchain analytics to machine learning \*

Nadia Pocher<sup>1,2,\*</sup>, Mirko Zichichi<sup>2,3</sup> and Stefano Ferretti<sup>4,5</sup>

<sup>1</sup>*Institute of Law and Technology, Universitat Autònoma de Barcelona, Bellaterra, Spain*

<sup>2</sup>*Department of Legal Studies, University of Bologna, Bologna, Italy*

<sup>3</sup>*Ontology Engineering Group, Universidad Politécnica de Madrid, Madrid, Spain*

<sup>4</sup>*Dipartimento di Scienze Pure e Applicate, University of Urbino Carlo Bo, Urbino, Italy*

<sup>5</sup>*Department of Computer Science and Engineering, University of Bologna, Bologna, Italy*

## Abstract

Financial applications of distributed ledger technologies (DLTs) generate regulatory concerns. In the crypto sphere, pseudonymity may safeguard privacy and data protection, but lack of identifiability cripples investigation and enforcement. This challenges the fight against money laundering and the financing of terrorism and proliferation (AML/CFT/CPF). Nonetheless, forensic techniques trace transfers across blockchain ecosystems and provide intelligence to regulated entities. This working paper addresses anomaly detection in the crypto space, the role of machine learning, and the impact of disintermediation.

## Keywords

AML/CFT/CPF, blockchain, cryptocurrency, machine learning, forensics

## 1. Introduction

Ever since the launch of Bitcoin [1], the opportunities offered by distributed ledger technologies (DLTs) have driven a fierce excitement for technology [2]. Leveraging distributed systems and cryptography, Nakamoto's work opened the way to recording and managing information trustworthily without intermediaries. Although the 'blockchain hype' goes beyond the financial domain, its first large-scale implementation and leading regulatory debates are financial in nature. The perception of this space as inherently anonymous triggers substantial concerns, and some of its fundamentals clash with accountability. From the early 2010s to the present day, scandals and scams ignite fears of illicit exploitation (e.g., Silk Road, Tornado Cash [3, 4]).

A prime example concerns the fight against money laundering and the financing of terrorism and proliferation (AML/CFT/CPF). The field is overseen by the Financial Action Task Force (FATF), whose risk indicators guide the understanding of crypto risks [5, 6]. The EU has harmonised its rules since 1991, and a major reform is about to establish a regulation-based single rulebook. Currently, the consolidated version of the AML Directive (AMLD) is Directive

---

*Artificial Intelligence Governance Ethics and Law (AIGEL), November 02 - December 19, 2022, Barcelona, Spain*

\*The contribution of Nadia Pocher and Mirko Zichichi received funding from the EU H2020 R&D programme under the MSCA ITN EJD grant agreement No 814177 Law Science and Technology Joint Doctorate - RIoE.

\*Corresponding author.

✉ [nadia.pocher@uab.cat](mailto:nadia.pocher@uab.cat) (N. Pocher); [mirko.zichichi@upm.es](mailto:mirko.zichichi@upm.es) (M. Zichichi); [stefano.ferretti@uniurb.it](mailto:stefano.ferretti@uniurb.it) (S. Ferretti)

🆔 0000-0003-1472-2963 (N. Pocher); 0000-0002-4159-4269 (M. Zichichi); 0000-0002-1911-4708 (S. Ferretti)



© 2023 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

 CEUR Workshop Proceedings (CEUR-WS.org)

(EU) 2015/849 as amended by Directive (EU) 2018/843. The regime relies on ‘regulated entities’, on which duties are imposed to prevent misuses of the financial systems and draw the attention of authorities when suspicions arise. They comprise financial entities – e.g., banks, but also cryptoasset service providers (CASPs) – but also non financial businesses and professions.<sup>1</sup>

In AML/CFT/CPF compliance, operations are typically screened in a partially automated way by software solutions that for crypto transactions are based on blockchain analytics. Nowadays, research displays how this sphere is less anonymous, disintermediated and decentralised than what the hype would suggest [8, 9]. Meanwhile, crypto-related crime seemingly decreased from USD 4.5 to 1.9 billion between 2019 and 2020 [10]. However, industry estimates keep reporting unsettling numbers: in 2021 crypto-related laundering amounted to USD 8.6 billion and illicit addresses were holding at least USD 10 billion [11]. The shortcomings of early-stage monitoring systems led to explore the use of artificial intelligence to enhance anomaly detection.

Against this backdrop, this paper provides an AML/CFT/CPF overview of blockchain forensics and introduces the role of machine learning. In particular, Section II outlines blockchain specifics and elaborates on the concepts of pseudonymity and de-anonymization. Section III dives into the AML/CFT/CPF regime and gives an interdisciplinary account of anomaly detection. Section IV addresses analytic techniques, thus introducing the role of machine learning solutions described in Section V. Section VI presents open issues, and Section VII concludes the work.

## 2. Pseudonymity and AML/CFT/CPF

The Bitcoin system [1] showed there is no need of a centralized party to reliably keep records of transactions. A distributed ledger is shared, replicated, and synchronised in a distributed and decentralized way, which in principle means control is distributed among participants [12]. In turn, a blockchain is a type of distributed ledger where data is recorded in a tamper-proof chain of blocks linked cryptographically. Blockchain types vary depending on whether the ledger is *public* (publicly readable) or *private* (readable only by authorised actors), and *permissionless* (everyone can add transactions) or *permissioned* (only authorised parties can).

Different ledger types manage identity differently. In public permissionless systems with no centralised authority, such as Bitcoin and Ethereum, the nodes that maintain the network “operate without association to a particular given identity” [13] and “are structurally designed as devices allowing anonymous transactions between peers” [14]. On the contrary, in permissioned ones there is a centralised entity/consortium that identifies the nodes, and key-pairs tend to be associated with real-world identities. In blockchains such as Bitcoin’s two elements co-exist: *ledger transparency* and *user pseudonymity* – i.e., using pseudonyms as identifiers [15]. Typically, a blockchain system manages identifiers through key pairs that identify the wallet holder uniquely [13]. Hence, the history of Bitcoin transactions is transparent, but participants are only related to addresses [16]. These alone do not convey any personal identifying information, unless there is an association with additional data [13]. However, they can be used to connect transactions to their history, and de-anonymization techniques can help establish links to real-world identities, or identify entity types, for the sake of compliance or investigation.

---

<sup>1</sup>The Markets in Crypto-Assets Regulation defines CASPs as providers of various crypto-related services including custody, administration, trading, exchange, advice. The definition includes FATF’s virtual asset service providers [7].

A key component of the AML/CFT/CPF regime is customer due diligence (CDD), including know your customer (KYC). These provisions pivot on the *identification* of some subjects – primarily, customers and beneficial owners,<sup>2</sup> and the *verification* or *authentication* of these identities. CDD also comprises assessing purpose and intended nature of the relationship and ongoing monitoring. As per the AMLD, all operations must be consistent with the entity’s knowledge of the customer, business and risk profile. *Identification* means establishing a real-world identity and a blockchain address that acts as a pseudonym is not sufficient to hold users accountable. Hence, *identifiability* safeguards *accountability*. On the other hand, the authenticity of said identity must be *verified* against a(n) (set of) identifiers [17].

In AML/CFT/CPF compliance, *identifiability* plays a crucial role in all risk-based assessments performed by the regulated entity – e.g., to decide whether to accept a client, perform an operation, assess the risk of the client and/or the operation, if enhanced due diligence is required. Hence, it is not surprising the primary concerns about crypto misuse were linked to the lack of *identifiability* of the parties involved, due to the absence of real-world identities.

### 3. Anomaly Detection

Since Bitcoin proved not to possess key anonymity properties [16], new cryptographic techniques were deployed in new currencies, services, and networks – e.g., ‘privacy coins’ such as Monero and Zcash,<sup>3</sup> which typically pursue anonymity explicitly [18]. Authorities outline these scenarios in *red flag indicators* that describe suspicious activities to guide compliance and supervision. They are usually provided in a rule-based format as templates of sequences of actions, and FATF’s indicators inform national and institutional guidance. A list was published in 2020 [5] and complemented in 2021 [6].<sup>4</sup> In some countries, these templates are named anomaly indicators [19]. In data science, anomaly detection consists of processing data to pinpoint events significantly different from the dataset [20]. The concept can also be analyzed from a regulatory perspective, and in compliance technology the two viewpoints merge.

Indeed, risk indicators are the basis of transaction monitoring solutions, whose hits are generated through a process of rule-matching. In other words, operations are screened in real-time to detect anomalous activities in an automated way and the tools usually rely on customizable rules – i.e., alerts are produced if a transaction meets predefined standards of suspicion. Accordingly, transaction monitoring solutions were defined as “predominantly rules-based thresholding protocols tuned for volume and velocity of transactions with tiered escalation procedures” [21]. This means that the preliminary review of a flagged transfer usually relies on suspiciousness heuristics such as political exposure, geographical dynamics, transaction type and properties, behavioral logic [21], as enshrined by risk indicators.

Examples of types of rules are: (a) high risk or non-permitted jurisdiction -> alert rule:

---

<sup>2</sup>Identification is based on data from a reliable and independent source, which includes means of electronic identification such as the eIDAS framework.

<sup>3</sup>The privacy motive was to obtain *fungibility*. If the history of transfers can be traced, a given unit is tainted by previous actions. If the transaction history is obfuscated, each unit is equal, just as physical coins and banknotes [18].

<sup>4</sup>It outlines indicators pertaining to transaction types and features, transaction patterns, anonymity, features of senders/recipients, specifics of source of funds/wealth, geographic risks. Anonymity-related indicators include cases of obfuscation (e.g., privacy coins) and disintermediation (e.g., self-hosted wallets).

transfers from/to the jurisdiction, based on the IP address; (b) transfers above EUR 1,000 -> alert rule: transfers above the aggregated value within a time frame; (c) transfers unusual for a specific customer -> customized alert rule: transfers exceeding by 30% the average transaction pattern of the customer. If the system finds a match with a rule, the transfer is flagged accordingly. In this context, a considerable effort in terms of time and human resources is dedicated to reviewing the alerts generated by the rule-matching process. To this end, regulated entities have internal procedures according to which multiple layers of analysts decide whether to escalate the alert.

## 4. Blockchain Forensics and Network Analysis

While new techniques of anonymization were developed, the private sector and law enforcement agencies (LEAs) started tracing crypto transfers through *blockchain forensics* or *analysis*.<sup>5</sup> Indeed, even if the set of publicly accessible data in (certain) blockchain systems offers great material to investigators, a specialized knowledge is needed for a useful interpretation. This is because the details of the various networks usually translate into misleading pieces of information to non-expert eyes [23]. Furthermore, there are often preliminary activities of acquisition or extraction of private keys, public addresses and wallet files [24]. Forensic techniques determine the likelihood of linking a real-world identity to a (set of) transaction(s), and the degree of success depends on their effectiveness vis-à-vis privacy enhancements.

The presence of these two sets of actors pushing towards higher peaks of *obfuscation* and more efficient *accountability* generates mutual influences. Indeed, the implementation of innovative cryptographic techniques led to new investigative strategies. This, in turn, spurred increasingly sophisticated cryptographic methods in a race that seems never ending. Meanwhile, various analytic strategies leveraged the fact that transactions consist of flow relationship between entities and can be organized and visualized in the form of a network. These methods of analysis focus, primarily, on reusing an account for multiple transactions or co-using multiple accounts for a single transaction to match multiple accounts to the same user/service.

One should consider that the Bitcoin blockchain, but also IOTA's Tangle, employs a type of address-based data representation based on unspent transaction outputs (UTXOs).<sup>6</sup> This means there are no accounts at the protocol level, and transaction representation is based on *inputs*, amounts spent on a transaction, and *outputs*, amounts received. A wallet's balance equals to the outputs not yet spent. When making a transfer, the whole amount of coins of an UTXO corresponding to an address must be spent. The 'change', if any, is transferred to an address owned by the sender [24]. Thus, usually one of a transaction's outputs is the 'change' address.

Clustering algorithms enable statistical evaluations, especially to determine if a given address belongs to a specific identified cluster, such as an exchange, to a (yet) unidentified cluster, or to no cluster. They are often proprietary and owned by analytic companies. They allow to visualize the flow of funds between identified clusters instead of between individual addresses. This leads to inferences about the type of entities involved and, when the algorithm is applied

---

<sup>5</sup>*Blockchain forensics* was defined as the use of science and technology in the investigation and establishment of facts in court, dealing primarily with recovering and analyzing evidence generated by transacting on the blockchain [22].

<sup>6</sup>Other blockchains use an account-based system. Forensic techniques have been mostly tested on UTXO-based networks, but data-exploitation methods have been deployed on Ethereum [4, 25, 26] and other networks [27, 28, 29].

to huge datasets, about the degree of receiving and sending relationships between clusters. This is of great value in risk-sensitive assessments performed by regulated entities. In particular, when it comes to evaluating a specific exposure vis-à-vis their risk appetite.

A set of clustering methods are based on heuristics [30, 31], and aim to link more addresses to an identity [32], under the assumption that users can be associated with addresses through heuristics [33]. Clustering can focus on similar behavioural patterns, co-spending or sources other than the transaction history [34, 30], gathered through web-scraping and open source intelligence tools to find correlation between transactions and public user profiles [35, 16]. Other methods focus on mixers [36], and on cross-chain transactions [32].

The network generated by transaction flows can be visualized as a graph – i.e., a mathematical model comprising a set of nodes and a set of edges connecting nodes' couples. In blockchains, nodes can be (groups of) accounts and edges transaction between accounts. This means that specific methods can be applied to infer intelligence [35, 37] and the graph structure helps spot illicit transactions by exploring the network characteristics. Given a transaction  $t$ , it is possible to collect all connected transactions and recursively search for other ones up to a certain depth level. Within the connected graph, neighbouring transactions and their classified value aid the classification of  $t$  – i.e., each transaction has neighbours that influence its classification.

## 5. Machine Learning Applications

Indicators aim to provide a structure and clear benchmarks regarding AML/CFT/CPF risks. Often, however, rule-based indicators can be for the most part descriptive and distant from industry best practices. Although interpretability is an advantage, the simplicity of rule-based systems produces false positives estimated at around 95–98% [38]. Indeed, transaction datasets are massive, dynamic, high dimensional, non-linear, as well as often fragmented, inaccurate, incomplete, or inconsistent. The difficulty to automate synthesis from various data streams leaves the task up to human analysts. This generates a vicious circle of over-reporting due to the cost asymmetry between false positives and false negatives [21].

The insufficiency of rule-based systems suggested to automate several processes. Some machine learning-based methodologies are deployed and investigated not only to detect anomalies and optimize alerts, but also to draw intelligence from transaction and cluster classification. The underlying idea is that building models able to infer patterns from historical data increases detection rates and decreases false positives [39], while some approaches pursue to map and predict illicit transactions [37]. A main distinction in machine learning is between *supervised* methods, where labelled datasets are used to train algorithms, and *unsupervised* techniques, where the model works on its own to discover patterns and information previously undetected. Supervised learning needs an initial training dataset tagged and annotated.<sup>7</sup> These techniques are generally regarded as good for making predictions and they are used for transaction classification.

Meanwhile, unsupervised methods are usually deployed if there is label scarcity. In the crypto sphere, there is a considerable shortage of annotated datasets, due to the scale of the phenomena, the timing of investigations, and the cost of manual labelling. Therefore, analytic companies assume a crucial role in labelling datasets, where a transaction can be tagged as

---

<sup>7</sup>Some examples are Decision Trees, Random Forests, Boosting Algorithms, and Logistic Regression.

licit or illicit based on investigations, public information, or proprietary data. The resulting annotated dataset can be used to train algorithms. To mitigate the issue of label scarcity and the drawbacks of unsupervised methods, one can pursue alternative paths such as generating fully or partially synthetic datasets or improving algorithm performance by organizing the training data differently. Datasets of blockchain transactions can be organized in the form of a graph.

Graph analytics fits well the AML/CFT/CPF sphere because transactions involve relationships between entities that can be represented in graph structures. For instance, graph convolutional networks aim to learn a function of features on a dataset structured as a graph. The key idea is that each node receives and aggregates features from its neighbors to represent and compute its local state [40]. Further, graph attention networks give different importance to each node's edge by using attention coefficients [41]. Both models seem promising in predicting illicit transactions and the type of entity to which an unidentified one belongs. They combine transaction features with 'close' graph data. However, some labelling is required, and it is still challenging to state if there are specific graph patterns that suggest suspicious activities.

The researchers in [34] collaborated with Chainalysis to deploy a supervised approach to predict the type of entities yet not identified, concluding it is possible to predict if a cluster belongs to predefined categories such as exchange, gambling, shuffling. Further, [37] benchmarked graph convolutional networks against supervised methods, while [38] extended the work to a non-blockchain context with the aim to reduce false alerts through supervised methods, where the produced score enables alert suppression or prioritization. The GuiltyWalker [42] leverages random walks on a cryptocurrency graph to characterize distances to previous suspicious activity. With transaction graphs modelling illicit activity over time, however, it is difficult to apply methods that are efficient and whose results can be understood by humans. Indeed, literature is still lacking research into explainable AI techniques for anomaly detection [43].

## 6. Discussion

Analytics is largely deployed in intermediated contexts. This is not surprising, since AML/CFT/CPF explicitly does not apply to person-to-person transfers, and about 80% of crypto transfers go through centralized exchanges [44, 45]. Nonetheless, transfers enabled by unhosted wallets and decentralised finance (DeFi) are increasingly popular and require specific techniques to meet specific monitoring needs. A clear example of the tension between forensics and disintermediation can be found in the debate on the 'crypto travel rule', which mandates regulated entities to guarantee the traceability of crypto transfers. The rule expands the scope of application of measures concerning wire transfers, as required by the FATF. In the EU, it was implemented by recasting Regulation (EU) 2015/847, and CASPs/financial institutions have to collect, hold, submit and share specific data on originators and beneficiaries of crypto transfers.

However, wallets hosted by providers (typically regulated entities) are not the only way to store and transfer cryptoassets. Using self-hosted/unhosted wallets, users can have full control over their funds and transfer/receive them to/from another unhosted wallet or, if regulation allows it, to/from a hosted one. The EU recast regulates unhosted wallets only when they interact with hosted ones. Notably, transfers of EUR 1,000 or more are allowed only if the unhosted wallet is controlled by the customer. This poses the challenge of obtaining proof of

control from the customer and verifying it. Meanwhile, the industry denounced the absence of standards and technical solutions to effectively and affordably comply [46].

From a related perspective, unhosted-to-unhosted wallet transfers may be used to elude traceability and cash limits. This challenges the efficacy of the approach and shows how regulation has yet to capture P2P transfers. Meanwhile, the evolution of DeFi has displaced illicit activities. The total value of these projects reportedly amounted to USD 1 billion in January 2020, USD 27 billion in January 2021, USD 60 billion April 2021, and USD 40 billion in November 2022. Accordingly, the use of DeFi platforms for laundering increased by 1.964% between 2020 and 2021. They received 17% of funds originating from illicit addresses in 2021 (vs 2% in 2020), and in 2021 funds derived from crypto thefts were increasingly sent to DeFi platforms (51%) or risky services (25%), while only 15% went to centralized exchanges [11].

## 7. Final remarks

This paper addressed the role of machine learning to gather AML/CFT/CPF insights, and argued that the use of these methods can improve the efficiency of forensics. Considering the evolution of the crypto space, regulated entities and LEAs will increasingly analyze a large number of transactions whose transparency is obfuscated. While the use of unhosted wallets and decentralized platforms mean the lack of regulated counterparties, the industry denounces difficulties in complying with rules to ensure traceability. Complex solutions of compliance technology, however, are not enough, and they must be considered in an interdisciplinary fashion: it is pivotal to heed the relationship between any implemented approach and the regulatory environment. For instance, the effectiveness of an algorithm largely depends on the extent to which it generates useful alerts. Although the quantity of transaction data suggest machine learning will continue to be key, synergies between public and private stakeholders are needed to put forward innovative compliance tools and safeguard interpretability and explainability. The fact that labelled transaction datasets are currently proprietary in large part cannot help but impact also supervisory activities. Hence, it is crucial to establish multi-stakeholder dialogue to position blockchain analytic experimentation within initiatives that consider AML/CFT/CPF from a socio-technical, operational, and regulatory viewpoint.

## References

- [1] S. Nakamoto, Bitcoin: A Peer-to-Peer Electronic Cash System, [www.bitcoin.org](http://www.bitcoin.org) (2008).
- [2] O. Ali, M. Ally, Y. Dwivedi, et al., The state of play of blockchain technology in the financial services sector: A systematic literature review, *International Journal of Information Management* 54 (2020) 102199.
- [3] B. Akhgar, M. Gercke, S. Vrochidis, H. Gibson, *Dark Web Investigation*, 2021.
- [4] M. Wu, W. McTighe, K. Wang, I. A. Seres, N. Bax, M. Puebla, M. Mendez, F. Carrone, T. De Mattey, H. O. Demaestri, M. Nicolini, P. Fontana, *Tutela: An Open-Source Tool for Assessing User-Privacy on Ethereum and Tornado Cash* (2022).
- [5] FATF, *Virtual Assets Red Flag Indicators of Money Laundering and Terrorist Financing*, 2020. URL: <http://www.fatf-gafi.org/>.

- [6] FATF, Virtual Assets and Virtual Asset Service Providers - Updated Guidance for a Risk-Based Approach (2021). URL: [www.fatf-gafi.org](http://www.fatf-gafi.org).
- [7] FATF, International Standards on Combating Money Laundering and the Financing of Terrorism & Proliferation: FATF Recommendations, 2012. URL: <https://www.fatf-gafi.org/>.
- [8] T. Schrepel, Smart Contracts and the Digital Single Market Through the Lens of a 'Law + Technology' Approach, 2021. URL: <https://papers.ssrn.com/abstract=3947174>.
- [9] C. Campajola, R. Cristodaro, F. M. De Collibus, T. Yan, N. Vallarano, C. J. Tessone, The Evolution Of Centralisation on Cryptocurrency Platforms (2022) 1–14.
- [10] CipherTrace, Cryptocurrency Crime and Anti-Money Laundering Report, 2021.
- [11] Chainalysis Team, The 2022 Crypto Crime Report (2022).
- [12] ITU-T Focus Group on DLT, Technical Report FG DLT D1.2 - Distributed ledger technology overview, concepts, ecosystem, 2019. URL: <https://www.itu.int/>.
- [13] F. Wang, P. De Filippi, Self-Sovereign Identity in a Globalized World: Credentials-Based Identity Systems as a Driver for Economic Inclusion, *Frontiers in Blockchain* 2 (2020) 1–22.
- [14] M. Quiniou, *Blockchain: the advent of disintermediation*, ISTE Ltd, 2019.
- [15] A. Pfitzmann, M. Hansen, A terminology for talking about privacy by data minimization: Anonymity, Unlinkability, Undetectability, Unobservability, Pseudonymity, and Identity Management, *Technical University Dresden* (2010) 1–98.
- [16] N. Amarasinghe, X. Boyen, M. McKague, A Survey of Anonymity of Cryptocurrencies, in: *ACM International Conference Proceeding Series*, Sydney, 2019.
- [17] J. Grijpink, C. Prins, Digital anonymity on the Internet: New rules for anonymous electronic transactions? An exploration of the private law implications of digital anonymity, *Computer Law and Security Report* 17 (2001) 379–389.
- [18] J. Harvey, I. Branco-Illodo, Why Cryptocurrencies Want Privacy: A Review of Political Motivations and Branding Expressed in “Privacy Coin” Whitepapers, *Journal of Political Marketing* 19 (2020) 107–136.
- [19] Bank of Italy, *Provvedimento recante gli indicatori di anomalia per gli intermediari*, 2010. URL: <https://uif.bancaditalia.it/>.
- [20] A. Kamišalić, R. Kramberger, I. Fister, Synergy of blockchain technology and data mining techniques for anomaly detection, *Applied Sciences (Switzerland)* 11 (2021).
- [21] M. Weber, J. Chen, T. Suzumura, A. Pareja, T. Ma, H. Kanezashi, T. Kaler, C. E. Leiserson, T. B. Schardl, *Scalable Graph Learning for Anti-Money Laundering: A First Look* (2018). [arXiv:1812.00076](https://arxiv.org/abs/1812.00076).
- [22] T. Phan, *Exploring Blockchain Forensics*, 2021.
- [23] D. Silva Ramalho, N. Igreja Matos, What we do in the (digital) shadows: anti-money laundering regulation and a bitcoin-mixing criminal problem, *ERA Forum* 22 (2021).
- [24] N. Furneaux, *Investigating Cryptocurrencies: Understanding, Extracting, and Analyzing Blockchain Evidence*, Wiley, 2018.
- [25] M. Bartoletti, S. Carta, T. Cimoli, R. Saia, Dissecting Ponzi schemes on Ethereum: Identification, analysis, and impact, *Future Generation Computer Systems* 102 (2020) 259–277.
- [26] S. Ferretti, G. D’Angelo, On the ethereum blockchain structure: A complex networks theory perspective, *Concurrency and Computation: Practice and Experience* 32 (2020).
- [27] L. Tennant, Improving the Anonymity of the IOTA Cryptocurrency (2017) 1–20. URL: <https://laurencetennant.com/papers/anonymity-iota.pdf>.



- [28] P. Ince, J. K. Liu, P. Zhang, Adding confidential transactions to cryptocurrency IOTA with bulletproofs, volume 11058 LNCS, Springer International Publishing, 2018.
- [29] P. Moreno-Sanchez, M. Zafar, A. Kate, Listening to whispers of ripple: Linking wallets and deanonymizing transactions in the ripple network, Proceedings on PETs (2016).
- [30] M. Lischke, B. Fabian, Analyzing the bitcoin network: The First Four Years, Future Internet 8 (2016). doi:10.3390/fi8010007.
- [31] E. Androulaki, G. O. Karame, M. Roeschlin, T. Scherer, S. Capkun, Evaluating user privacy in Bitcoin, Lecture Notes in Computer Science 7859 (2013) 34–51.
- [32] H. Al Jawaheri, M. Al Sabah, Y. Boshmaf, A. Erbad, Deanonymizing Tor hidden service users through Bitcoin transactions analysis, Computers and Security 89 (2020).
- [33] S. Meiklejohn, M. Pomarole, G. Jordan, K. Levchenko, D. McCoy, G. M. Voelker, S. Savage, A fistful of Bitcoins: Characterizing payments among men with no names, Communications of the ACM 59 (2016) 86–93.
- [34] H. H. S. Yin, K. Langenheldt, M. Harlev, R. R. Mukkamala, R. Vatrupu, Regulating Cryptocurrencies: A Supervised Machine Learning Approach to De-Anonymizing the Bitcoin Blockchain, Journal of Management Information Systems 36 (2019) 37–73.
- [35] M. Fleder, M. S. Kester, S. Pillai, Bitcoin Transaction Graph Analysis (2015) 1–8. arXiv:1502.01657.
- [36] J. Wu, J. Liu, W. Chen, H. Huang, Z. Zheng, Y. Zhang, Detecting Mixing Services via Mining Bitcoin Transaction Network with Hybrid Motifs, Journal of Latex Class Files 14 (2020). arXiv:2001.05233.
- [37] M. Weber, G. Domeniconi, J. Chen, D. K. I. Weidele, C. Bellei, T. Robinson, C. E. Leiserson, Anti-Money Laundering in Bitcoin: Experimenting with Graph Convolutional Networks for Financial Forensics (2019). arXiv:1908.02591.
- [38] A. N. Eddin, J. Bono, D. Aparício, D. Polido, J. T. Ascensão, P. Bizarro, P. Ribeiro, Anti-money laundering alert optimization using machine learning with graphs, 2021. doi:10.48550/ARXIV.2112.07508.
- [39] J. S. Lorenz, Machine learning methods to detect money laundering in the Bitcoin blockchain in the presence of label scarcity, Ph.D. thesis, 2021.
- [40] T. N. Kipf, M. Welling, Semi-supervised classification with graph convolutional networks, arXiv:1609.02907 (2016).
- [41] P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Lio, Y. Bengio, Graph attention networks, arXiv:1710.10903 (2017).
- [42] C. Oliveira, J. Torres, M. I. Silva, D. Aparício, J. T. Ascensão, P. Bizarro, Guiltywalker: Distance to illicit nodes in the bitcoin network, arXiv:2102.05373 (2021).
- [43] D. V. Kute, B. Pradhan, N. Shukla, A. Alamri, Deep learning and explainable artificial intelligence techniques applied for detecting money laundering—a critical review, IEEE Access (2021).
- [44] BCB Group, Centralized vs. Decentralized Exchanges, 2022. URL: <https://www.bcbgroup.com/centralized-vs-decentralized-exchanges/>.
- [45] The Block, DEX to CEX Spot Trade Volume, 2022. URL: <https://www.theblock.co/data/decentralized-finance/dex-non-custodial/dex-to-cex-spot-trade-volume>.
- [46] C. R. Goforth, Crypto Assets : A Fintech Forecast (2020) 5–25.