






# Supporting Resilient, Ethical, and Verifiable Anonymous Identities Through Blockchains

Alberto De Marchi<sup>1</sup><sup>a</sup>, Lorenzo Gigli<sup>2</sup><sup>b</sup>, Andrea Melis<sup>2</sup><sup>c</sup>, Luca Sciallo<sup>2</sup><sup>d</sup> and Fabio Vitali<sup>2</sup><sup>e</sup>

<sup>1</sup>University of the Bundeswehr Munich, Department of Aerospace Engineering, Neubiberg, Germany

<sup>2</sup>University of Bologna, Department of Computer Science and Engineering, Bologna, Italy

**Keywords:** Online Anonymity, Self-Sovereign Identity, Cyberbullying, Blockchain-Based Identity, Blockchain.

**Abstract:** In recent years, anonymity on the internet has come under intense scrutiny for enabling criminal behaviors like cyberbullying, disinformation, child exploitation, and illicit financial activities. Nevertheless, strong advocates highlight its importance as a protective space for legitimate and ethical actions that individuals may prefer to keep separate from their real-world identities. This paper presents a protocol for authenticated anonymity, enabling anonymous usage that remains unlinkable to real identities unless criminal activity is detected. Blockchain offers a robust and secure framework to manage these needs. While existing solutions — e.g., self-sovereign identities — grant users full control over their disclosure, they lack proper accountability. To address this limitation, the proposed protocol employs a blockchain-driven mechanism that supports anonymous yet verifiable identities. De-anonymization is achieved exclusively through multi-party consensus on the blockchain, triggered by explicit and non-repudiable requests. We provide the formal mathematical model of the protocol and offer some evaluations of its robustness and fault tolerance, even under large-scale identity management scenarios.

## 1 INTRODUCTION

While (state-issued) public digital identities (ISO, 2011) are identifying codes assigned to persons or firms to enable full use of online public and private services under their own name, *online anonymity* refers to the capacity to behave on the Internet without disclosing one's real-world name, whereabouts, or personal information — including any other anonymous persona one may be using for different aims or in different settings.

For its supporters, online anonymity serves as a shield against unfair backlash toward individuals who merely seek to express themselves and live according to their own choices. It is viewed as protection against the overreach of powerful external forces:


- **Societal overreach**, protecting users from people in their circles (spouse, family, employers, clergy) who may reject their life decisions.


- **Infrastructure overreach**, guarding users from service providers and platforms that collect and trade personal data, often without consent.
- **Law enforcement overreach**, where even in liberal democracies, excessive surveillance practices are growing despite legal safeguards (Mateescu et al., 2015).
- **State overreach**, where harsh regimes legally silence dissent and control digital life through oppressive laws.


For its critics, anonymity hinders law enforcement from identifying people committing serious crimes:


- **Crimes Against Individuals:** hate speech, threats, cyberbullying, or impersonation — such as for scams or deceit.
- **Crimes Against Information Spaces:** fake news and orchestrated outrage manipulate public discourse by flooding networks with disinformation.
- **Unlawful Acts:** criminal trades benefit from anonymity, e.g., drug trafficking, laundering, terrorism, child abuse.


The debate has become highly polarized. Some organizations, even from fairly different political

<sup>a</sup> <https://orcid.org/0000-0002-3545-6898>

<sup>b</sup> <https://orcid.org/0000-0001-9714-3777>

<sup>c</sup> <https://orcid.org/0000-0002-0101-2551>

<sup>d</sup> <https://orcid.org/0000-0002-8973-4486>

<sup>e</sup> <https://orcid.org/0000-0002-7562-5203>

backgrounds, such as the Electronic Frontier Foundation (EFF, 2024) and the Cato Institute (Shapiro and Meyer, 2015), strongly defend anonymity as vital to freedom and democracy. Conversely, others highlight its dangers: for instance, (Fredheim et al., 2020) outlines how foreign powers and corporations manipulate debate using networks of fake accounts. Consequently, various laws now seek to curb anonymous Internet usage, including financial regulations (Pat, 2001; European Parliament and Council of the European Union, 2015) and the UK Online Safety Act 2023 (UK, 2023).

Meanwhile, new architectures for anonymous identities are emerging, like decentralized identifiers (DIDs) (Sporny et al., 2022) and self-sovereign identity frameworks such as ESSIF (eSSIF Lab, 2022), often relying on blockchain technologies to ensure robust anonymity while minimizing technical mistakes.

Unfortunately, ideological divisions leave little space for a shared solution. Yet a simple compromise could be imagined: an architecture enabling strong anonymity for legal use, but also fast, controlled deanonymization for crimes, with safeguards against abuse.

In this paper, we propose such an architecture—a blockchain-based anonymization protocol that:

1. connects public and anonymous identities securely for legitimate uses;
2. enables reliable deanonymization in case of crimes;
3. restricts deanonymization to authorized actors only;
4. enforces traceability and accountability in the deanonymization process;
5. requires multi-party consensus to prevent unethical access.

We envision a transnational network of trusted agencies (Union of Identity Providers, UIP) agreeing on technical and ethical standards to ensure that deanonymization is lawful, justified, and immune to covert or oppressive access. The system employs *authenticated anonymous* identities—accounts invisibly linked to real individuals and only unmasked under strict conditions. We also analyze attack scenarios, including coalitions of malicious actors attempting to bypass or misuse the protocol for unethical ends.

**Outline.** The rest of the paper is structured as follows: Section 2 reviews blockchain-based anonymity and identity, Section 3 introduces the protocol, Section 4 discusses threat models, Section 5 presents an analytical model, and Section 6 concludes with final thoughts and future directions.

## 2 RELATED WORK

Research on anonymity spans privacy, security, accountability, and traceability. A core challenge is balancing user privacy with the ability to attribute actions when needed. Blockchain offers a decentralized, tamper-resistant infrastructure supporting anonymous yet verifiable interactions. Combined with zero-knowledge proofs (ZKPs), blockchains can validate statements without disclosing identities, enabling systems where users remain anonymous but actions are auditable under defined conditions.

Zk-creds (Rosenberg et al., 2023) is an anonymous credential system based on general-purpose ZKPs. It avoids assumptions of uniform formats or issuer cooperation by transforming existing identity documents into anonymous credentials without requiring issuer modifications. It supports various accumulator schemes (e.g., Merkle trees, RSA accumulators, Verkle trees), each with distinct tradeoffs.

AttriChain (Shao et al., 2020) enables anonymous yet traceable identities in permissioned blockchains. It uses attribute-based signatures and threshold encryption to trace transactions, with unforgeability and anonymity preserved through ZKPs, offering a balance between privacy and accountability.

ChainAnchor (Hardjono et al., 2014) extends the Bitcoin blockchain with an identity layer that uses ZKPs to prove anonymous group membership. Only verified anonymous users can write to the blockchain, while transactions remain publicly readable and verifiable.

SilentProof (Mosakheil and Yang, 2024) addresses the computational limits of ZKP-based anonymous credentials on constrained devices. It delegates complex operations to the blockchain, acting on user consent without accessing private data, and employs blind signatures with proxy re-identification to preserve privacy.

Some use cases require identity recovery. Lafourcade et al. (Lafourcade et al., 2024) present a ticketing system allowing secure and anonymous transfers. It ensures privacy, prevents double spending, and enables controlled deanonymization by a judge under strict conditions. Karantaidou et al. (Karantaidou et al., 2024) introduce blind multi-signatures (BMS), enabling verifiable, unlinkable credentials from a dynamic group of signers. Even if signers collude, tokens remain untraceable. Two BMS constructions are proposed: one using BLS signatures, the other based on discrete logs, both proven secure in the Algebraic Group Model.

Our solution is independent of specific blockchain platforms. It uses only standard blockchain fea-

tures—public verifiability, integrity, decentralization—without requiring protocol extensions. It supports many authenticated anonymous identities per real identity, improving privacy and resilience against statistical attacks. Identity recovery requires collective consensus, ensuring no single authority can break anonymity unilaterally.

### 3 ANONYMIZATION PROTOCOL

In this section, we summarize the anonymization protocol from (Sciullo et al., 2024), focusing on the flow relevant to deanonymization—i.e., recovering anonymous identities linked to a user. The full deanonymization process is detailed in Section 3.1 as a novel contribution (see fig. 1).

The identity setup begins when a user submits personal documents and a public key to her National Identity Provider (NIP). After verifying identity, the NIP issues a Public Identity Data (PID), a token linking the user to the system without revealing personal details. The user then requests a seed phrase by sending her PID and public key to a smart contract, which generates random words, encrypts them, and stores them on-chain. It also computes a Secret Identity Data (SID) from the hash of the full phrase. The encrypted words are distributed among multiple UIP nodes to allow future recovery via consensus. A symmetric key links PID and SID on-chain.

To authenticate anonymously, the user requests a Secret Authentication Code (SAC) by sending a signed and encrypted SID with a new public key. The NIP validates SID ownership, maps the SAC to the SID, and records the mapping on the blockchain. When using the service, the user presents the SAC, corresponding public key, and signature—proving identity without revealing it.

Encryption and digital signatures ensure confidentiality and authenticity throughout.

#### 3.1 Deanonymization

The protocol keeps user anonymity intact unless law enforcement requires identification (e.g., in criminal cases). In such cases, a formal deanonymization protocol is available, with safeguards against abuse.

The process starts when a NIP requests to identify anonymous accounts linked to a user.

In Phase 1, the NIP submits a signed request including the user's PID, NIP's public key, legal reference, and reason code. Phase 2 verifies the signature and checks NIP status:

- **White-listed:** request proceeds automatically.

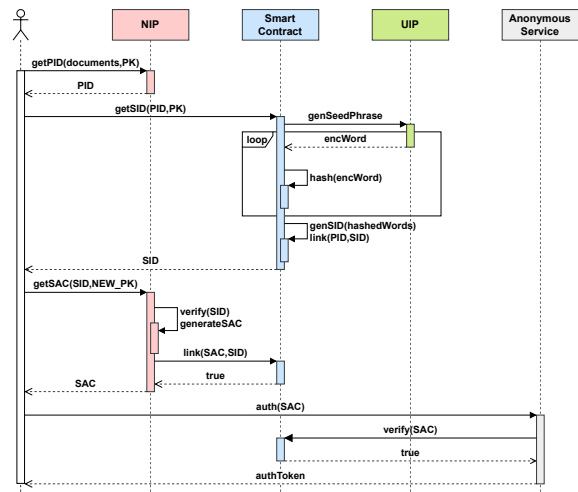


Figure 1: Sequence diagram of the anonymization flow.

- **Gray-listed:** manual review is triggered before proceeding.
- **Black-listed:** request is rejected and logged.

If approved, Phase 3 starts seed phrase recovery. The contract emits a `DeanonymizationEvent` with PID and NIP public key. UIP nodes holding fragments encrypt and upload them to the blockchain with signatures and indexes.

In Phase 4, the NIP collects encrypted words. For each index in the 24-word seed, it decrypts all submissions and selects the majority word. This redundancy ensures resilience to faulty or malicious nodes.

With the phrase recovered, the NIP recomputes the SID by hashing all word hashes. In Phase 5, the system retrieves all SACs tied to the SID. Each SAC corresponds to a public key representing one anonymous identity.

The overall complexity is:

Phase 1–2:  $O(1)$  (fixed cryptographic operations).

Phase 3:  $O(\bar{Q} \times K)$ , with  $\bar{Q}$  UIP nodes holding fragments and  $K$  words per node.

Phase 4:  $O(N \times M)$ , where  $N = 24$  and  $M$  is redundancy level.

Phase 5:  $O(A)$ , where  $A$  is the number of identities.

Thus, total time complexity is  $O(\bar{Q} \times K + N \times M + A)$ , dominated by node count and identity volume. The protocol remains efficient and secure for legitimate deanonymizations.

### 4 THREAT MODEL

We define the threat model following the OWASP methodology and the model used in Identity Man-

## Algorithm 1: User Deanonimization Protocol.

---

**Input:** User PID  $pid$ , NIP Public Key  $pk_{NIP}$ , NIP Digital Signature  $sig_{NIP}$   
**Output:** Collection of anonymous identities associated with the target user

---

```

// Phase 1: Request Submission
1   $req_{deAnon} \leftarrow (pid, pk_{NIP}, reason, legal\_ref)$ 
2   $signed\_req \leftarrow SIGN(req_{deAnon}, sk_{NIP})$ 
3  Submit  $(req_{deAnon}, signed\_req)$  to Deanonimization Contract
// Phase 2: Request Authorization
4  Smart Contract verifies  $sig_{NIP}$  using  $pk_{NIP}$ 
5   $status \leftarrow CHECKNIPSTATUS(NIP\_id)$ 
6  if  $status = WhiteListed$  then
7      Proceed to Phase 3
8  else if  $status = GrayListed$  then
9      Pause protocol, trigger manual review, Proceed to Phase 3
10 else
11     Abort deanonymization, log rejection reason, return Failure
// Phase 3: Seed Phrase Recovery
12 Smart Contract emits DeanonimizationEvent( $pid, pk_{NIP}$ )
13 foreach UIP Node  $n_i$  holding words for  $pid$  do
14     foreach word  $w_{ij}$  at index  $j$  held by  $n_i$  do
15          $enc\_word_{ij} \leftarrow ENCRYPT(w_{ij}, pk_{NIP})$ 
16         Write  $(pid, j, enc\_word_{ij}, SIGN(enc\_word_{ij}, sk_{n_i}))$  to blockchain
// Phase 4: Seed Phrase Reconstruction by NIP
17  $words[24] \leftarrow \emptyset$ 
18 for  $j \leftarrow 1$  to 24 do
19      $candidates_j \leftarrow \{enc\_word_{ij} \mid \text{index } j \text{ from all nodes}\}$ 
20     foreach  $enc\_word \in candidates_j$  do
21         Verify signature of providing node
22          $w \leftarrow DECRYPT(enc\_word, sk_{NIP})$ 
23         Add  $w$  to frequency count for index  $j$ 
24      $words[j] \leftarrow$  Most frequent word at index  $j$ 
25  $seed\_phrase \leftarrow CONCAT(words[1..24])$ 
26  $SID \leftarrow HASH(CONCAT(HASH(words[1]), \dots, HASH(words[24])))$ 
// Phase 5: Identities retrieval
27 Retrieve all SACs associated with  $SID$ 
28 foreach SAC do
29     Retrieve associated PK from  $(SAC : PK)$ 
30 return All PKs (anonymous identities)
    
```

---

agement Systems like OAuth 2.0 (Fett et al., 2016). The analysis focuses on adversaries aiming to compromise user anonymity through surveillance, misuse, or collusion (see fig. 2).

We consider three adversary categories:

- **External Observers:** Entities monitoring blockchain traffic and metadata to infer user identities (Meiklejohn et al., 2013).
- **Malicious Participants:** Users, validators, or providers engaging in fraud, identity replication, or unauthorized access (Bonneau et al., 2014).
- **Compromised Infrastructure:** Partial control

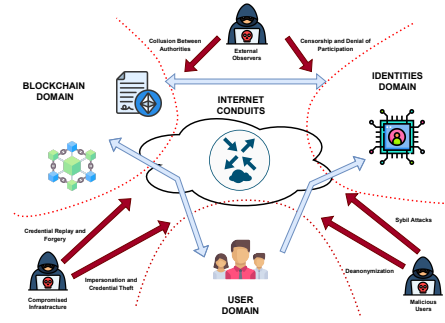


Figure 2: Threat model of the proposed protocol.

over issuers or smart contracts enabling surveillance, censorship, or collusion (Heilman et al., 2015).

Their capabilities include:

- Full-node monitoring and transaction graph analysis.
- Partial control of validators or identity issuers.
- Cross-correlation of metadata (timestamps, IPs) with blockchain events (Koshy et al., 2014).
- Coordinated activity among adversarial entities.

Blockchain-level attacks are excluded, as our model is agnostic to specific chain implementations.

#### 4.1 Attack Scenarios

We consider six representative threats:

- **Deanonimization:** Correlating pseudonyms and real identities via analysis (Meiklejohn et al., 2013).
- **Sybil Attacks:** Generating multiple identities to skew decision processes (Douceur, 2002).
- **Impersonation:** Using stolen or leaked credentials to act as another user.
- **Replay/Forgery:** Reusing or forging credentials for unauthorized access.
- **Authority Collusion:** Issuers conspiring to track or exclude specific users (Zyskind et al., 2015).
- **Censorship:** Preventing access to credentials via infrastructure control (Heilman et al., 2015).

We adopt the OWASP-based methodology for evaluating Indicators of Compromise (IoCs), based on four categories: threat agents, technical impact, vulnerability, and business impact. Each risk is rated LOW to HIGH depending on likelihood and severity.

Table 1: Risk Assessment Categories and Factors.

Risk Category	Factors
Threat Agent Factors	Skill level, Motive, Opportunity, Size
Technical Impact Factors	Confidentiality, Integrity, Availability, Accountability
Vulnerability Factors	Discovery, Exploitability, Awareness, Detection
Business Impact Factors	Financial, Reputation, Compliance, Privacy

## 4.2 Risk Scenarios

**(1) Anonymous Service Compromise.** Even if an anonymous platform is breached, user identities cannot be recovered. Risk: **Low**.

**(2) Consensus Attack Without Origin NIP.** Colluding nodes reconstruct an identity without the user’s NIP. SID exposure is possible, but PID remains secure – see fig. 3(a). Risk: **Low**.

**(3) Consensus Attack With Origin NIP.** If the original NIP joins the attack, full deanonymization is feasible. Still, coordination complexity keeps likelihood low – see fig. 3(b). Risk: **Medium-low**.

**(4) Full NIP Compromise.** A full breach of a NIP exposes all managed PIDs. This has **High** impact but very low probability, needing state-level access.

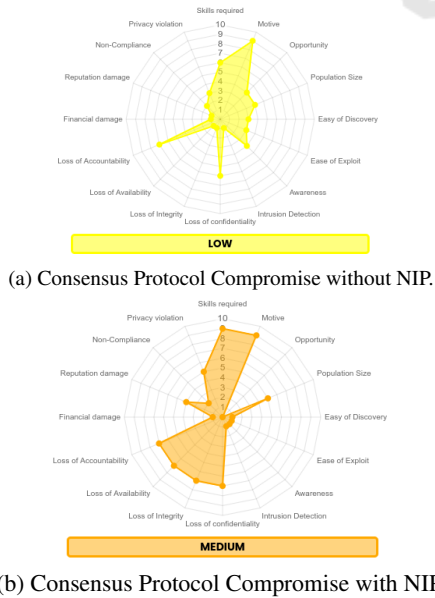


Figure 3: Risk assessment of different compromise scenarios.

## 5 ANALYTICAL MODEL

This section presents a compact mathematical model of the anonymization architecture, adapted from (Sciullo et al., 2024, §5–6), to evaluate the impact of structural parameters on system robustness. We focus on the risks of malicious attacks, node failures, and their combined effect, analyzing both per-user and system-wide behavior.

**Key Metrics.** We define three probabilities:

- $\pi_{\text{evil}}(q)$ : a coalition of  $q$  malicious nodes reconstructs a full seed phrase:

$$\pi_{\text{evil}}(q) = \left[ 1 - \frac{C(Q-q, M)}{C(Q, M)} \right]^N$$

- $\pi_{\text{fault}}(q)$ : phrase loss due to  $q$  node failures:

$$\pi_{\text{fault}}(q) = 1 - \left[ 1 - \frac{C(q, M)}{C(Q, M)} \right]^N$$

- $\pi_{\Theta}(q)$ : overall susceptibility to either risk:

$$\pi_{\Theta}(q) = 1 - [1 - \pi_{\text{evil}}(q)][1 - \pi_{\text{fault}}(q)]$$

These probabilities depend on:

- $Q$ : total UIP nodes (e.g., in EU,  $20 \leq Q \leq 40$ ),
- $M$ : word redundancy (default  $M = 5$ ),
- $N$ : number of seed words (default  $N = 24$ ),
- $q$ : compromised or faulty nodes ( $q \in [1, 10]$ ).

**Trade-offs and Parameter Tuning.** Increasing  $M$  improves fault tolerance but weakens resistance to attacks, while increasing  $N$  strengthens attack resistance but reduces fault tolerance. Simulations show that  $M = 5$  and  $N = 24$  strike a practical balance across reasonable values of  $q$  and  $Q$ .

**System-Wide Susceptibility.** Given  $T$  seed phrases and  $\pi_{\Theta}(q)$ , the likelihood that exactly  $k$  identities are compromised follows a binomial distribution:

$$P[\pi_{\Theta}](k) = C(T, k) \pi_{\Theta}^k (1 - \pi_{\Theta})^{T-k}$$

The cumulative distribution allows estimating the number of affected users. For  $T = 10^9$ :

- With  $Q = 20$ ,  $q = 2$ , fewer than 10 identities are compromised with 99% confidence;
- With  $Q = 40$ ,  $q = 4$ , the risk remains negligible (under 5 accounts at 99%);
- Only with  $q = 5$  or more do attacks affect thousands of users, a tiny fraction of total accounts.



The architecture's security hinges on tuning  $N$  and  $M$  against a target  $\pi_\theta$  based on acceptable risk levels. Under realistic assumptions, the system remains highly resilient even in the presence of partial compromise or failure.

## 6 CONCLUSIONS

In this paper, we introduced a resilient, ethical, and verifiable protocol that leverages blockchain technology to create authenticated anonymous identities for secure and private access to online services. We formalized the architecture through a mathematical model to evaluate its resilience against malicious attacks and faults, demonstrating its robustness and scalability for global adoption. Additionally, we analyzed potential threats, showing that its vulnerabilities are minimal. Future work will explore mechanisms for reconstructing all anonymous identities linked to a single user, credential recovery strategies, and real-world testbed implementation.

## ACKNOWLEDGEMENTS

This research is funded by the EU in the framework of the NGI Sargasso project, grant no. 101092887.

## REFERENCES

(2001). Uniting and strengthening america by providing appropriate tools required to intercept and obstruct terrorism (usa patriot act) act of 2001. Public Law 107-56.

(2011). ISO/IEC 24760-1:2011: Information technology – security techniques – a framework for identity management – part 1: Terminology and concepts.

Bonneau, J., Narayanan, A., Miller, A., Clark, J., Kroll, J. A., and Felten, E. W. (2014). Mixcoin: Anonymity for bitcoin with accountable mixes. In *International Conference on Financial Cryptography and Data Security*. Springer.

Douceur, J. R. (2002). The sybil attack. In *International Workshop on Peer-to-Peer Systems*. Springer.

EFF (2024). Anonymity.

eSSIF Lab (2022). European self sovereign identity framework laboratory.

European Parliament and Council of the European Union (2015). Directive (EU) 2015/849 of the European Parliament and of the Council.

Fett, D., Küsters, R., and Schmitz, G. (2016). A comprehensive formal security analysis of oauth 2.0. In *Proceedings of the 2016 ACM SIGSAC conference on computer and communications security*, pages 1204–1215.

Fredheim, R., Bay, S., Dek, A., Dek, I., and Singularex (2020). Social media manipulation report 2020. Report, NATO Strategic Communications Centre of Excellence.

Hardjono, T., Smith, N., and Pentland, A. S. (2014). Anonymous identities for permissioned blockchains.

Heilman, E., Kendler, A., Zohar, A., and Goldberg, S. (2015). Eclipse attacks on bitcoin's peer-to-peer network. In *24th USENIX Security Symposium (USENIX Security 15)*.

Karantaidou, I., Renawi, O., Baldimtsi, F., Kamarinakis, N., Katz, J., and Loss, J. (2024). Blind multisignatures for anonymous tokens with decentralized issuance. In *Proceedings of the 2024 on ACM SIGSAC Conference on Computer and Communications Security*, pages 1508–1522.

Koshy, P., Koshy, D., and McDaniel, P. (2014). An analysis of anonymity in bitcoin using p2p network traffic. In *Financial Cryptography and Data Security*. Springer.

Lafourcade, P., Mahmoud, D., Marcadet, G., and Olivier-Anclin, C. (2024). Transferable, auditable and anonymous ticketing protocol. In *Proceedings of the 19th ACM Asia Conference on Computer and Communications Security*, pages 1911–1927.

Mateescu, A., Brunton, D., Rosenblat, A., Patton, D., Gold, Z., and Boyd, D. (2015). Social media surveillance and law enforcement. *Data & Civil Rights*, 27:2015–2027.

Meiklejohn, S., Pomarole, M., Jordan, G., Levchenko, K., McCoy, D., Voelker, G. M., and Savage, S. (2013). A fistful of bitcoins: characterizing payments among men with no names. In *Proceedings of the 2013 ACM SIGCOMM conference on Internet measurement (IMC)*, pages 86–93.

Mosakheil, J. H. and Yang, K. (2024). Silentproof: Anonymous authentication with blockchain-backed offloading. In *Proceedings of the 19th ACM Asia Conference on Computer and Communications Security*, pages 1361–1377.

Rosenberg, M., White, J., Garman, C., and Miers, I. (2023). zk-creds: Flexible anonymous credentials from zk-snarks and existing identity infrastructure. In *2023 IEEE Symposium on Security and Privacy (SP)*, pages 790–808. IEEE.

Sciullo, L., De Marchi, A., Gigli, L., Palmirani, M., and Vitali, F. (2024). AAA: A blockchain-based architecture for ethical, robust authenticated anonymity. In *Proceedings of the 2024 International Conference on Information Technology for Social Good, GoodIT '24*, pages 1–9, New York, NY, USA. Association for Computing Machinery.

Shao, W., Jia, C., Xu, Y., Qiu, K., Gao, Y., and He, Y. (2020). Attrichain: Decentralized traceable anonymous identities in privacy-preserving permissioned blockchain. *Computers & Security*, 99:102069.

Shapiro, I. and Meyer, R. J. (2015). The right to anonymous speech and association.

Sporny, M., Guy, A., Sabadello, M., and Reed, D. (2022). Decentralized identifiers (dids) v1.0. Technical report, W3C.

UK (2023). Online safety act 2023.

Zyskind, G., Nathan, O., and Pentland, A. (2015). Decentralizing privacy: Using blockchain to protect personal data. In *Proceedings of the 2015 IEEE Security and Privacy Workshops (SPW)*.