ADVANCES IN CYBER-PHYSICAL SYSTEMS

Vol. 10, No. 2, 2025

BLOCKCHAIN-BASED ANONYMIZATION METHODS WITH SMART CONTRACT FOR DATA EXPIRY: TOWARD GDPR-COMPLIANT LIFECYCLE MANAGEMENT

Andrii Pavliv

Lviv Polytechnic National University, 12, S. Bandery str., Lviv, 79013, Ukraine,
Author's e-mail: *andrii.s.pavliv@lpnu.ua*

https://doi.org/10.23939/acps2025.02.173

Submitted on 11.09.2025

***Abstract*: This paper introduces a privacy-preserving framework for blockchain systems using the Smart Contract for Data Expiry (SCDE). SCDE governs data registration, retention, and erasure through on-chain policies and off-chain encrypted storage. It combines AES-256 encryption, a Key Management System (KMS) for cryptographic erasure, and Zero-Knowledge Proofs (ZKPs) for verifiable deletion without revealing data. Decentralized Identifiers (DIDs) enable pseudonymization and user accountability.**

Comparative results show that traditional and off-chain approaches lack automated, verifiable erasure. SCDE achieves full GDPR compliance with moderate overhead, demonstrating that privacy, transparency, and immutability can coexist in decentralized environments.

***Index terms:* blockchain, data anonymization, Smart Contract for Data Expiry (SCDE), Zero-Knowledge Proofs (ZKP), Key Management System (KMS), Decentralized Identifiers (DID), cryptographic erasure, GDPR compliance.**

I. INTRODUCTION

The exponential growth in the volume and sensitivity of personal data processed by organizations has significantly increased risks to confidentiality and integrity. Recent studies highlight that cyber-physical systems and digital platforms continuously accumulate heterogeneous data, intensifying the challenges of protecting user privacy [1]. Regulatory frameworks such as the General Data Protection Regulation (GDPR) establish strict principles of lawfulness, accountability, data minimization, and enforceable rights, including the "right to erasure" [2]. At the same time, blockchain environments, valued for their immutability, decentralization, and public verifiability—are being actively adopted in finance, healthcare, and smart city infrastructures [3-4]. However, the very immutability that ensures integrity and auditability of blockchain records creates a structural contradiction with GDPR requirements for data deletion and lifecycle control [5].

A variety of anonymization techniques have been explored to mitigate identifiability risks, including masking, pseudonymization, aggregation, and differential privacy [6-7]. Despite their widespread use, these methods often remain vulnerable to re-identification when auxiliary datasets are available [8]. Moreover, they lack machine-verifiable guarantees of data lifecycle management in decentralized systems, which limits their regulatory applicability [9]. Purely on-chain storage exacerbates the conflict with GDPR's erasure principle, while manual off-chain deletion depends on third-party trust and offers only limited transparency [10]. These shortcomings point to the urgent need for an integrated framework that unites automated retention policies, cryptographic erasure mechanisms, and verifiable audit trails, while remaining fully decentralized.

This article builds on previous research in block-chain-based anonymization architectures [11] and focuses on the methodological foundation of privacy-preserving lifecycle management. Specifically, it introduces a Smart Contract for Data Expiry (SCDE), a mechanism that encodes retention policies directly on-chain while coordinating encrypted off-chain storage. The SCDE enforces irreversible key destruction through a Key Management System (KMS) and records Zero-Knowledge Proofs (ZKPs) that confirm lifecycle events without revealing sensitive content [12-13]. In addition, Decentralized Identifiers (DIDs) and salted cryptographic hashes provide pseudonymization and integrity anchoring without disclosing personal attributes [14].

The objective of this study is to demonstrate the coherence and regulatory alignment of these methods, justify their integration into a unified framework, and show how SCDE operationalizes GDPR-compliant lifecycle management in immutable blockchain ledgers. The scope of the research includes the selection and integration of encryption schemes, KMS, DID, hashing, and ZKP, together with lifecycle automation via SCDE. Evaluation metrics such as proof latency, gas overhead, KMS throughput, and erasure success rate are discussed, alongside trade-offs including proof overheads, KMS robustness, and the complexity of system integration [15].

II. LITERATURE REVIEW AND PROBLEM STATEMENT

Recent scholarship has continued to investigate the structural tensions between blockchain immutability and data protection rights under the General Data Protection Regulation (GDPR). Zafar (2025) [3] analyzed this

contradiction in depth, highlighting that principles such as minimization, accountability, and the "right to be forgotten" remain inherently difficult to enforce on immutable ledgers. Similarly, Gołdyn et al. (2022) [5] examined architectural trade-offs, showing that compliance risks depend strongly on whether personal data is stored on-chain or off-chain. The European Data Protection Board (2025) [2] has issued updated recommendations for blockchain deployments, advocating designs that minimize exposure by delegating responsibilities to defined roles and shifting sensitive content off-chain.

More recent research has expanded on privacy-preserving computation in decentralized contexts. Quang et al. (2025) [8] demonstrated the use of homomorphic encryption to allow computations over encrypted data without decryption, balancing confidentiality and analytical utility. Guo et al. (2025) [12] proposed linkable ring signatures to enhance anonymity in blockchain transactions, while Bao et al. (2024) [9] integrated multiparty computation to increase resilience against inference attacks. In parallel, studies of differential privacy and anonymization models have highlighted the limitations of traditional techniques, including susceptibility to re-identification when auxiliary data is available, as shown by Monteiro et al. (2024) [6] and Shao et al. (2019) [7]. These techniques provide strong confidentiality but still lack lifecycle-level erasure guarantees.

Zero-Knowledge Proofs (ZKPs) have emerged as one of the most practical cryptographic mechanisms for reconciling confidentiality and verifiability. Shashidhara et al. (2024) [11] showed that ZKPs can transparently prove deletion events without revealing underlying data, while Capraz and Ozsoy (2021) [10] demonstrated the automation of compliance checks using ZKPs in GDPR-sensitive contexts. Sun et al. (2022) [14] introduced smart-contract-based cryptographic erasure schemes enabling verifiable deletion with low overheads. Despite these advances, current ZKP systems primarily address correctness of operations rather than holistic lifecycle automation.

Hybrid blockchain–off-chain models have also attracted significant attention. Boumaouche et al. (2020) [13] developed architectures integrating blockchain with external storage, storing only hashes on-chain to facilitate GDPR-compliant deletion. Broader reviews of blockchain decentralization and governance, such as Bodó et al. (2021) [4], have highlighted ongoing tensions between transparency, accountability, and privacy protections. Recent surveys of blockchain privacy mechanisms, including Zhang et al. (2023) [16], document advances in anonymization, erasure, and encrypted computation but emphasize fragmentation among existing approaches.

The notion of cryptographic erasure, defined as rendering ciphertext inaccessible through irreversible key destruction, has been revisited with stronger guarantees in recent years. Kumar et al. (2021) [15] analyzed blockchain key-management challenges and emphasized secure and auditable deletion as a critical unresolved issue. Complementary work on verifiable deletion and lifecycle auditability (e.g., Sun et al. (2022) [14]) demonstrates progress yet underscores the absence of unified automation frameworks.

Across these studies, three shortcomings are consistently observed. First, retention and erasure mechanisms are still predominantly manual or off-chain, limiting automation **(**Godyn et al., 2022 [5]; Boumaouche et al., 2020 [13]). Second, only a small number of systems provide publicly verifiable proofs of deletion, which are essential for regulatory audits (Sun et al., 2022 [14]; Shashidhara et al., 2024 [11]). Third, many solutions rely on centralized or semi-centralized key management, undermining the decentralization principle of distributed ledgers (Kumar et al., 2021 [15]). While emerging prototypes address subsets of these challenges, none yet integrate automated retention policies, cryptographic erasure, decentralized key management, and ZKP-based attestations within a unified lifecycle-governed architecture.

The present study addresses this gap by proposing a Smart Contract for Data Expiry (SCDE), which encodes retention policies directly on-chain, coordinates encrypted off-chain storage, initiates verifiable cryptographic erasure via decentralized KMS, and records ZKP attestations. This integration offers a regulator-aligned anonymization framework for blockchain systems that unifies automation, compliance, and transparency.

III. ABBREVIATIONS AND ACRONYMS

GDPR (General Data Protection Regulation): The European Union regulation on data protection and privacy, establishing principles such as lawfulness, transparency, and the right to erasure.

SCDE (Smart Contract for Data Expiry): A proposed smart contract model that enforces automated data retention, cryptographic erasure, and auditability in blockchain systems.

DLT (Distributed Ledger Technology): A family of technologies, including blockchain, that enable decentralized and immutable record-keeping.

KMS (Key Management System): A cryptographic subsystem responsible for secure generation, storage, distribution, and destruction of encryption keys.

ZKP (Zero-Knowledge Proofs): Cryptographic protocols that allow proving the validity of a statement without revealing the underlying data.

DID (Decentralized Identifier): A unique cryptographic identifier for entities in decentralized systems, used for pseudonymization and identity management.

AES (Advanced Encryption Standard): A widely used symmetric encryption standard for securing sensitive data.

IPFS (InterPlanetary File System): A decentralized off-chain storage system often used together with blockchain for storing large encrypted data objects.

IV. SCOPE AND OBJECTIVES

The scope of this research is defined by the methodological development and justification of anonymization techniques applied within blockchain systems. The focus is placed on the Smart Contract for Data Expiry (SCDE),

which coordinates the retention and erasure of encrypted personal data in compliance with the General Data Protection Regulation (GDPR).

The main objectives of the study are as follows:

1. Systematization of existing methods. To analyze data anonymization approaches such as masking, pseudonymization, aggregation, differential privacy, homomorphic encryption, and zero-knowledge proofs, emphasizing their applicability and limitations in decentralized environments.
2. Conceptual design of SCDE. To introduce a smart contract model that automates lifecycle management, integrates erasure policies, and enables audit transparency in distributed ledger systems.
3. Cryptographic justification. To substantiate the choice of encryption, key destruction, and zero-knowledge proofs as fundamental methods ensuring irreversible anonymization and verifiable compliance.
4. Hybrid architecture development. To propose a framework combining on-chain verifiability with off-chain encrypted storage, balancing scalability, privacy, and regulatory requirements.
5. Scientific contribution. To demonstrate that SCDE provides a novel methodological foundation for reconciling blockchain immutability with GDPR, offering a structured and auditable approach to anonymization.

The expected outcome of the research is a validated methodological framework that defines how anonymization techniques can be systematically applied to blockchain-based data management while ensuring both privacy and compliance.

V. METHODS

Conceptual Framework. The methodological foundation of this research is the Smart Contract for Data Expiry (SCDE)**,** a privacy-preserving mechanism that integrates blockchain technology with cryptographic anonymization techniques to achieve GDPR-compliant lifecycle management. SCDE encodes retention policies directly on-chain while coordinating encrypted off-chain storage and automated key destruction. This combination ensures that personal data can be erased in a verifiable manner while maintaining blockchain immutability and transparency. To validate its effectiveness, SCDE is evaluated against classical anonymization approaches through both theoretical analysis and comparative assessment.

Mathematical Model of Data Lifecycle. Let $D \in R^n$ - dataset containing personal information, k - symmetric encryption key generated by the Key Management System (KMS), T_{exp} - expiration time enforced by the Smart Contract for Data Expiry (SCDE), $E_k(D)$ - symmetric encryption function (AES-256 or Kalyna DSTU 7624), $E_k^{-1}(D)$ - corresponding decryption function, $\emptyset$ - the empty set, denoting irreversible key destruction.

The ciphertext is generated as:

$$C = E_k(D), D \in R_n . \quad (1)$$

The data remain accessible only while the key exists:

$$D = E_k^{-1}, if\ k \neq \emptyset . \quad (2)$$

Once SCDE enforces key destruction, decryption becomes impossible:

$$D = E_k^{-1}, is\ impossible\ if\ k \to \emptyset . \quad (3)$$

This property constitutes cryptographic erasure**,** ensuring compliance with GDPR ("right to be forgotten").

Zero-Knowledge Proofs for Erasure. To verify that encryption keys have been destroyed without revealing them, a Zero-Knowledge Proof (ZKP) is generated. Let $H(\cdot)$ be a cryptographic hash function (e.g., Keccak-256 or BLAKE3).

The erasure proof is constructed as:

$$\pi_{erase} = Prove\ pk, H\ k\ , \text{erasure confirmed} \quad (4)$$

and verified on-chain by:

$$Verify(pk, H(k), \pi_{erase}) = 1 \quad (5)$$

where pk is the public verification key. This provides an immutable, regulator-auditable log of erasure events.

Workflow of SCDE. The lifecycle of anonymization in SCDE can be formalized as a finite-state machine:

$$S = \{Registered, Active, Expiring, Erased\} \quad (6)$$

with transitions:

- Registered: D encrypted, metadata recorded on-chain.
- Active: D accessible until $t < T_{exp}$.
- Expiring: SCDE signals KMS to initiate $k \to \emptyset$.
- Erased: Key destruction completed, π_{erase} generated, proof logged on-chain.

This workflow guarantees transparent and auditable verification of each lifecycle event.

Comparative Analysis. To assess the novelty of the Smart Contract for Data Expiry (SCDE), its features were compared with common anonymization approaches in blockchain systems.

Traditional smart contracts ensure immutability but lack automated retention and cryptographic erasure, leaving compliance gaps.

Off-chain deletion enables partial removal but depends on storage provider trust and offers limited transparency.

ZKP-only models improve verifiability but lack lifecycle automation and often face scalability issues due to computational overhead.

In contrast, SCDE unifies these strengths: it automates retention by binding keys to policy rules, enforces erasure through cryptographic destruction, and employs Zero-Knowledge Proofs (ZKPs) for auditable compliance without data disclosure. Decentralized Identifiers (DIDs) further enhance pseudonymization and ownership management.

Evaluation was based on six criteria: automation, GDPR compliance, transparency, ZKP support, decentralization, and scalability. Results show that conventional methods address isolated aspects but lack completeness, while SCDE achieves balanced integration**,**

making it a regulator-aligned pathway for practical anonymization in blockchain ecosystems. Table 1 summarizes the comparative benchmarking.

Comparison of anonymization approaches in blockchain systems

Approach	Automation	GDPR Compliance	Transparency	ZKP Support	Decentralization	Scalability
Traditional Smart Contracts	–	Partial	Limited	–	✓	High
Off-chain Manual Deletion	Partial	Partial	–	–	–	Medium
ZKP-only Proof Models	–	Partial	✓	✓	?	Low
Proposed SCDE	✓	Full	✓	✓	✓	Medium–High

The comparative analysis confirms that no conventional method achieves full alignment with GDPR requirements while preserving decentralization. Traditional smart contracts and off-chain deletion offer only partial compliance, whereas ZKP-only models provide verifiability but at the cost of automation and scalability. By contrast, SCDE uniquely integrates automated lifecycle management, cryptographic erasure, and regulator-auditable ZKP proofs into a unified framework.

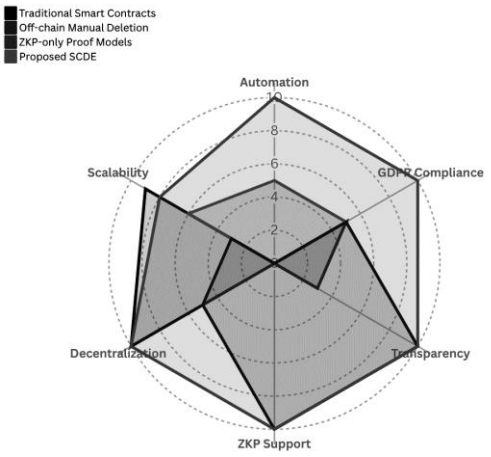
This positions SCDE as a practical and balanced solution that advances beyond existing approaches, ensuring both GDPR compliance and operational scalability in blockchain ecosystems.

Visualization of Comparative Characteristics. To illustrate the relative performance of the proposed Smart Contract for Data Expiry (SCDE), a radar chart was constructed (Fig. 1). The chart maps six evaluation dimensions-automation, GDPR compliance, transparency, ZKP support, decentralization, and scalability-which collectively represent both technical capability and regulatory alignment. Each approach (traditional smart contracts, off-chain manual deletion, ZKP-only models, and SCDE) was normalized to a 0-1 scale, where higher values indicate stronger support for the criterion.

Interpretation of Results. The visualization highlights several important insights:

1. Traditional smart contracts achieve strong decentralization and scalability but provide no meaningful automation, GDPR compliance, or ZKP support, leaving substantial regulatory gaps.

2. Off-chain manual deletion improves compliance marginally but relies heavily on trusted third parties, which undermines transparency and contradicts the decentralization principle.



Radar chart comparing anonymization approaches in blockchain systems.

3. ZKP-only models excel in verifiability and confidentiality but lack lifecycle governance and are limited by computational overhead, constraining scalability.

4. SCDE exhibits the most balanced profile, covering all six dimensions. By combining automated lifecycle management, cryptographic erasure, decentralized identifiers, and regulator-auditable proofs, SCDE achieves comprehensive GDPR compliance without sacrificing decentralization or scalability.

Overall, the radar chart provides visual confirmation that SCDE surpasses existing strategies by addressing their individual shortcomings. It demonstrates that lifecycle automation and verifiable erasure can be effectively integrated, establishing SCDE as a coherent and regulator-aligned solution for anonymization in blockchain environments.

VI. CONCLUSION

This study presented a method-centric framework for privacy preservation in blockchain systems, built around the Smart Contract for Data Expiry (SCDE). The proposed framework addresses the fundamental challenge of reconciling blockchain immutability with the GDPR "right to be forgotten" by combining automated retention policies, cryptographic erasure, and verifiable proofs of compliance. The main contributions can be summarized as follows:

- Development of a unified methodology that integrates symmetric encryption, decentralized key lifecycle control through a Key Management System (KMS), Zero-Knowledge Proofs (ZKPs) for verifiable erasure, and Decentralized Identifiers (DIDs) for pseudonymization and user control.
- A comparative analysis demonstrating that existing approaches - traditional smart contracts, off-chain deletion, and ZKP-only models - provide only partial solutions, whereas SCDE achieves a balanced integration of automation, transparency, and compliance.
- Formalization of anonymization processes through mathematical modeling and workflow definition, ensuring regulator-auditable lifecycle management.

Practical implications: SCDE offers regulators, system architects, and organizations handling sensitive data a coherent, regulator-aligned mechanism for lifecycle governance. It is particularly relevant to domains such as healthcare, financial services, and cyber-physical systems, where privacy protection and auditability are equally critical.

Limitations: As a methodological contribution, the framework has been presented in conceptual and comparative terms. Further work is needed to validate scalability in production environments and to test interoperability across heterogeneous blockchain ecosystems.

Future directions include optimizing ZKP protocols for lower verification cost, extending lifecycle management to multi-chain infrastructures, and exploring hardware-assisted secure enclaves to strengthen guarantees of cryptographic erasure.

In conclusion, SCDE demonstrates that GDPR-compliant anonymization and blockchain immutability are not mutually exclusive. By embedding lifecycle automation and cryptographic proofs into blockchain governance, SCDE provides a pathway toward trustworthy, privacy-preserving decentralized systems.

VII. CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

VIII. DECLARATION ON GENERATIVE AI

During the preparation of this work, the author(s) used ChatGPT, Grammarly in order to: Grammar and spelling check, Paraphrase and reword. After using this tool/service, the author reviewed and edited the content as needed and take full responsibility for the publication's content

References

[1] Tripathi, G., Ahad, M. A., & Casalino, G. (2023). A comprehensive review of blockchain technology: Underlying principles and historical background with future challenges. *Digital Applications and Technology*, 3, 100344. https://doi.org/10.1016/j.dajour.2023.100344

[2] European Data Protection Board (2025). *Guidelines 02/2025 on processing of personal data through blockchain technologies*. Official publication.

[3] Zafar, A. (2025). Reconciling blockchain technology and data protection laws. *Journal of Cybersecurity*, 11(1), tyaf002. https://doi.org/10.1093/cybsec/tyaf002

[4] Bodó, B., Brekke, J. K., & Hoepman, J. H. (2021). Decentralisation: A multidisciplinary perspective. *Internet Policy Review*, 10(2). https://doi.org/10.14763/2021.2.1563

[5] Godyn, M., Kedziora, M., Ren, Y., Liu, Y., & Song, H. H. (2022). Analysis of solutions for a blockchain compliance with GDPR. *Scientific Reports*, 12, 15021. https://doi.org/10.1038/s41598-022-19341-y

[6] Monteiro, S., Oliveira, D., António, J., Martins, P., & Abbasi, M. (2024). Data anonymization: Techniques and models. In *ICMarkTech 2022*. Springer. https://doi.org/10.1007/978-981-99-0333-7_6

[7] Shao, Y., Liu, J., Shi, S., & Zhang, Y. (2019). Fast de-anonymization of social networks with structural information. *Data Science and Engineering*, 4, 76–92. https://doi.org/10.1007/s41019-019-0086-8

[8] Quang, Q. N., Pham, A., & Nguyen, T. (2025). Privacy-Preserving Cyberattack Detection in Blockchain-Based IoT Systems Using AI and Homomorphic Encryption. *IEEE Internet of Things Journal*, 12(11), 16478–16492. https://doi.org/10.1109/JIOT.2025.3535792

[9] Bao, H., Yuan, M., Deng, H., Xu, J., & Zhao, Y. (2024). Secure multiparty computation protocol based on homomorphic encryption and its application in blockchain. *Future Generation Computer Systems*, 153, 22–35. https://doi.org/10.1016/j.heliyon.2024.e34458

[10] Capraz, S., & Ozsoy, A. (2021). Personal data protection in blockchain with zero-knowledge proof. In *Blockchain Technology and Innovations in Business Processes*. Springer. https://doi.org/10.1007/978-981-33-6470-7_7

[11] Shashidhara, R., Nair, R. C., & Panakalapati, P. (2024). Promise of Zero-Knowledge Proofs (ZKPs) for blockchain privacy and security. *Security and Privacy*, e461. https://doi.org/10.1002/spy2.461

[12] Guo, F., Gao, Y., Jiang, J., Chen, X., Chen, X., & Jiang, Z. (2025). Linkable Ring Signature for Privacy Protection in Blockchain-Enabled IIoT. *Sensors*, 25(12), 3684. https://doi.org/10.3390/s25123684

[13] Boumaouche, O., Ghenai, A., & Zeghib, N. (2020). Data-Oriented Blockchain: Off-chain storage with data-dedicated and prunable transactions. In *ACOSIS 2019*. Springer. https://doi.org/10.1007/978-3-030-61143-9_16

[14] Sun, S., Zhang, Y., Wu, Q., & Huang, Z. (2022). Blockchain-based verifiable data deletion using smart contracts and cryptographic erasure. *Computers & Security*, 120, 102806. https://doi.org/10.1016/j.cose.2022.102806

[15] Kumar, R., Ruj, S., & Nayak, A. (2021). Decentralized key management in blockchain-based systems. *IEEE Communications Surveys & Tutorials*, 23(4), 2154–2180. https://doi.org/10.1109/COMST.2021.3106764

[16] Zhang, X., Wang, T., Li, J., & Chen, L. (2023). A survey on blockchain-based privacy-preserving mechanisms: Cryptographic foundations and applications. *Information Sciences*, 621, 72–99. https://doi.org/10.1016/j.ins.2022.12.044

Andrii Pavliv was born in Vasyuchin, Ukraine, in 1997. He received the Bachelor's and Master's degrees in computer engineering, specializing in software engineering, at Lviv Polytechnic National University (NULP), Lviv, Ukraine.

He is currently an Assistant at the Department of Computer Engineering at NULP. With more than seven years of combined academic research and industry experience, his research interests include software engineering, artificial intelligence, and data security.