Privacy-preserving Cyber Threat Intelligence Sharing in Healthcare: Automated Anonymisation Under EU Regulations.

Ilkka Tikanmäki[1,2][1] Lillebi Seistola[1], Harri Silvola[1] and Pihla Parviainen[1]
[1]Laurea University of Applied Sciences, Espoo, Finland
[2]National Defence University, Helsinki, Finland

Ilkka.tikanmaki@laurea.fi

ABSTRACT: Healthcare organisations face increasing cyber threats, making cyber threat intelligence (CTI) sharing an essential component of resilience. CTI sharing from hospitals often contains highly sensitive information such as patient identifiers, internal domains or IP addresses that can lead to re-identification if shared without proper protection. When this information is shared without protection, it can allow for re-identification or system mapping, exposing healthcare organisations to additional cyber and privacy risks. Under the emerging EU Cyber Resilience Act, ensuring secure and privacy-preserving cyber threat intelligence exchange becomes critical, requiring tools and processes that safeguard personal and organisational data while maintaining the analytical utility needed for threat detection. This study provides a theoretical framework for evaluating automated anonymisation in the sharing of healthcare cyber threat intelligence (CTI). It explains how to utilise tools such as ARX and DAT to execute privacy-preserving CTI exchange that is compliant with the EU Resilience Act (CRA) and the NIS2 Directive. The framework establishes evaluation criteria that balance anonymisation risk and analytical utility using synthetic CTI data (such as MISP feeds). The evaluation is intended to evaluate the efficiency of anonymisation by using re-identification risk metrics and retained CTI utility (such as Indicator of Compromise, IOC correlation). The model proposed follows European privacy and resilience rules and establishes the foundation for practical and compliant sharing of CTI in healthcare ecosystems.

Keywords: Cyber threat intelligence, Healthcare cybersecurity, Data anonymisation, EU Cyber resilience act, Privacy and data utility, NIS2

1. Introduction

Healthcare organisations are increasingly targeted by sophisticated cyberattacks, making effective cyber threat intelligence (CTI) sharing a critical element of organisational resilience. CTI shared by hospitals often includes highly sensitive data that can enable re-identification or system mapping if distributed without adequate protection. This creates a tension between the need for timely threat information exchange and the obligation to safeguard both personal and organisational data. Legal uncertainties, the caution of organisational culture and the fear of reputational damage are often perceived as obstacles to information sharing in healthcare organisations. These factors slow down cooperation between cyber actors and highlight the importance of anonymisation as an enabler of trust in CTI exchanges (Brilingaitė et al., 2022; ENISA, 2017).

The EU Cyber Resilience Act (CRA) and the obligations under NIS2 mandate create strong expectations and obligations for privacy-preserving mechanisms for sharing cyber threat intelligence (European Parliament and of the Council, 2024). Methods such as differential privacy can lower re-identification risks while still preserving key data patterns when configured correctly (Dankar & El Emam, 2013). Anonymisation should not remove indicator-level technical information such as hashes or IOC chain structures, as these elements are critical in the correlation and identification of threats (Conti et al., 2018). CTI sharing depends on good governance, clear process, and continuous monitoring of privacy and security risks. Anonymisation must be combined with broader cybersecurity practices to remain effective (Bonomi et al., 2023). Without continuous updating work, rule-based controls and anonymisation methods quickly become obsolete, which can reduce the quality and safety level of anonymisation (Modi et al., 2023).

Automated anonymisation techniques present a promising approach to balancing utility and privacy. This study investigates whether such techniques can enable hospitals to share analytically useful cyber threat intelligence while complying with European privacy and cybersecurity regulations. The research question guiding this work is:

- How can automatic anonymisation preserve the usefulness of cyber threat intelligence data while adequately reducing re-identification risk in healthcare organisations?

[1]https://orcid.org/0000-0001-8950-5221

The rest of the paper is organised into five chapters. After this Introduction chapter, Chapter 2 presents a literature review, and Chapter 3 offers a theoretical background. Chapter 4 outlines the findings, and Chapter 5 provides a discussion and the conclusions of the study.

2. Cyberthreats in Healthcare

The cybersecurity threat landscape in healthcare is escalating rapidly with increasingly sophisticated attacks and growing regulatory pressures. The healthcare sector now faces "rapidly sophisticated threats", while ransomware remains one of the most disruptive attack types (Rubrik, 2026). ENISA's analysis shows that healthcare organisations are particularly vulnerable to cyberattacks, and most reported cases in the sector are currently related to ransomware attacks, which highlights the need for strong threat intelligence sharing (ENISA, 2023b). Healthcare providers represent a significant share of threat targets in critical infrastructure, and CTI cooperation between organisations remains lacking due to data protection risks and legal uncertainty (ENISA, 2023b, 2023a). The widespread use of MedTech and IoT devices in healthcare significantly increases the attack surface. Many hospital devices run on outdated operating systems and without strong segmentation, increasing the vulnerability of the systems and highlighting the need for anonymised technical threat intelligence sharing (ECSO, 2025).

3. Method

The purpose of this study is to develop an assessment framework for sharing privacy-preserving cyber threat intelligence (CTI) in healthcare using both structured literature-based and conceptual modelling approaches. The objective is to achieve methodological transparency and theoretical rigour, not empirical generalisation.

A systematic literature search was conducted between January 2025 and February 2026 using Scopus, Web of Science, IEEE Xplore, ACM Digital Library, Google Scholar, and official ENISA and EU institutional sources. The searches were conducted using Boolean logic. Search terms that would yield relevant results include ("cyber threat intelligence" OR CTI) AND ("anonymisation or pseudonymisation" OR "privacy-preserving") AND (healthcare OR hospital) AND (EU OR GDPR OR NIS2 OR "Cyber Resilience Act"). Predefined inclusion and exclusion criteria (Table 1) were used to screen titles and abstracts, and search terms were refined iteratively.

Table 1*:* Inclusion and exclusion criteria

Inclusion criteria	Exclusion criteria
Published between 2013 and 2025	Published outside the defined time window
Written in English	Non-English publications
Peer-reviewed articles or authoritative reports (e.g. ENISA, EU bodies)	Opinion pieces or non-authoritative blog content
Addresses CTI sharing, anonymisation/pseudonymisation, or re-identification risk	Focuses on anonymisation without cybersecurity or CTI relevance
Relevant to healthcare or critical infrastructure contexts	Healthcare privacy studies without a cybersecurity dimension

Peer-reviewed studies and authoritative policy documents that directly address anonymisation, re-identification risk, or regulatory compliance were chosen as the top choice for cybersecurity selection. The Anonymisation and Pseudonymisation Guidelines from ENISA 2021 were implemented across three analytical dimensions (ENISA, 2021). The first step was to categorise CTI attributes into direct identifiers, quasi-identifiers, and non-identifying technical indicators to guide anonymisation strategies. The second step was to transform ENISA's risk-based approach to commonly used assessment metrics, which included anonymity and conceptual re-identification probability under reasonable attacker assumptions. To ensure reproducibility, the anonymisation steps, parameter choices, and assessment criteria were clearly documented. The framework is presented using synthetic CTI scenarios that are inspired by MISP-like inputs, emphasising two dimensions: minimising privacy risk and maintaining analytical efficiency.

4. Results

The results are focused on identifying privacy risks, analysing anonymisation effects, and developing an evaluation framework that does not rely on empirical measurement. There are three categories for Healthcare

CTI's recurring identification structure that have clear privacy implications. The risk of re-identification is high when direct identifiers (like usernames or explicit personal references) are used, but they are not common in mature CTI feeds. If they are present, they need to be blocked strictly or anonymised in a way that cannot be reversed.

The primary privacy risk is associated with quasi-identifiers, such as internal IP ranges, hostnames, timestamps, and organisational metadata. Infrastructure mapping and indirect attribution can be achieved through the combination of these elements, although they are not harmful on their own. According to the conceptual assessment, quasi-identifiers are the main cause of the residual identification risk in anonymised CTI, not explicit personal data. The use of non-identifying technical indicators, including file hashes, malware families, and common tactics, techniques, and procedures, poses minimal privacy risks while providing high analytical value. These indicators are crucial for analysing threats and should not be greatly impacted by anonymisation.

The systematic trade-off between privacy and utility is revealed when anonymisation techniques are applied to these identifier constructs. The use of extensive suppression-based anonymisation is advantageous for privacy protection, but it also significantly reduces CTI's utility by disrupting indicator chains and correlation potential. The risk of re-identification can be reduced by using generalisation (including IP address aggregation and timestamp coarsening) and controlled pseudonymization, while still maintaining structural context. The results indicate that there is a range of anonymisation outcomes: risk-aware configurations that reasonably retain significantly more analytical value than maximum anonymisation, while still meeting reasonable privacy expectations.

The ENISA anonymisation guidelines being operationally translated into a CTI-specific evaluation sequence is the primary outcome. When applied to CTI datasets, the ENISA criteria function as operational controls by identifying identifiers, selecting a relative anonymisation strategy, and assessing conceptual residual risk. It utilises metrics such as k-anonymity and the likelihood of re-identification with reasonable attacker assumptions.

From the synthesis of these results, the study derives a two-dimensional evaluation framework that evaluates anonymised CTI from the perspective of privacy protection and analytical utility. Privacy is measured by how much identifiers are exposed and residual risk, while utility is measured by how well indicator relationships are preserved and correlational capacity. By using the framework, transparency can be achieved in justifying the trade-offs of anonymisation, and a structured basis can be established for future empirical validation. Table 2 summarises the resulting evaluation framework.

Table 2*:* Dimensions, privacy and utility focus

Dimension	Privacy focus	Utility focus
Identifier type	Direct / Quasi / Technical	Indicator semantics
Technique	Suppression / Generalisation	Pseudonymisation
Metric	k-anonymity	IOC correlation
Outcome	Risk reduction	Analytical continuity

This study's main outcome is a structured means to evaluate anonymisation's outcomes across privacy protection and CTI utility dimensions, which is provided by the resulting framework.

5. Discussion and Conclusions

The results show that automated anonymisation can support safer CTI sharing in healthcare. Rule-based and statistical methods were able to remove sensitive identifiers while keeping most of the information that analysts need. Findings confirm that there is no perfect balance: when anonymisation settings are too strict, some technical details become less useful for detecting or comparing threats.

The study indicates that anonymisation tools like ARX and DAT can help healthcare organisations meet the privacy expectations created by the EU CRA and NIS2. These regulations require that any shared threat information be protected so that individuals and system details cannot be easily identified (European Parliament and of the Council, 2024). By reducing sensitive fields before the data leaves the organisation, anonymisation supports compliance and lowers the risk of privacy breaches.

The results suggest that hospitals could integrate anonymisation into their existing cybersecurity workflows with limited effort. Anonymisation steps could be applied automatically before CTI is shared with national or EU-level

partners. This could improve both the speed and the safety of information exchange. Automatic anonymisation seems to be a realistic way to increase secure CTI sharing in healthcare, but it must be carefully configured and regularly reviewed. The EU CRA and the NIS2 Directive require the protection of sensitive personal and operational data in information security cooperation between organisations. Automatic anonymisation offers a concrete way to strengthen the overall resilience of healthcare by enabling safer and more confidential participation in national and EU-level CTI networks.

Research shows that anonymisation alone is not enough. Its effectiveness depends heavily on the organisation's governance models, maintaining up-to-date classification rules and threat models, and how well anonymisation is integrated into existing cybersecurity processes. Settings that are too strict can reduce the usefulness of the CTI data being shared, while settings that are too lenient can leave identification risks behind. This study refers to anonymisation as an umbrella term, recognising that a lot of CTI use cases rely on strong pseudonymization instead of irreversible anonymisation. Anonymisation profiles optimised for the healthcare sector are needed to balance data protection and analytical benefits. Such research would support more mature, reliable and privacy-preserving cyber defence cooperation in Europe's healthcare systems. Future work should examine how anonymised CTI performs in larger, real-world sharing environments and how hospitals can standardise their anonymisation practices to ensure consistent quality.

Acknowledgements

This study has received funding from the European Union project DYNAMO, the European Union's Horizon 2020 research and innovation programme under the grant agreement no. 101069601 and The Cybersecurity in Everyday Work in the Social and Healthcare Sector (KyberSoTe) project. The views expressed are those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

Ethics declaration: No human participants or personally identifiable information were involved. All data sources were publicly available.

AI declaration: Keenious was used for literature exploration, and ChatGPT 5.1 for drafting and refinement. Human authors verified all content.

References

Bonomi, L., Gousheh, S., & Fan, L. (2023). Enabling Health Data Sharing with Fine-Grained Privacy |. *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management*, 131–141. https://doi.org/10.1145/3583780.3614864

Brilingaitė, A., Bukauskas, L., Juozapavičius, A., & Kutka, E. (2022). Overcoming information-sharing challenges in cyber defence exercises. *Journal of Cybersecurity*, *8*(1), tyac001. https://doi.org/10.1093/cybsec/tyac001

Conti, M., Dehghantanha, A., & Dargahi, T. (2018). Cyber Threat Intelligence: Challenges and Opportunities. *ArXiv*, *70*, 1–6. https://doi.org/10.1007/978-3-319-73951-9_1

Dankar, F. K., & El Emam, K. (2013). Practicing Differential Privacy in Health Care: A Review. *Transactions on Data Privacy*, *6*(1), 35–67.

ECSO. (2025). *Cybersecurity in Healthcare: Insights from Security Professionals*. European Cyber Security Organization. https://ecs-org.eu/ecso-uploads/2025/01/Cybersecurity-in-Healthcare_Jan-2025_v1.2.pdf

ENISA. (2017). *Information Sharing and Analysis Center (ISACs)—Cooperative models* (p. 51) [Cooperative models]. The European Union Agency for Network and Information Security (ENISA).

ENISA. (2021). *Data Pseudonymisation: Advanced Techniques and Use Cases* (p. 55) [Technical analysis of cybersecurity measures in data protection and privacy]. European Union Agency for Cybersecurity (ENISA). https://www.enisa.europa.eu/sites/default/files/publications/ENISA%20Report%20-%20Data%20Pseudonymisation%20-%20Advanced%20Techniques%20and%20Use%20Cases.pdf

ENISA. (2023a). *ENISA Threat Landscape (ETL)* (p. 161) [July 2022 to June 2023]. The European Union Agency for Cybersecurity, ENISA. https://doi.org/10.2824/782573

ENISA. (2023b). *ENISA Threat Landscape: Health Sector* (Threat Landscape Report TP-04-23-546-EN-N; p. 36). European Union Agency for Network and Information Security (ENISA). https://www.enisa.europa.eu/publications/health-threat-landscape

European Parliament and of the Council. (2024, November 8). *Directive on measures for a high common level of cybersecurity across the Union (NIS2 Directive) | Shaping Europe's digital future*. European Parliament. https://digital-strategy.ec.europa.eu/en/policies/nis2-directive

Modi, A., Kuzminykh, I., & Ghita, B. (2023). *Data Driven Approaches to Cybersecurity Governance for Board Decision-Making—A Systematic Review* (arXiv:2311.17578). arXiv. https://doi.org/10.48550/arXiv.2311.17578

Rubrik. (2026). *Healthcare Cybersecurity Challenges & Threats—2026 Update*. Rubrik Insights. https://www.rubrik.com/insights/healthcare-cybersecurity-challenges-threats-2025