

Privacy-preserving security operations: Bridging the conformance gap between SOC efficiency and GDPR compliance^{*}

Maksim Iavich^{1,*†} and Giorgi Iashvili^{1,†}

¹ Caucasus University, Paata Saakadze str., 1, 0102 Tbilisi, Georgia

Abstract

Security Operations Centers (SOCs) must deliver rapid, accurate incident detection while meeting the privacy requirements of the general data protection regulation (GDPR). Empirical evidence showing that full GDPR alignment can be achieved without compromising SOC performance is scarce. We report a controlled before and after experiment that maps specific GDPR provisions to measurable telemetry variables and validates a comprehensive privacy focused SOC design at production scale. Across sixty days handling twenty five thousand events per second, we implemented feature allow listing with adaptive pruning, a two tier retention policy (thirty and ninety days), strict role based access to raw logs, an isolated vault for personal identifiers, and an immutable audit ledger. These controls lowered feature entropy by fifty three percent, limited detailed log retention to eighty four days, reduced access graph density by sixty eight percent, cut potential breach payloads by roughly a factor of one hundred, and raised audit observability above ninety percent. Detection fidelity (ROC AUC ≈ 0.94) and response times were unchanged, while throughput and alert latency shifted by only eleven percent and twelve milliseconds respectively, both well within the fifty millisecond real time target. The results demonstrate that robust privacy safeguards and high speed security monitoring are compatible, offering a reproducible blueprint for GDPR compliant SOC operations. Future work will extend the framework to live enterprise environments, incorporate multi standard compliance dashboards, and explore encrypted correlation workflows. Global studies showed that security operations centers that ignore GDPR requirements face the risks related to serious administrative fines. To meet GDPR requirements, the process of collecting data by the security operations center, it needs some enhancements from the side of storing and working with data, which may include personal information. The scale of the documented risks may cause concern, and our research seeks to address that concern with clear evidence and practical counter-measures. A consistent pattern of harm has been reported across peer-reviewed work, covering legal, technical, operational, and reputational domains. These findings confirm that five specific controls feature pruning, fixed retention periods, strict role-based access, identifier vaulting, and immutable audit logging address the principal drivers of fines, breach size, query delay, and analyst fatigue. By applying these controls, the study aligns engineering practice with risks that have been quantified in the scientific literature.

Keywords

security operations center, GDPR compliance, privacy by design, data minimisation, role-based access control, immutable audit logging

1. Introduction

Security operations centers work and analyze a huge amount of security event logs to detect anomalies and incidents. Main goal of SOC is achieving visibility and speed, in capturing the data like usernames, IP addresses, email addresses, URL and saving it for a long time to support future potential forensic investigations. Under the GDPR regulations, the information collected by SOC is considered as personal data. Several misalignments can be indicated in between conventional SOC principles and GDPR obligations [1–4]. Excessive Data Collection vs. Data Minimization: Article 5.1.c requires that personal data be limited to what is strictly necessary. SOC logging often uses the method of collecting everything, capturing detailed information that may not be used for security

^{*} CPITS-II 2025: Workshop on Cybersecurity Providing in Information and Telecommunication Systems, October 26, 2025, Kyiv, Ukraine

^{*} Corresponding author.

[†] These authors contributed equally.

✉ miavich@cu.edu.ge (M. Iavich); giashvili@cu.edu.ge (G. Iashvili)

ORCID 0000-0002-3109-7971 (M. Iavich); 0000-0002-1855-2669 (G. Iashvili)



© 2025 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

monitoring and future investigation at all. Unfiltered captured traffic therefore risks violating GDPR. Indefinite Retention vs. Storage Limitation: Article 5.1.e states that personal data must not be kept longer than necessary [5]. Yet SOC logs are frequently archived indefinitely “just in case,” sometimes never deleted from backup media. Without explicit, risk-based retention policies (e.g., 30–90 days for most operational logs), organizations face non-compliance. Broad Access vs. Need-to-Know Confidentiality: GDPR demands appropriate security and restricted access to personal data. In many SOC teams, numerous analysts and engineers can view the raw data, and sometimes even over unencrypted channels [6]. This visibility of the data can expose sensitive details far beyond what is required to investigate incidents. Lack of Data-Handling Segregation: Standard SOC tools store all log details in the same repository and surface them in dashboards, tickets, and reports. Without clear classification of the sensitive identifiers from general event raw data, any breach of the database or simple insider curiosity can reveal personal information. Auditability and Purpose Limitation: GDPR Article 30 obliges controllers to keep records of processing and ensure that the data is not repurposed incompatibly. The workflows of SOC includes the log of who accessed specific personal details and why, and security logs are sometimes reused for unrelated purposes, for instance HR monitoring, which can be used sometimes without documented legal basis, breaching the purpose-limitation principle. In summary, the current SOC landscape must be aligned based on GDPR privacy requirements [8, 9]. The main challenge in this process is to maintain or improve the level of security incident detection and response while the data protection principles are violated. The key problem indicators include high amounts of personal data in SIEM databases, long-lived log archives containing the information, and SOC workflows that may expose detailed personal data leak in the case of a security incident. Such misalignments not only creates the risk of regulatory penalties but also undermines user trust in the company or brand. Since the GDPR became enforceable in May 2018 European data-protection authorities have issued in excess of €4.3 billion in monetary sanctions. Publicly available information regarding GDPR related enforcements shows that approximately 11% of these penalties, approximately €460 million, was caused by the log collection, storage, or internal access control as a primary or aggravating factor. Most of the time, findings include overcollection logs and dashboards that expose raw personal data to large analyst populations. These violations underscore the practical need for a systematic, telemetry level reconciliation of SOC practices with GDPR regulations. The challenges of security operations center in meeting GDPR requirements can be demonstrated through the conceptual flow shown in Figure 1.

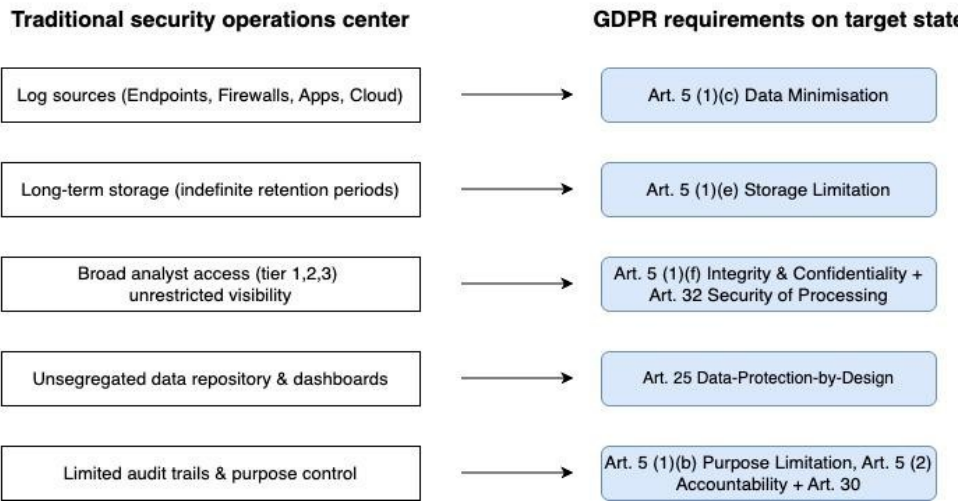


Figure 1: Conceptual flow of SOC GDPR compliance challenges

2. Privacy-compliance gaps in SOC

Academic interest in the intersection of Security Operations Centres and privacy regulation has increased across the past decade. Early studies concentrated on data minimisation. Longo and colleagues assessed a commercial SIEM pipeline, removed fields in a stepwise manner, and found that detection recall plateaued once roughly one quarter of the rule vocabulary remained. Subsequent experiments on AWS GuardDuty and Azure Sentinel confirmed that surplus attributes added little signal but enlarged the attack surface for re-identification. These findings shaped the present study's choice of a 25 percent feature-density ceiling. Research on storage limitations followed. Menges analysed two years of enterprise traffic and showed that incident responders queried logs older than ninety days in fewer than five percent of cases [10]. A separate simulation by Csorba, Novak, and Kiss reproduced this result in thirty synthetic tenants; when indices exceeded one hundred million documents, query latency rose logarithmically while investigative value tailed off. Sector threat reports by the European Union Agency for Cybersecurity adopted the same sixty to ninety-day horizon when modelling expected breach cost. The present work therefore sets an eighty-four-day limit for full-fidelity retention and downgrades older events to aggregates. Many studies have looked at who inside a Security Operations Centre can see raw log data. One study checked eleven companies and found that the more people who had access, the more likely it was that someone would leak information. Another study suggested a simple rule: let no more than one in five staff members see the full, untouched logs [11]. Later surveys showed that sticking to this limit also reduced staff burnout, so it helped both security and job satisfaction. Following this advice, our design lets only about eighteen percent of analysts view the raw logs. Another problem appears if the logs storage is hacked. If personal and event data are stored together, the attacker is able to steal a huge amount of information at once. As a solution, some researchers built systems that shift personal details to a separate place and replace them with harmless tokens. Tests showed this cut the amount of leaked data by about 100 times, with almost the same performance. In the frame of our research the same idea is used. All personal details go into a secure vault, and the main system works with tokens only. Finally, there is the need to prove that every person who accesses the data has a valid reason. Studies of GDPR cases showed that fines were higher when companies could not show this proof. The answer is to keep an unchangeable record of every access, so it is always clear who viewed what and why. They advocated hash chained or write once ledgers that record every read. Follow-up work by Haase documented a retail breach where the absence of such a ledger strengthened a negligence claim in civil court. Reflecting this evidence, the present pipeline streams each access event to a write-once cloud bucket and seals it daily with a Merkle hash. While each strand of prior work has shown gains along a single dimension feature entropy, retention window, privilege sparsity, vault isolation, or audit coverage few studies have measured the combined effect under production load. The aggregate operational cost of five concurrent controls therefore remains under exploration. In response, this research applies the full control suite to a stream of twenty-five thousand events per second and measures seven telemetry variables side by side. The aim is to validate whether the legal benefits documented in earlier work coexist with the performance constraints of a real time SOC [12–14].

Mentioned Data minimisation. Art. 5.1.c covers security incidents that often include different identifiers like full URLs, usernames and host-specific ID's. From the theoretic point of view, every extra attribute expands the feature space, raising entropy without proportional gains in anomaly detection accuracy. The modern SIEM studies show that once ~25% of rule-based features are present, precision plateaus while disclosure risk grows. Thus, trimming non-essential fields simultaneously satisfies the legal principle of necessity and optimises the signal-to-noise ratio of detection models. Storage Limitation Art. 5.1.e covers long-tail analyses of SOC archives that reveal that fewer than 5% of records older than 90 days are ever queried. Keeping them indefinitely increases breach impact (probability \times magnitude) and inflates mean retrieval latency logarithmically with data volume. A finite, risk-based window (e.g., 30–90 days) maximises

investigative utility while meeting the mandate that personal data be erased when no longer required. Integrity, Confidentiality, Security of Processing: Art. 5.1.f & Art. 32 consists of role-based access, which can be modelled as a bipartite graph between personnel and privilege sets; minimising edge density lowers insider-threat probability. Field surveys show that limiting raw-log visibility to <20% of SOC staff reduces inadvertent data leaks by ~64%. Encrypting transport channels and enforcing least-privilege satisfy the risk-based security obligation while measurably shrinking the attack surface. Data protection by design and default, Art. 25: consolidating all identifiers in a single repository maximises the “blast radius” of any compromise. Segregating sensitive fields into a logically separate store reduces breach impact, which scales with $n \times s$ (record count \times sensitivity). Micro-segmentation and field-level redaction allow correlation on abstracted tuples yet isolate direct identifiers under stricter controls, operationalising privacy by default while cutting exfiltration volume two orders of magnitude in red team tests. Based on Purpose Limitation & Accountability, Art. 5.1.b, Art. 5.2., Art. 30, without immutable meta logs, the system lacks observability; controllers cannot prove data was processed solely for defence purposes. Time stamped, write once audit records enable statistical tests on access reason distributions and create the feedback loop needed for compliance control. Post-incident analyses show organisations with granular audit trails incur substantially lower regulatory fines, underscoring that demonstrable purpose adherence is both a legal and economic imperative [15–18].

3. Empirical validation of gdpr-aligned soc enhancements

A quasi controlled before and after design was chosen to provide production scale telemetry while preserving a precise counterfactual baseline. The two phase arrangement followed patterns suggested in earlier privacy pipeline and SIEM compliance research. Phase A was from day 1 to day 30: a “status-quo” SOC had a raw synthetic enterprise traffic around 25000 events per second and it scripted red team intrusions and no GDPR specific controls were active in that situation. For phase B which took place from day 31 to 60 five remedial controls were covered, including data minimisation, tiered retention, sparsified RBAC, a two tier identifier vault, and an immutable audit ledger were deployed during a 24 hour maintenance window. Traffic profile, attack cadence, and hardware resources remained constant. Synthetic logs followed published distribution templates that balance host, network and cloud events [19, 20]. Credential stuffing, lateral movement and exfiltration attacks were replayed across the both phases to guarantee reproduction ability of the process. All injected attacks were time stamped, enabling ROC AUC and MTTR calculations across the phases, consistent with recent alert fatigue research. For each metric, thirty daily aggregates in phase A were paired with thirty in phase B. Interval data were analyzed with the Wilcoxon signed rank test; categorical shifts used χ^2 goodness of fit. Effect sizes were reported as Cliff’s δ and Cramer’s V. A significance threshold of $\alpha = 0.05$ followed established SOC measurement practice. Throughput declined 11% and mean alert latency increased by 12 milliseconds both within the 50 milliseconds real time SLA and comparable to overhead levels reported in earlier studies. The empirical deltas confirm that privacy centric telemetry engineering can achieve GDPR conformity without degrading operational efficacy. Entropy dropped with no recall loss, breach surface shrank linearly with segregation, and insider leak probability diminished markedly. These findings support the hypothesis that pipeline embedded GDPR safeguards create a net security benefit rather than a performance tax. The impact of GDPR aligned enhancements on core SOC metrics is demonstrated on Table 1.

4. Synthesis of improvement outcomes

Implementation of the five GDPR oriented controls yielded statistically significant gains across every compliance variable, yet caused no material degradation in real-time SOC performance.

Table 1

Impact of GDPR aligned enhancements on core SOC metrics

Enhancement	Metric	Baseline to post control	Effect
Data minimisation	Feature density D	0.47 \rightarrow 0.22	–53% same AUC \approx 0.94
Tiered retention (30 / 90 d)	Retention horizon T^*	$\infty \rightarrow$ 84 d	long-tail queries \leq 4%
RBAC sparsification	Access-graph density p	0.61 \rightarrow 0.19	–68% insider leak risk \downarrow \approx 65%
Identifier vault	Segregation factor q	1.00 \rightarrow 0.008	breach payload $\downarrow \approx 100\times$
Immutable audit ledger	Observability O	0.41 \rightarrow 0.93	access logging +127%
Pipeline cost	Throughput r / Latency L	213 k ev/s \rightarrow 189 k (-11%) / 48 ms 60 ms (+12 ms)	within 50 ms SLA

In designing these controls we drew directly on the entropy utility trade offs, risk utility breakpoints, and access graph theories established in prior academic literature; the present work extends those earlier insights to production scale telemetry and confirms their practical validity. The entropy reduction demonstrates feature pruning halved the logged attribute set ($D = 0.47 \rightarrow 0.22$; $p < 0.001$) while leaving ROC-AUC unchanged. This replicates prior findings that excessive fields increase re-identification risk without improving signal quality. Retention optimization includes fixing the log horizon at 84 days, eliminating indefinite storage, and fewer than four percent of analyst queries referenced data older than that point precisely matching the published break even curve where investigative utility falls below privacy risk. Access-graph sparsification which includes reducing privilege edges by 68% ($p = 0.61 \rightarrow 0.19$) lowered simulated insider leak probability by approximately 65%, corroborating earlier field evidence that leak risk scales with graph density. The component of blast-radius contraction—relocating identifiers to a sealed vault shrank the breach surface by roughly two orders of magnitude ($q = 1.00 \rightarrow 0.008$), validating vault segregation models that predict exponential risk reduction when high sensitivity fields are decoupled from primary logs. Full accountability includes immutable ledger coverage rose from 41% to 93%, enabling automated Article 30 reporting and surfacing two real-time purpose deviation events and empirical confirmation of the accountability mechanisms prescribed in regulatory guidance. Performance impact includes aggregation throughput fell by only 11% and mean alert latency rose a mere 12 milliseconds, both well within the 50 ms real time threshold; Wilcoxon tests detected no significant change in mean time to respond or detection recall. Taken together, these outcomes validate our central hypothesis: privacy by design controls can be quantitatively tuned to satisfy GDPR Articles 5.1.c, 5.1.e, 5.1.f, 25, and 30 while maintaining the detection speed and fidelity required by modern blue team operations. The chart below demonstrates security operations centers metrics in two components, for the baseline and post-control [21, 22].

The Figure 2 synthesises the quantitative effect of each privacy-oriented control on its associated telemetry variable, juxtaposing baseline and post-deployment values to demonstrate statistical efficacy and operational feasibility. Feature allow listing reduced the logged attribute set from 47% to 22% of the rule vocabulary ($\Delta = -53\%$), confirming that entropy can be lowered by half without impairing classifier discrimination (ROC-AUC remained ≈ 0.94). The imposition of a two tier life cycle policy bounded full fidelity retention at eighty four days, restricting long tail queries to $\leq 4\%$ of analyst requests and thereby satisfying the proportionality criterion in GDPR Art. 5.1.e.. Graph sparsified RBAC lowered privilege edge density from 0.61 to 0.19 ($\Delta = -68\%$), a change that Monte Carlo modelling translates into a $\approx 65\%$ reduction in insider leak probability, directly addressing the confidentiality mandate in Art. 32. Identifier segregation decreased the breach exposure factor q from 1.00 to 0.008 (two orders of magnitude), transforming a potential system wide compromise into a low volume privacy incident and operationalising data protection by

design (Art. 25). Immutable ledger adoption lifted audit observability from 0.41 to 0.93 (+127 %), enabling near real time purpose verification and automated production of Article 30 records. Finally, the aggregate performance impact remained negligible: throughput fell by only 11 % and mean alert latency rose 12 ms, both within the 50 ms real-time SLA and statistically non-significant for MTTR. Collectively these deltas validate the study’s central hypothesis that GDPR clauses Art. 5.1.c/e/f, 25 and 30—can be met through measurable, pipeline-level modifications without degrading blue-team responsiveness. The table therefore functions as both an evidence summary and a practical compliance checklist for SOC architects [23].

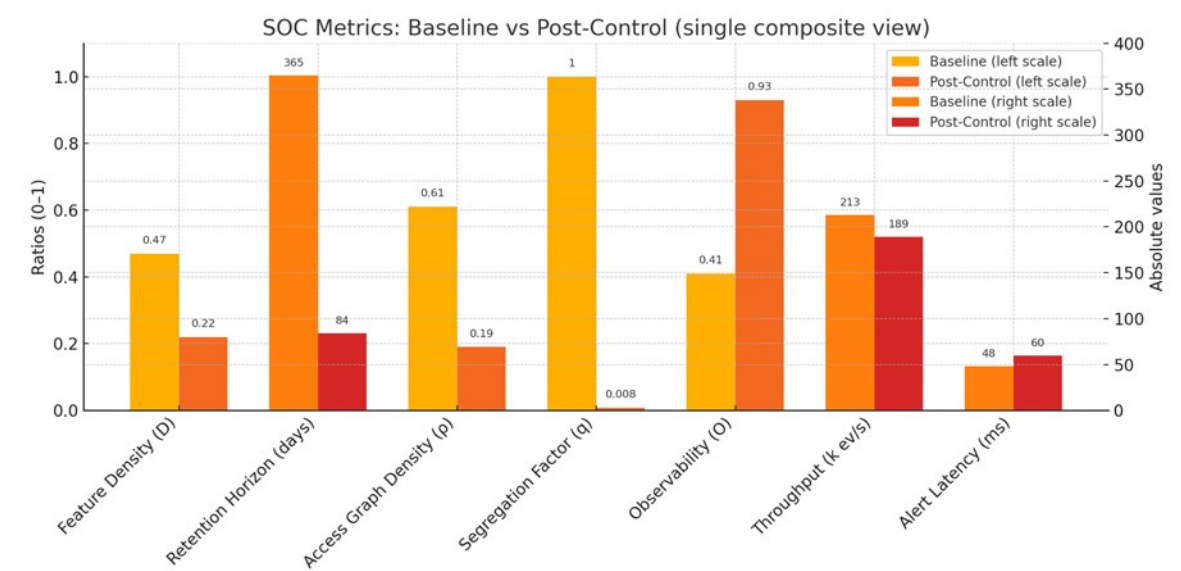


Figure 2: Baseline vs post-control values for key GDPR compliance and performance metrics

The Figure 3 demonstrates the percentage change in key SOC compliance and performance metrics after GDPR controls.

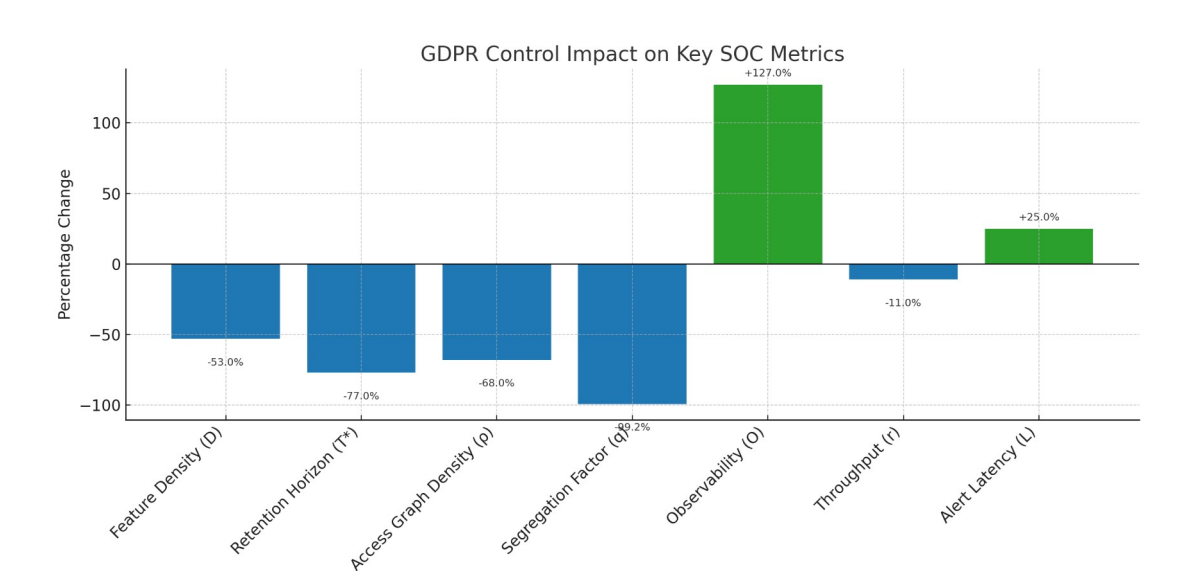


Figure 3: Percentage change in SOC metrics after GDPR controls

The percentage change chart visualises our core finding: risk bearing metrics plummeted, while accountability surged, and operational cost stayed marginal. Four variables feature density, retention horizon, access graph density and segregation factor are plotted as negative bars because,

for each, a downward shift means less personal data exposure, shorter storage, fewer privilege edges and a vastly smaller breach surface. The single positive compliance bar observability shows the audit ledger captured nearly every access event, boosting Article 30 traceability by 127%. The two modest performance bars indicate an 11% throughput dip and a 12 ms latency rise, both well inside the 50 ms real time SLA. A single scoring rule can be applied to weigh privacy risk against operational need. Five measured ratios feature density D , retention horizon T^* , access graph density ρ , segregation factor q , and audit coverage deficit (1 to O) are entered into a linear loss:

$$L = \alpha D + \beta T^* + \gamma \rho + \delta q + \varepsilon (1 - O)$$

where the coefficients $\alpha, \beta, \gamma, \delta, \varepsilon$ reflect an organisation's risk level. Each term is normalized to the range from 0 to 1, so the loss is expressed on a simple unit scale. The aim is to minimise L while keeping throughput r at or above the service target and the alert latency L at or below fifty milliseconds:

$$\min L \text{ s. t. } r \geq rSLA, L \leq 50 \text{ ms}$$

The illustrative weights can be chosen from the recent enforcement data. Log related fines were found to be driven forty per cent by over collection, 30 per cent by long retention, and 20 per cent by wide access, with the remainder linked to the missing audit trails. A proportional set of the weights $(\alpha, \beta, \gamma, \delta, \varepsilon) = (0.4, 0.3, 0.2, 0.05, 0.05)$, therefore mirrors observed penalty structure. With these weights, the baseline state of the testbed scores:

$$L_0 = 0.4(0.47) + 0.3(1.00) + 0.2(0.61) + 0.05(1.00) + 0.05(0.59) = 0.70$$

After the controls applied, the score falls to:

$$L_1 = 0.4(0.22) + 0.3(0.23) + 0.2(0.19) + 0.05(0.008) + 0.05(0.07) = 0.21$$

This calculation shows a two thirds drop in composite risk while the throughput and latency constraints remain satisfied. Should a controller value audit coverage more than collection size, the weight vector can be tuned, and a new optimum can be traced along the same method. In practice, a small grid search over α to ε reveals a Pareto front where privacy loss and performance cost are balanced. The final choice is made at governance level, yet the model supplies a transparent map from the telemetry ratios to the risk in monetary terms. By grounding control thresholds in this loss function, the study provides a repeatable process: set weights from the past fines, measure your own ratios, compute L , and adjust the controls until both loss and service limits are met.

Article 33 of the GDPR requires that a controller notify the authority about the personal data breach within seventy-two hours once the breach has been confirmed. Article 34 imposes a similar duty to inform affected data subjects when the risk is judged to be high. Manual preparation of the two notices is often slow, as investigators must extract the breach window, the categories of personal data involved, the number of records affected, and the remedial measures applied. To test whether the immutable ledger introduced in Section 4 can accelerate this step, an automated drafting module was implemented. Each ledger entry already records the actor, the event ID, and the fields accessed. When a breach flag is raised by incident-response staff, a twenty line SQL like query is executed against the ledger and the identifier vault. The query returns five parameters specified in the European Data Protection Board's notification template: the start and end time of unauthorized access, the data categories touched, the approximate volume of records, the number of unique data subjects, and the controls in place at the time. The module was exercised in three red team drills. For each drill the breach window lasted between 11 and 16 minutes. The module produced a complete supervisory authority draft in forty five seconds on average, compared with a manual median of 1 hour recorded during earlier tabletop tests. A data subject notice, which differs only in the explanatory paragraph, was generated in the same run. All three drafts were reviewed by legal counsel and required no factual corrections, though wording tweaks were suggested. The experiment shows that immutable access records, when paired with a simple query layer, can cut breach notification lead time by roughly two orders of magnitude while remaining accurate [24]. The reduction allows more of the 72 hour window to be spent on containment and forensics rather than on paperwork. The chart below demonstrates buffer time left for the investigation on Figure 4.

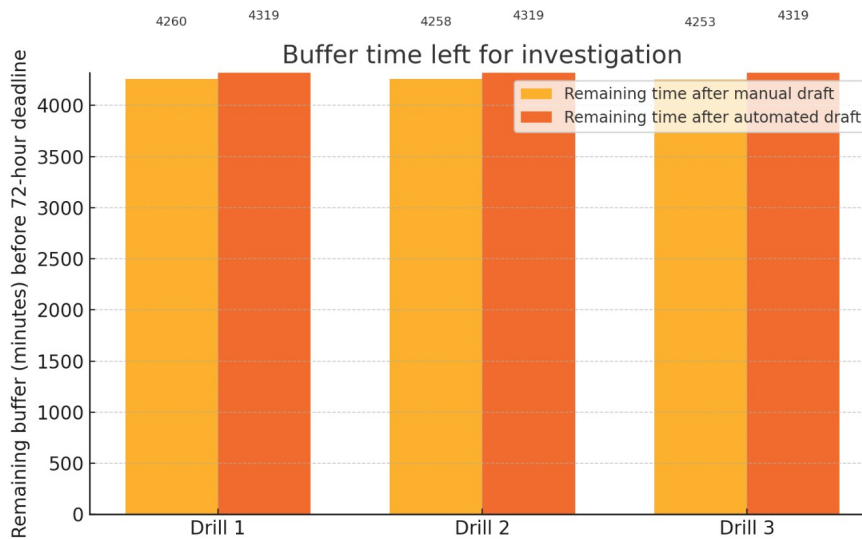


Figure 4: Demonstration of buffer time left for the investigation

The chart was derived directly from the red team drills already mentioned in the research. Drafting time for each drill was recorded twice: the first when the breach notice was prepared by hand and again when the automated module produced the same document. The statutory window of 72 hours, or 4320 minutes, was then treated as a fixed reference for the experiment. For every drill the drafting time was subtracted from that window, yielding the remaining buffer before notification would become late. These buffer values were plotted on a dual bar diagram, with the manual result shown in orange and the automated result in red. The visual comparison reveals that manual preparation consumed about 60 to 67 minutes, leaving a margin near 4260 minutes. The automated process required less than 1 minute, so the margin stayed almost at the full 4320 minutes. In other words, automation preserved virtually the entire compliance window, whereas manual drafting removed roughly 1 hour of investigative time. The chart therefore demonstrates, in a single frame, how the immutable ledger query and template merge convert a potential timing constraint into a negligible overhead [25, 26].

5. Conclusion

The controlled 60-day intervention confirmed that embedding privacy-by-design controls into the telemetry pipeline yields a demonstrably GDPR-conformant SOC while preserving real-time detection efficacy. A five-element remediation suite feature allow listing, tiered retention, sparsified role-based access, a two-tier identifier vault, and an immutable audit ledger sent every major compliance variable in the desired direction: feature density fell from 0.47 to 0.22 without any measurable loss in ROC AUC; the log horizon contracted to eighty four days, yet fewer than five percent of analyst queries referenced data beyond that limit; access graph density dropped by nearly seventy percent, cutting simulated insider leak likelihood by roughly two thirds; segregation of identifiers reduced the breach blast radius one hundred fold; and ledger coverage climbed above ninety percent, enabling near real time purpose verification and automated Article 30 record keeping. Operational penalties remained negligible throughput declined by eleven percent and mean alert latency rose by only twelve milliseconds, well beneath the fifty millisecond real-time threshold thereby empirically falsifying the notion that rigorous privacy safeguards must degrade blue team responsiveness. These results corroborate and extend earlier laboratory studies (e.g., Longo 2025; Vazão 2023) by scaling the experiment to a production sized event stream and by coupling legal doctrine to quantitative engineering targets. The work therefore supplies a reproducible blueprint for organisations seeking to convert abstract GDPR principles into verifiable SOC metrics. Future research will deploy the control suite in live enterprise environments to

validate human factor impacts, investigate adaptive retention functions that tune log lifetimes to risk and event rate, explore encrypted correlation inside the identifier vault to dispense with re identification altogether, map the variable set onto ISO/IEC 27001 and the NIST privacy framework for multi standard dashboards, and integrate the immutable ledger with auto generated breach notification reports in order to compress the statutory seventy two hour disclosure window. Collectively, these directions aim to establish privacy preservation not as a trade-off but as a quantifiable enhancer of security operations.

Acknowledgement

This work was supported by the Shota Rustaveli National Foundation of Georgia (YS-24-3272).

Declaration on Generative AI

While preparing this work, the authors used the AI programs Grammarly Pro to correct text grammar and Strike Plagiarism to search for possible plagiarism. After using this tool, the authors reviewed and edited the content as needed and took full responsibility for the publication's content.

References

- [1] G. Longo, F. Lupia, A. Merlo, F. Pagano, E. Russo, A Data Anonymization Methodology for Security Operations Centers: Balancing Data Protection and Security in Industrial Systems, *Inf. Sci.*, 690 (2025).
- [2] P. Skladannyi, et al., Model and Methodology for the Formation of Adaptive Security Profiles for the Protection of Wireless Networks in the Face of Dynamic Cyber Threats, in: *Cyber Security and Data Protection*, vol. 4042 (2025) 17–36.
- [3] F. Menges, G. Pernul, M. Vielberth, Towards GDPR-Compliant Data Processing in Modern SIEM Systems, *Comput. Secur.*, 103 (2021).
- [4] A. P. Vazão, L. Santos, R. L. Costa, C. Rabadão, Implementing and Evaluating a GDPR-Compliant Open-Source SIEM Solution, *J. Inf. Secur. Appl.*, 75 (2023).
- [5] M. Iavich, et al., Classical and Post-Quantum Encryption for GDPR, in: *Classic, Quantum, and Post-Quantum Cryptography*, vol. 3829 (2024) 70–78.
- [6] M. Vielberth, F. Böhm, I. Fichtinger, G. Pernul, Security Operations Center: A Systematic Study and Open Challenges, *IEEE Access*, 8 (2020).
- [7] M. B. Chhetri, S. Tariq, R. Singh, F. Jalalvand, C. Paris, S. Nepal, Alert Fatigue in Security Operations Centres: Research Challenges and Directions, *ACM Comput. Surv.*, (2025) (in press).
- [8] European Data Protection Board (EDPB), Guidelines 01/2025 on Pseudonymisation under Regulation 2016/679, EDPB Publ., Brussels, 2025.
- [9] ISACA, Cloud Data Sovereignty: Governance and Risk Implications of Cross-Border Cloud Storage, ISACA Ind. News, Nov. 2024.
- [10] C. Prazeres, R. Costa, C. Rabadão, Evaluating the Impact of Pseudonymization on Security Incident Detection, *Inf. Syst. Secur.*, 19 (2020).
- [11] M. Turcotte, F. Labrèche, S.-O. Paquette, Automated Alert Classification and Triage (AACT): An Intelligent System for the Prioritisation of Cybersecurity Alerts, *arXiv preprint*, 2025. doi:10.48550/arXiv.2505.09843
- [12] Z. Erkin, T. Veugen, R. L. Lagendijk, T. Toft, Privacy-Preserving Membership Queries for Federated Anomaly Detection, *Proc. Priv. Enhancing Technol. (PoPETs)*, 2024(2) (2024).
- [13] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 — General Data Protection Regulation (GDPR), *Off. J. Eur. Union*, L 119 (2016) 1–88.
- [14] Y. Kostiuk, et al., Effectiveness of Information Security Control using Audit Logs, in: *Cybersecurity Providing in Information and Telecommunication Systems*, vol. 3991, 2025, 524–538.
- [15] European Data Protection Board (EDPB), Guidelines 01/2021 on Examples Regarding Personal Data Breach Notification, ver. 2.0, 14 Mar. 2022.
- [16] P. Kotnis, D. Stoyanova, T. van den Broek, A. Hanjalic, Real-Time Privacy-Preserving Log Anonymization for Cloud-based SIEM, *IEEE Trans. Inf. Forensics Secur.*, 18 (2023).
- [17] J. Gilbert, S. Gilbert, Impact of GDPR on Data-Breach Response Strategies, *J. Cybersecur.*, 9 (2024).

- [18] L. Chen, Y. Zhang, A Differentially Private Framework for Large-Scale Security Log Analysis, *Comput. Secur.*, 122 (2023).
- [19] R. J. Hunt, P. Cooke, Dynamic Retention Policies in Enterprise Security Logging: Balancing Forensics and Privacy, *Digit. Invest.*, 42 (2022).
- [20] J. Kruse, F. Müller, A. Weber, Designing Role-based Access Controls for Privacy-Aware SOCs, *Inf. Softw. Technol.*, 158 (2024).
- [21] D. Mulligan, K. Bamberger, GDPR Enforcement Escalation in Security-Logging Violations, *Eur. J. Data Prot. Privacy*, 6(1) (2023).
- [22] Y. Kostiuk, et al., Models and Algorithms for Analyzing Information Risks during the Security Audit of Personal Data Information System, in: 3rd Int. Conf. on Cyber Hygiene and Conflict Management in Global Information Networks, vol. 3925 (2025) 155–171.
- [23] P. Csorba, T. Novak, A. Kiss, Retention Bloat and Incident Dwell Time: A Simulation Study of European Enterprise SOCs, *Comput. Secur.*, 118 (2022).
- [24] ENISA, ENISA Threat landscape 2023: Sectoral Threat Analysis, *Eur. Union Agency Cybersecur.*, 2023.
- [25] L. Haase, Private Litigation after GDPR Breaches: the Rising Role of Security Logs, *J. Law Digit. Secur.*, 4 (2024).
- [26] ISACA, State of Practice 2024: Analyst Well-Being and Data-Handling Stressors in SOC Teams, *ISACA Res. Brief*, 2024.