Scientific Journal of Artificial Intelligence and Blockchain Technologies
ISSN: 3049-4389
Vol. 1, Issue 3, Jul – Sep 2024 || PP. 29-38 https://doi.org/10.63345/sjaibt.v1.i3.104

GDPR Compliance Challenges in Blockchain-Based Systems

Dr Sandeep Kumar

SR University

Hasanparthy, Telangana 506371 India

er.sandeepsahratia@kluniversity.in

Date of Submission: 25-06-2024 Date of Acceptance: 27-06-2024 Date of Publication: 07-07-2024

ABSTRACT

Blockchain's decentralization, transparency, and tamper-resistance are celebrated properties for auditability and trust, yet they collide with core data protection duties under the EU General Data Protection Regulation (GDPR). This manuscript analyzes the principal compliance challenges that arise when blockchain processes personal data and proposes a practical, design-oriented framework to address them. First, we synthesize legal and regulatory positions on what counts as "personal data," the difference between anonymization and pseudonymization, and the implications of the right to erasure, data protection by design and by default, allocation of controller/processor roles, and international data transfers. We then map these requirements to blockchain architectures (public permissionless, public permissioned, and private permissioned) and data patterns (on-chain, off-chain, hybrid). Building on recent guidance from the European Data Protection Board (EDPB) and national authorities, we outline concrete technical and governance controls—off-chain storage and on-chain commitments, keyed hashing, encryption/key-revocation strategies, chameleon-hash/redactable-ledger designs, selective-disclosure credentials/zero-knowledge proofs, and robust consortium governance—to reduce risk and improve demonstrable compliance. Applying a six-step assessment methodology to three realistic use cases (NFT profile registry, supply-chain provenance, and consortium KYC), we show that while no single pattern fully reconciles immutability with erasure, practicable combinations can align processing with GDPR's principles of minimization, purpose limitation, storage limitation, and accountability. The paper concludes with a prioritized checklist for engineering "compliance-by-design" blockchains, and delineates scope and limitations for practitioners and researchers.

Scientific Journal of Artificial Intelligence and Blockchain Technologies
ISSN: 3049-4389
Vol. 1, Issue 3, Jul – Sep 2024 || PP. 29-38 https://doi.org/10.63345/sjaibt.v1.i3.104

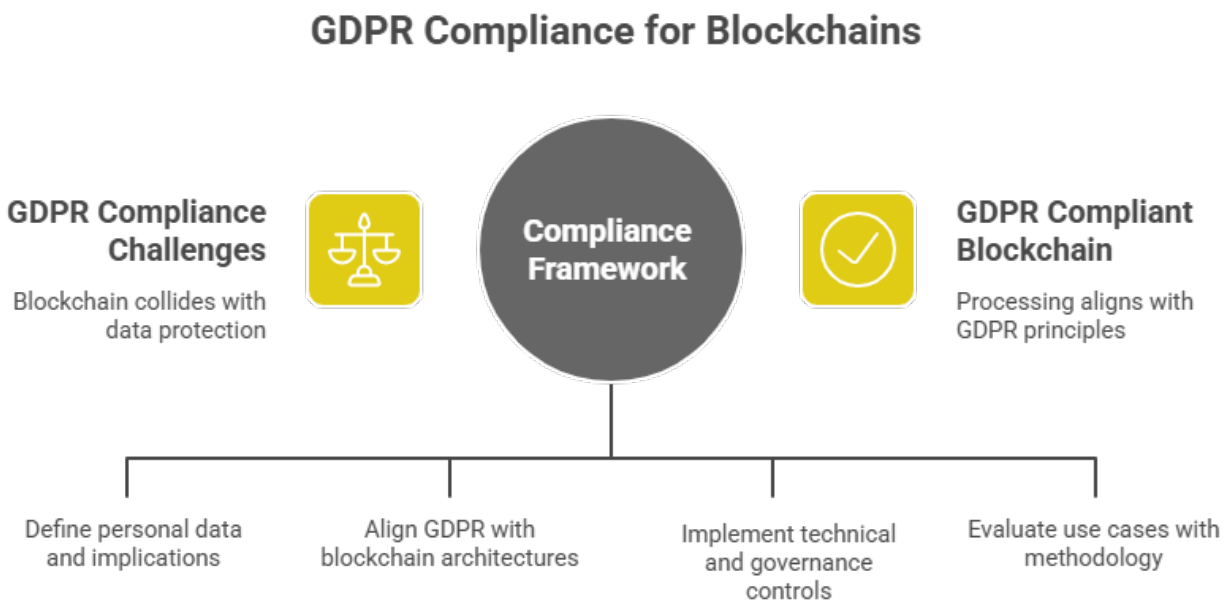

Figure-1.GDPR Compliance for Blockchains

KEYWORDS

GDPR, Blockchain, Right to Erasure, Pseudonymization, Anonymization, Data Protection by Design, Controller/Processor, Cross-Border Transfers, Off-Chain Storage, Zero-Knowledge Proofs

INTRODUCTION

The GDPR applies whenever information relates to an identified or identifiable natural person (personal data). Under Article 4(1), identifiability is interpreted broadly and includes indirect identifiers and online identifiers; consequently, even seemingly technical artifacts (addresses, transaction metadata, or linkable hashes) can become personal data if they can be tied back to an individual. The Regulation also embeds proactive duties like data protection by design and by default (Article 25) and reactive data subject rights such as rectification (Article 16) and erasure (Article 17). These provisions collide with public, append-only ledgers whose very purpose is to make records durable and widely replicated.

Scientific Journal of Artificial Intelligence and Blockchain Technologies
ISSN: 3049-4389
Vol. 1, Issue 3, Jul – Sep 2024 || PP. 29-38 https://doi.org/10.63345/sjaibt.v1.i3.104

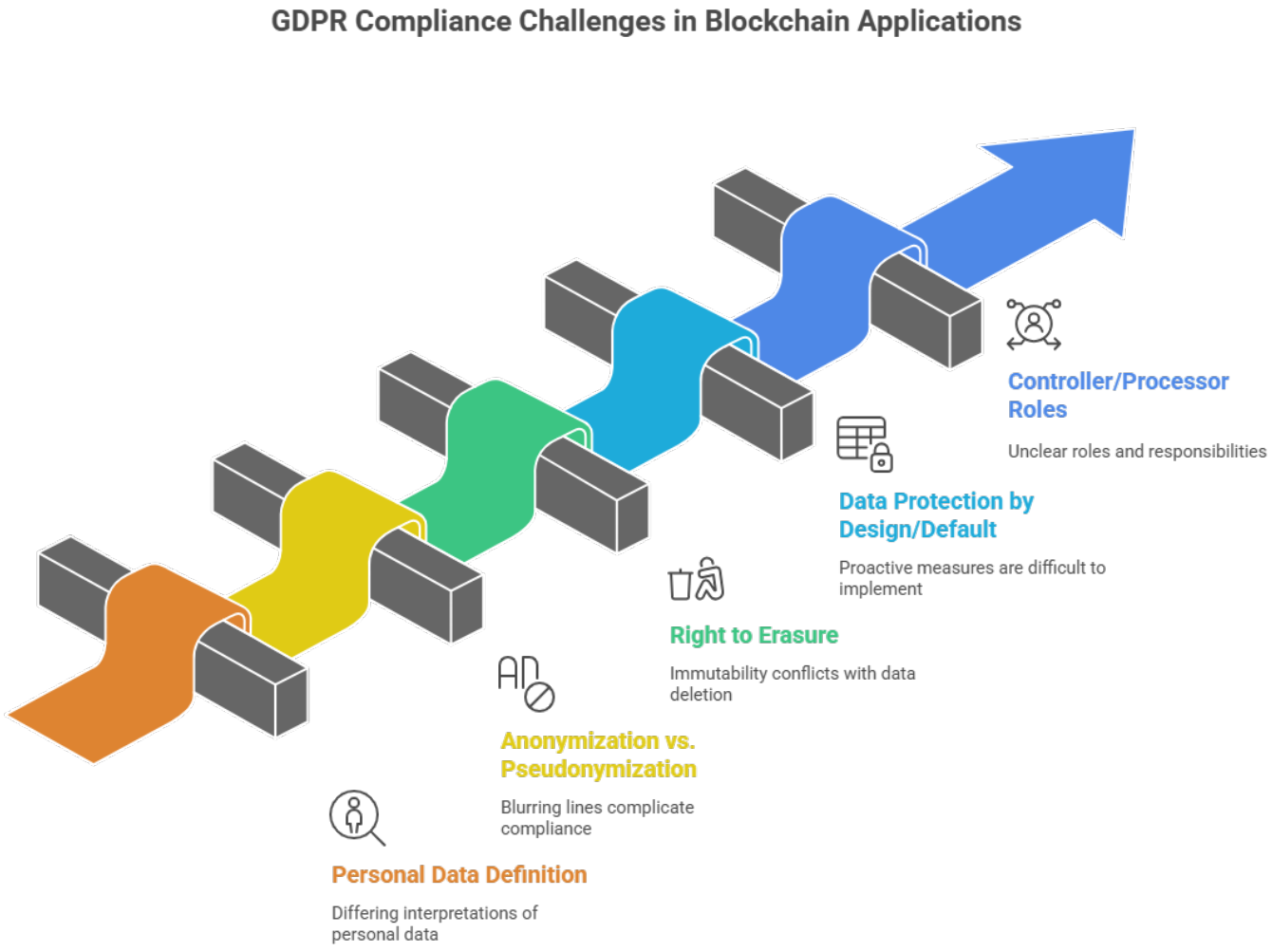

Figure-2.GDPR Compliance Challenges in Blockchain Applications

Recent regulators have sharpened their view of blockchain. In April 2025 the EDPB adopted Guidelines 02/2025 on processing personal data through blockchain technologies, clarifying role allocation, recommending off-chain storage, and calling for early Data Protection Impact Assessments (DPIAs) where high risk is likely. National authorities like France's CNIL have long warned that hashes and public keys often remain personal data (pseudonymized, not anonymized) and that controller responsibility cannot be "outsourced" to a protocol. Parallel guidance from the UK ICO stresses the distinction between pseudonymization and anonymization—pseudonymization reduces risk but does not remove data from GDPR's scope. Courts have also nuanced what counts as "personal data" in contexts involving pseudonymization and the recipient's realistic ability to re-identify, a trend that can affect how on-chain/off-chain splits are assessed.

This paper (i) synthesizes the legal tensions that arise from immutability, transparency, and global replication; (ii) proposes a practical methodology for GDPR-aware blockchain engineering; and (iii) reports application "results" by stress-testing three representative use cases.

Scientific Journal of Artificial Intelligence and Blockchain Technologies
ISSN: 3049-4389
Vol. 1, Issue 3, Jul – Sep 2024 || PP. 29-38 https://doi.org/10.63345/sjaibt.v1.i3.104

LITERATURE REVIEW

Early commentary recognized the structural tension between the GDPR's rights framework and blockchain immutability—particularly the right to rectification and erasure, and the difficulty of pinpointing a controller in decentralized networks. Practitioner analyses and authority guidance converged on several themes: keep personal data off-chain where possible; if unavoidable, use keyed hashing, commitments, or strong encryption; and prefer permissioned architectures with clear governance.

Regulators and expert bodies:

- **GDPR text and principles:** Articles 4, 17, 25, and Chapter V on international transfers define the baseline legal duties and principles (minimization, storage limitation, accountability) that any blockchain system must satisfy.
- **CNIL (France):** CNIL's landmark 2018 guidance explains roles (accessors, participants, miners), signals that participants who decide to write personal data are typically controllers, and recommends off-chain storage or keyed hashing when blockchain is necessary.
- **EDPB (EU):** The EDPB's 2025 Guidelines 02/2025 provide the most granular and up-to-date view: they analyze permissionless vs permissioned architectures, emphasize DPIAs, and advise controllers to avoid storing personal data directly on-chain, favoring salted/ keyed hashes and off-chain retrieval.
- **ICO (UK):** The ICO's anonymization guidance clarifies that pseudonymized data remain within data protection law; anonymization depends on context and reasonable means of re-identification.

Academic and policy analyses deepen these themes. Systematic reviews catalog the key pressure points: (1) exercising data subject rights on immutable records; (2) allocating controller/processor roles across heterogeneous node operators and smart-contract developers; and (3) applicable law and cross-border transfers in a multi-jurisdictional mesh. Technical proposals include redactable blockchains via chameleon hashes, though legal commentators note that redaction cannot purge historical replicas and often re-introduces governance trust assumptions. ENISA's series on pseudonymisation and data protection engineering provides patterns for risk-reduction but cautions that pseudonymisation does not equate to anonymization. fora (EU Blockchain Observatory & Forum) repeatedly call for clarity on the GDPR-blockchain intersection.

Finally, case-law developments around whether pseudonymized data constitute personal data for a recipient who lacks the key (e.g., General Court in *SRB v. EDPS*) illustrate a contextual turn that controllers may invoke when designing off-chain/ on-chain splits and determining the scope of transfers.

METHODOLOGY

We adopt a normative-technical methodology aimed at practitioners:

Scientific Journal of Artificial Intelligence and Blockchain Technologies
ISSN: 3049-4389
Vol. 1, Issue 3, Jul – Sep 2024 || PP. 29-38 https://doi.org/10.63345/sjaibt.v1.i3.104

(A) Regulatory Mapping: Identify GDPR obligations most sensitive to blockchain: definition of personal data (Art. 4), data subject rights (Arts. 15–18), data protection by design/default (Art. 25), and international transfers (Chapter V).

(B) Architectural Classification: Categorize the target system as:

1. **Public permissionless** (open participation, global replication, weakest control over geography/governance);
2. **Public permissioned** (open read, restricted write/validate);
3. **Private permissioned/consortium** (restricted read/write, contractual governance).

(C) Data Pattern Inventory: For each processing purpose, classify data flows as: on-chain content; on-chain pointer/commitment; off-chain content in controlled storage; and linkage secrets (keys/salts). Flag any on-chain personal data or personal-data inferences (e.g., linkage via analytics).

(D) Control Catalogue: Assemble technical and governance controls aligned to guidance: store personal data off-chain; if on-chain is unavoidable, use keyed hashing or commitments; encrypt robustly and manage keys for revocation; consider redactable patterns (with governance); apply selective-disclosure credentials/zero-knowledge techniques; implement strict role governance, logging, and access controls; and plan DPIAs and incident response.

(E) Risk/Compliance Scoring: Score each processing purpose against GDPR principles (lawfulness, minimization, storage limitation, integrity/confidentiality, accountability) using a simple ordinal scale (0–3) and document mitigations.

(F) Use-Case Stress Test: Apply the framework to representative scenarios to produce "results" on feasibility, residual risk, and compliance posture.

This methodology reflects recent EDPB expectations: define roles, avoid direct on-chain personal data, and embed DPIAs early.

RESULTS

Use Case 1: NFT Profile Registry (Public Permissionless)

Context

A consumer-facing dApp mints NFTs representing user profiles; on-chain metadata include display name and links to social handles.

Findings

- **Personal data scope:** Public wallet addresses and linkable profile metadata constitute personal data when reasonable re-identification is possible (e.g., handles that point to a person).

- **Right to erasure & rectification:** Immutable token metadata resist deletion or correction. Even if a "burn" function hides references, historical state persists on full nodes and archival services. Redactable/chameleon-hash techniques could enable post-facto edits but demand centralized governance and may not purge historical replicas.
- **Data minimization by design:** The compliant pattern is to store personal data **off-chain** in a controlled store; keep only a keyed hash/commitment on-chain so that the controller can delete or correct off-chain records and rotate keys. The EDPB expressly recommends off-chain storage and salted/keyed hashes where blockchain is necessary.
- **International transfers:** Global replication across non-EU nodes combined with personal data on-chain triggers transfer rules (Chapter V). With off-chain storage in the EEA and on-chain commitments that do **not** enable re-identification by foreign nodes, the transfer risk narrows.

  Residual risk: Medium–high on public chains unless data are confined to off-chain stores with on-chain commitments and robust key management.

Use Case 2: Supply-Chain Provenance (Private Permissioned)

Context

A consortium tracks product provenance. Participants include EU and non-EU manufacturers, logistics providers, and auditors.

Findings

- **Role allocation:** Consortium members who jointly determine purposes/means are **joint controllers**; node operators processing on behalf of the consortium may be **processors**, necessitating Article 28 agreements. EDPB and CNIL urge explicit allocation of responsibilities in consortium charters.
- **Data design:** Most records can be non-personal (SKUs, batch hashes). Personal data (e.g., driver IDs) must be kept off-chain; on-chain entries store commitments/pointers. ENISA's pseudonymisation patterns help manage linkage risk across domains.
- **Erasure & storage limitation:** Off-chain stores enforce retention; rotating or destroying the mapping secret (keys/salts) diminishes linkage risk. ICO guidance warns that pseudonymization alone does not make data anonymous; thus retention and deletion must operate on the off-chain layer.
- **Transfers:** A private permissioned ledger with EU-hosted nodes and standard contractual controls can substantially attenuate Chapter V issues compared to public chains.

  Residual risk: Low–medium with disciplined governance, minimization, and off-chain controls.

Use Case 3: Consortium KYC Utility (Public-Read / Permissioned-Write)

Context

Banks write attestations ("KYC completed", "sanctions-screened") to a shared ledger; counterparties read proofs.

Findings

- **Lawful basis & purpose limitation:** Clear legal bases (legal obligation/legitimate interests) and strict purpose limitation are essential. On-chain attestations should avoid attributes that identify a person; prefer selective-disclosure or zero-knowledge proof systems to confirm facts without revealing raw personal data.
- **Pseudonymization vs anonymization:** Storing hashed identifiers is still personal data if a controller can re-link; whether a **recipient** without keys holds personal data depends on realistic re-identification means—a nuance highlighted by recent EU case-law.
- **DPIA & accountability:** Given scale and sensitivity, a DPIA is expected; controllers should record threat models for re-identification, linkage attacks across datasets, and governance-level failures. EDPB guidance makes DPIAs and role clarity explicit expectations.

  Residual risk: Medium; acceptable with selective-disclosure credentials, clear roles, and strict off-chain data custody.

DISCUSSION

Immutability vs. Erasure/Rectification

Article 17 remains the sharpest friction point. Off-chain storage with on-chain commitments enables functional deletion and correction; "key-destroy" and pointer-invalidations reduce residual risk, while redactable-blockchain designs add a last-resort edit path at the cost of governance complexity and potential fragmentation (old replicas).

Transparency vs. Data Minimization

Public blockchains' default transparency clashes with "need-to-know" exposure. Privacy-preserving primitives (zero-knowledge proofs, selective disclosure) allow verification of facts without publishing personal data. Regulators emphasize building such protections in at design time (Article 25; EDPB 2025).

Decentralization vs. Accountability

"Code is law" is not a defense under GDPR. Someone determines the purposes and means—often the application developer, consortium, or the participant submitting a transaction—and thus assumes controller duties; others may be processors or joint controllers. CNIL and the EDPB press for explicit role mapping and contracts.

Global Replication vs. Chapter V

If personal data appear on public chains with non-EU nodes, every block broadcast is potentially a transfer. Minimization plus off-chain custody within the EEA (or with appropriate safeguards) is the safer route.

Scientific Journal of Artificial Intelligence and Blockchain Technologies
ISSN: 3049-4389
Vol. 1, Issue 3, Jul – Sep 2024 || PP. 29-38 https://doi.org/10.63345/sjaibt.v1.i3.104

Pseudonymization vs. Anonymization

ICO and ENISA emphasize that pseudonymization reduces risk but keeps GDPR in scope; some case-law nuances whether a recipient **without** the key holds personal data, underscoring context-sensitive analysis—useful for hybrid patterns.

A Practical Compliance-by-Design Playbook

1. **Challenge the Need for a Blockchain:** Choose it only when auditability/disintermediation are essential; otherwise, conventional databases with append-only logs may better satisfy GDPR (minimization/storage limitation).
2. **Prefer Permissioned Architectures:** They offer clearer role assignment, controllable geography, and configurable privacy. CNIL and EDPB lean toward permissioned patterns where personal data are in scope.
3. **Keep Personal Data Off-Chain:** Put only commitments or keyed hashes on-chain; store raw data in controlled repositories with retention/deletion enforcement. EDPB 2025 explicitly recommends this.
4. **Use Strong Pseudonymisation & Cryptography:** Apply long, random, per-record salts or secret keys; rotate keys; segregate mapping tables; and monitor re-identification risk. ENISA provides actionable patterns.
5. **Engineer for Erasure and Correction:** Design pointer invalidation, key destruction, or redactable patterns (with governance) to achieve functional RTBF where feasible, and document limits candidly.
6. **Define Roles and Paper the Governance:** Name the controller(s); appoint processors; execute Article 28 terms; establish a consortium charter with incident response, DPIA cadence, and DSR (data subject request) playbooks.
7. **Manage International Transfers:** Keep nodes in the EEA when possible; if not, implement Chapter V tools (adequacy, SCCs) and assess on-chain elements to ensure they are non-identifying for foreign recipients.
8. **Adopt Selective-Disclosure Credentials/ZKPs:** Replace raw attributes with verifiable claims or proofs; log proofs, not personal data.
9. **Run DPIAs Early and Often:** Treat architectural changes (e.g., new smart-contract modules) as potential DPIA triggers per EDPB expectations.
10. **Document Everything:** Accountability is a first-class requirement: keep design records, threat models, and DSR handling logs.

CONCLUSION

GDPR compliance in blockchain systems is neither automatic nor impossible. The sharpest frictions—erasure in the face of immutability, minimization under default transparency, controller accountability in decentralized topologies, and cross-border replication—are tractable when system designers treat privacy as a core architectural invariant rather than an afterthought. The EDPB's 2025 guidelines, combined with long-standing CNIL, ICO, and ENISA positions, set clear expectations: avoid direct on-chain personal data; use off-chain storage plus cryptographic references; assign roles explicitly; and prove (with DPIAs and documentation) that re-identification risks are controlled.

In practice, permissioned or hybrid architectures with off-chain custody, selective-disclosure credentials, robust key management, and governance that supports data subject rights offer the best path to compliance. Purely public, permissionless deployments that publish personal data are the hardest to reconcile with GDPR and should be avoided for regulated processing. Continued standardization—especially around redactable ledgers, verifiable deletion proofs, and interoperable DSR tooling—will further narrow the gap, but today's compliance-by-design is already achievable with prudent engineering and governance choices.

SCOPE AND LIMITATIONS

Scope

This paper focuses on GDPR-specific issues for blockchain systems that **process personal data**. It assesses mainstream architectures and controls that are currently implementable, drawing on authoritative guidance (EDPB 2025; CNIL; ICO) and widely cited technical/policy analyses.

Limitations

First, jurisprudence on pseudonymization and contextual identifiability is evolving; organizations must monitor decisions that refine when a recipient holds "personal data." Second, guidance referenced here is EU-centric; non-EU regimes (e.g., UK GDPR, state privacy laws) are related but not identical. Third, this is a conceptual evaluation rather than an empirical measurement of re-identification risk; real-world adversaries and data linkages can behave unpredictably. Fourth, redactable-ledger techniques remain emergent and may introduce governance trade-offs that regulators scrutinize differently across Member States. Finally, cryptography and protocol ecosystems evolve quickly; DPIAs should be iterative to reflect new threats and mitigations.

REFERENCES

- *Arthur Cox. (2025, May 29). Personal data on the chain: EDPB guidelines for blockchain technologies. https://www.arthurcox.com/knowledge/personal-data-on-the-chain-edpb-guidelines-for-blockchain-technologies/*
- *CNIL. (2018, October 29). Blockchain and the GDPR: Solutions for a responsible use of the blockchain in the context of personal data. https://www.cnil.fr/en/blockchain-and-gdpr-solutions-responsible-use-blockchain-context-personal-data*
- *CNIL. (2018). Blockchain and the GDPR (English PDF guidance). https://www.cnil.fr/sites/default/files/atoms/files/blockchain_en.pdf*
- *Dechert LLP. (2023, May 25). EU General Court examines data anonymisation and pseudonymisation (SRB v EDPS). https://www.dechert.com/knowledge/onpoint/2023/5/eu-court-examines-data-anonymisation-and-pseudonymisation.html*
- *ENISA. (2019). Pseudonymisation techniques and best practices. https://www.enisa.europa.eu/publications/pseudonymisation-techniques-and-best-practices*
- *ENISA. (2021). Deploying pseudonymisation techniques: Guidance and use cases. https://collab.dpa.gr/wp-content/uploads/2023/07/enisa_DEPLOYING-PSEUDONYMISATION-TECHNIQUES_en.pdf*
- *EDPB. (2025, April 8). Guidelines 02/2025 on processing of personal data through blockchain technologies (Version for public consultation). https://www.edpb.europa.eu/system/files/2025-04/edpb_guidelines_202502_blockchain_en.pdf*
- *EDPB. (2025, April 14). EDPB adopts guidelines on processing of personal data through blockchains (News release). https://www.edpb.europa.eu/news/news/2025/edpb-adopts-guidelines-processing-personal-data-through-blockchains-and-ready_en*
- *European Parliament & Council. (2016). Regulation (EU) 2016/679 (General Data Protection Regulation). https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng*

- *GDPR-Text.com. (n.d.). Article 17 GDPR: Right to erasure (right to be forgotten). https://gdpr-text.com/en/read/article-17/*
- *GDPR-Info.eu. (n.d.). Article 25 GDPR: Data protection by design and by default. https://gdpr-info.eu/art-25-gdpr/*
- *GDPR-Info.eu. (n.d.). Article 4 GDPR: Definitions. https://gdpr-info.eu/art-4-gdpr/*
- *GDPR-Info.eu. (n.d.). Article 44 GDPR: General principle for transfers. https://gdpr-info.eu/art-44-gdpr/*
- *ICO. (2025, March 28). Anonymisation and pseudonymisation guidance (About this guidance). https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/data-sharing/anonymisation/about-this-guidance/*
- *ICO. (n.d.). Pseudonymisation. https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/data-sharing/anonymisation/pseudonymisation/*
- *Lyons, T., Courcelas, L., & Timsit, K. (2018). EU Blockchain Observatory & Forum: Blockchain and the GDPR (Workshop report). https://afyonluoglu.org/PublicWebFiles/Reports/Blockchain/EU//20180608-EU%20Blockchain%20Forum-GDPR%20Report.pdf*
- *Oxford Business Law Blog. (2018, April 20). Blockchains and the right to be forgotten. https://blogs.law.ox.ac.uk/business-law-blog/blog/2018/04/law-and-autonomous-systems-series-blockchains-and-right-be-forgotten*
- *Pinsent Masons. (2025, April 15). ICO anonymisation guide aids UK data protection compliance. https://www.pinsentmasons.com/out-law/analysis/ico-anonymisation-guide-uk-data-protection-compliance*
- *Zafar, A. (2025). Reconciling blockchain technology and data protection laws: A closer look at the GDPR. Journal of Cybersecurity, 11(1). https://academic.oup.com/cybersecurity/article/11/1/tyaf002/8024082*
- *Belen-Saglam, R., Altuncu, E., Lu, Y., & Li, S. (2023). A systematic literature review of the tension between the GDPR and public blockchain systems. Digital Communications and Networks, 9(4), 1223–1246. https://www.sciencedirect.com/science/article/pii/S2096720923000040*