Invisible Data Subjects, Visible Personal Data: A Systematic Review of Privacy Implications for Organs-on-a-Chip in the EU

Dmitry Prokhorenkov*, Dimitrios Tsolovos

Stelar Security Technology Law Research, Versmannstr. 4, 20457 Hamburg, Germany, dp@stelar.de

Abstract. Organs-on-a-Chip (OoC) have emerged as a promising alternative to animal testing and offer an innovative approach to personalised drug development. However, the rapid advancement and implementation of OoC technologies have outpaced existing EU regulations, leaving various privacy risks related to potential re-identification insufficiently addressed and raising concerns regarding the responsibilities of data controllers and processors. The systematic review examines both the dedicated privacy risks and the technologies, such as PET DP, and their implications for the OoC domain. The study considers privacy risks and mitigation measures, with a specific focus on the intersection of GDPR, AIA, and the Digital Omnibus proposal. We systematically analyse the selected corpus study across ten years (n=374), covering a wide range of academic disciplines. By applying best-fit framework synthesis to develop ten high-level risk categories. This study demonstrates that the approach is functional for identifying and evaluating complex privacy risks in the OoC domain across diverse EU regulations. However, the development of risk categories and their alignment with regulatory accuracy appear to be complex processes. The findings indicate a lack of OoC-specific, systematically examined privacy risks that are explicitly allocated and studied, with existing studies focusing primarily on governance and standardisation requirements. We demonstrate the implications for interpreting the relationship between technical and legal requirements, for adopting evolving legal definitions, and for identifying the need to utilise PETs within the OoC domain during the transition from research and development to industry. While we observe some consideration of privacy risks for data subjects within the dedicated OoC research environment, these risks are not addressed in the industry domain. The study also identifies current challenges and proposes directions for future research to strengthen technical and legal measures, enhance data subjects' privacy, and build resilience to re-identification risks.

Keywords: Organ-on-chips · Privacy · GDPR · AIA · Differential Privacy · microphysicalogical systems

1 Introduction

Organs-on-a-Chip (OoC) have emerged as promising technologies that offer an alternative to animal testing and provide innovative approaches to personalised treatment [75] and drug development [38]. OoC could be described as microphysicalogical systems (MPSs): physiological, bioengineered systems based on microfluidic chips that contain human cells [49], [38] and emulate various specific organ-level functions and processes [35]. OoC applications are not limited to such human organs as lung-on-a-chip [31], gut-on-a-chip, multi-organ chip systems [68], and other essential body processes and critical organs, such as heart [43], liver, retina, brain [57], and bones. The rapid development and

broad applicability of OoC raise significant privacy and policy concerns, as well as ethical considerations [79]. The OoC domain is inherently multidisciplinary and extends beyond materials science, integrating techniques and technologies from biology, engineering, policy, physiology, and computer science. Beyond technical challenges, the OoC domain faces significant issues with standardisation, taxonomy, and data reporting [59]. This multidisciplinary nature broadens the applicability of research findings across diverse domains and fosters collaboration that improves scientific, technical, and clinical outcomes, particularly regarding privacy. Such collaboration represents a rapidly demanded trend within the European Union (EU) [65].

Within current EU regulations, the processing of health and genetic data, classified as special category data (Art. 9, GDPR), requires strict compliance with core principles such as purpose limitation, data minimisation, and storage limitation. Responsible data controllers, processors and relevant third parties are required to safeguard data subjects' rights and privacy throughout the entire data processing lifecycle. To meet regulatory requirements and mitigate privacy risks such as data linkage and re-identification, privacy-enhancing technologies (PETs) are essential. PETs utilisation is particularly significant for processing data from human-derived samples and molecular readouts [14], as well as for enhancing privacy and enabling secure data transfers among partners and third parties [19], [54], [6].

However, regulatory frameworks and PET-based support for OoC implementation remain underdeveloped [82], [71]. In particular, privacy risks related to the potential re-identification of coded or pseudonymised donor data, as well as the responsibilities of data controllers and processors, require further investigation. These challenges significantly impact the OoC domain [79], especially in applications such as personalised drug development [73]. Although consent, data transfer, and accountability pose some of the most challenging privacy risks to forecast and control [3], these areas require further analysis. Building on these considerations, current approaches to biomedical data mainly address and encode privacy risk through governance requirements, such as controlled-access repositories [45], and treat PETs primarily as a complementary layer for specific data flows and use cases [14]. The increasing integration of AI-driven methods and the use of complex, high-dimensional, and sensitive datasets further exacerbate this gap.

Furthermore, despite extensive academic attention to privacy-preserving methods and regulatory compliance [14], [11], a comprehensive systematic literature review (SLR) and critical analysis examining the intersection of the OoC domain with current and emerging EU legislation remains lacking. Specifically, there is limited research that systematically maps the impact of the General Data Protection Regulation (GDPR) [25], the EU Artificial Intelligence Act (AIA) [26], and the Digital Omnibus proposal [20] on privacy risk management, with particular attention to the PETs such as differential privacy (DP) [18] and evolving interpretations of the concepts of personal data and the applicable thresholds for anonymisation.

The objectives of the current study are threefold: (1) to address the lack of systematic reviews on privacy risks at the intersection of OoC technologies and emerging EU regulations; (2) to study risks affecting data subjects and data controllers and to provide recommendations on PET-based mitigation strategies; and (3) to identify research priorities and practical implications for organisations involved in privacy and data protection.

To enhance understanding of the current state of the art, we conduct an SLR utilising academic evidence relevant to the objectives of this study. The review is conducted in accordance with recognised guidelines [41], emphasising privacy risks and mitigation strategies within the OoC domain, and discussing the interactions among the GDPR, AIA, and the Digital Omnibus proposal. Our SLR is guided by the following research questions (RQs):

- **RQ1**: What privacy risks emerge in the OoC domain within the EU regulations? *By answering RQ1, we aim to identify, describe, and classify privacy risks in the context of the OoC domain, and to provide an overview of identified risks.*
- **RQ2**: How and to what extent do PETs utilised in the OoC domain mitigate privacy risks? *By answering RQ2, we account for the applicability and necessity of PETs in mitigating the identified privacy risks in the OoC domain and examine the mitigation strategies reported for addressing these risks.*

The analysis examines the privacy risks associated with the OoC domain and discusses the corresponding mitigation strategies. It also debates risk management insights that address the identified privacy risks, considering EU regulations and interpretations of recommendations from data protection authorities. Our findings are intended to support practitioners and researchers by facilitating the broader adoption of OoC technologies and PETs while reducing privacy risks for data subjects. Furthermore, these insights support policy and standardisation efforts, enabling diverse stakeholders to meet goal-based EU regulatory requirements across multiple domains. The systematic nature of the review supports the reproducibility of the study and enhances the applicability of the analysis for practitioners, policy-makers, and technical and legal experts, while also facilitating further academic research.

This study delivers four principal contributions at the intersections of the *OoC domain, EU regulations and privacy risks*: (i) We investigate and uncover the complexity of privacy risks in the OoC domain and organise them into dedicated privacy risk categories. (ii) We demonstrate the applicability of the best-fit framework synthesis for conceptualising high-level privacy risk categories. (iii) We highlight the importance of the proposed analysis, alongside EU goal-based regulations (GDPR, AIA, Digital Omnibus proposal), to assess privacy risks for data subjects in the OoC domain. (iv) We provide evidence of the urgent need to consider privacy risks in the OoC domain and to develop policy requirements that address emerging risks from the early stages of development.

The remainder of this study is structured as follows. **Section 2** establishes the context, background, and related work, including key definitions and concepts. **Section 3** outlines the research methodology and details the SLR process. **Sections 4** and **5** present the results and discussion, respectively. **Section 6, 7** presents the limitations and conclusion.

2 Background

OoC is a multifaceted research domain, particularly focused on advanced tissue culture models. The development of OoC can be traced back to the 1960s and to related research on micro-electro-mechanical systems [13]. The promising OoC technology challenges existing experimental approaches and advances knowledge of human biology and diseases [50], including the discovery and testing of new drugs, with specific personalisation to the human subject. Accordingly, this section outlines the relevant intersections among privacy, OoC technologies, and EU regulation.

Privacy Risks and Mitigation Privacy challenges and implications associated with biomedical data have evolved significantly over the past several decades and remain complex [14]. These issues are shaped by multiple factors, including technological advancements and the expanding role of biomedical data, particularly in personalised drug development within the OoC domain. On the one hand, privacy concerns persist and continue to expand into new biomedical domains; on the other

hand, the EU aims to clarify and navigate the associated privacy risks, including the responsibilities of the respective stakeholders, by developing and implementing tailored, applicable regulations (GDPR, AIA, Digital Omnibus proposal). The expansion of large-scale biobanks, biomedical technologies, and associated data repositories, along with the increased sharing and distribution [62] of personally identifiable information (PII), further amplifies privacy-related challenges. Given the complexity and scope of genetic and related biomedical data, privacy violations may affect entire families [32], [17] or span multiple generations. Furthermore, as reported in [21], preserving the full utility of original data while mitigating privacy risks remains an unresolved challenge. Moreover, the structural complexity of genetic data presents significant challenges for assessing re-identification risk and implementing effective privacy safeguards. As explained in [74], the corresponding risk re-identification assessment to guide data processors, particularly for genetic data, is not yet available. Furthermore, to address the varying legislation across jurisdictions, data processors are not required to fully disclose the referenced risks [74]. However, as further argued in [10], the methods utilised must go beyond adopting technical and organisational methods applied to genetic data, which are more likely to fulfil the legal requirements. This partly underscores the importance of PETs in addressing privacy risks and the complexity of fulfilling imposed regulations.

Practitioners have employed a range of strategies to address privacy risks. The significant role of PETs is widely recognised across various application domains [40]; nevertheless, their adoption is complex and factor-dependent. The configuration and selection of PETs [56] facilitate and establish appropriate protection of clinical or genetic privacy, and support collaboration among a distributed network of healthcare institutions, seeking to develop AI models for further personalised drug and medicine development. Nevertheless, as reported [56], in the absence of concrete measures related to standardisation, policy, or technical or organisational measures to address privacy risks, improper selection and adoption of PETs may outweigh the benefits of their integration into clinical workflows. This makes the selection and adoption of PETs a relatively complex exercise for the OoC domain.

Current approaches to biomedical data predominantly address privacy risk primarily through governance requirements, focusing on controlled-access repositories [45]. These measures are primarily implemented during the data storage and data utilisation stages of the privacy-preserving lifecycle of medical data [47]. This highlights the importance of PETs, such as DP, for the remaining stages (data publishing and data mining) of the referred lifecycle. Integrating governance controls with DP, despite inherent trade-offs and complexities [56], [4], can substantially enhance privacy management by addressing privacy risks continuously at each stage of the relevant lifecycle. Together with a legal opinion, privacy risk requires case-by-case consideration and largely converges on the notion of *singling out/re-identification* under the GDPR, whereas, under the AIA, it is addressed from a *risk-management perspective* [4]. In this sense, DP demonstrates one of the most practical approaches to addressing a range of privacy risks and related attack vectors. Collectively, DP provides formal privacy guarantees for addressing certain privacy risks, supporting its inclusion in the current study.

EU Legislation GDPR does not separately define genomic data as a special category; rather, genomic data will generally be qualified as genetic data (*Art. 4(13), GDPR*). In addition, it could be related to health data (*Art. 4(15), GDPR*) when it reveals health-related information. *Art. 9, GDPR* defines genetic data and health data as special categories of personal data. Under *Art. 9(1), GDPR*, the processing of these data is prohibited, unless the exemption under (*Art. 9(2), GDPR*) applies. In the OoC context, related genomic data can be generated or arise from various sources and experimental outputs [72], [36], including, but not limited to, the patient-derived cells, disease-

specific variants, and omics readout. To maximise the value of such data and translate them into practical, tailored medicine, these data may be shared, for example, among institutions and labs, with the limitations discussed [52].

Furthermore, given the OoC context and the previously discussed GDPR constraints, when such genetically related data are reused [9] for secondary research purposes or shared via biobanks, platforms, institutional or research infrastructures, and repositories, these activities may require an additional assessment of the applicable legal basis. This includes considerations of applicable exemptions under (*Art. 9, GDPR*) and further clarification of controllers' roles and implementation of corresponding safeguards. Such safeguards include pseudoanonymisation, access controls, purpose limitation, data minimisation and governance measures, unless the data are fully anonymised. Full anonymisation may, however, introduce the trade-offs and restrictions discussed previously. For these reasons, PETs, such as DP, serve as appropriate *technical and organisational measures* and may be interpreted as reasonable safeguards under (*Art. 89(1), GDPR*). Such safeguards are particularly critical [71] when applied to the genetic data in biobanks and utilised within the OoC domain.

Current EU legislative regulations, including various data protection agencies' opinions and recommendations [23], [27], [22], as well as the Digital Omnibus proposal, aim to facilitate and clarify: (a) increased transparency for data controllers and processors, and (b) the interaction between the AIA and the GDPR. These efforts also seek to enhance privacy protections for data subjects by strengthening safeguards for sensitive data while providing reasonable, controlled access to diverse entities. Nevertheless, despite significant attention from academia and industry, the implications of these regulations for privacy and data protection remain (a) insufficiently examined, and (b) complex and challenging to implement. When processing occurs under public interest (Article 6(1)(e) GDPR) or legitimate interest (Article 6(1)(f) GDPR), further safeguards defined in Article 9(2)(j) GDPR must be implemented. Furthermore, AI systems utilised in these contexts must be evaluated under the AIA, particularly where high-risk classification or validation requirements apply.

AIA does not directly apply to OoC technologies; rather, it applies where an OoC-related application qualifies as an AI system within its scope and its associated risks (*Recital 26, AIA*). As highlighted in [53], AI systems used as medical devices or as components of medical devices may fall within the high-risk category and therefore be subject to stricter requirements. Furthermore, as outlined in *Art. 10(5)(b) and (c)*, it is required to enforce "*the special categories of personal data are subject to technical limitations on the re-use of the personal data, and state-of-the-art security and privacy-preserving measures, including pseudonymisation*" and "*the special categories of personal data are subject to measures to ensure that the personal data processed are secured, protected, subject to suitable safeguards...*" to prevent missues or unauthorised use. Although *Art. 10, AIA* overlaps with the GDPR, the AIA does not displace the GDPR requirements applicable to health and genetic data, including where such data are processed strictly for research and development purposes. Nevertheless, it raises further concerns that training on data from biobanks can increase privacy risks, including attacks such as inversion, membership inference, and data leakage through model updates, which are not related to the breach of the original dataset. These risks may be addressed, at least in part, through DP, within declared drawbacks [46].

The following observation of the Digital Omnibus proposal is subject to critical review and should be approached with caution, given the ongoing legislative process. Furthermore, academic sources assessing the implications of the Digital Omnibus proposal are limited. Nevertheless, we aim to highlight the most relevant changes that may affect the OoC domain. The Digital Omnibus proposal presents significant modifications to the definition of personal data under Article 4(1), which have already attracted objections regarding its adoption [24]. If the proposal is adopted in its

original form, diverse OoC actors may face increased scrutiny to demonstrate that their data are non-personal. This will require additional effort to document and defend the classification of non-personal data. Although the proposed improvements aim to reduce privacy risk by imposing obligations on OoC actors, these obligations may simply be transferred to downstream actors without effectively reducing privacy risk.

The OoC domain is particularly exposed to this uncertainty because data often move across a multi-stakeholder ecosystem that includes researchers, platform developers, biobanks, clinical partners, commercial providers, and regulatory users. In these settings, related data can generate derived datasets, sensor streams, imaging outputs, omics profiles, and AI-ready analytical outputs, making the classification of data as personal, pseudonymised, anonymised, or non-personal highly context-dependent.

Related work Unfortunately, biomedical data inherently contain personal and sensitive information, which can enable the identification of individuals [52]. Even when such data are coded or pseudonymised, linking them with other datasets can enable the re-identification of individuals, including individuals represented in small samples, due to ongoing technical advancements [60].

The authors [83] systematically assess security and privacy risks associated with biomedical data, focusing on the period from 2014 to 2025. Their approach addresses privacy concerns across the entire data lifecycle, from development frameworks to data analysis, utilising both the CIA [1] and NIST frameworks. A notable contribution of this work is the assessment and classification of health data by sensitivity. This classification is particularly relevant to the OoC domain, although its broader applicability remains unclear. However, the proposed framework has been validated solely through a case study, while real-world validation is pending. Although the review aims to encompass both EU and non-EU regulations; however, the focus on the GDPR is limited and does not address the AIA or the Digital Omnibus. Similarly, the review [1] examines broad issues of security and privacy in health data, with limited emphasis on EU regulations such as GDPR. The reviewer covers studies up to 2022 and underscores the significance of privacy risks.

Complementing the broader reviews discussed above, [77] primarily focused on identifying and examining existing methodologies for Privacy Impact Assessments (PIAs) and Privacy Risk Assessments (PRAs). The analysis demonstrates a lack of systematic evaluation and a need for further industry validation. Likewise, review [37], also highlights the need for further research on PIA methodologies. It confirms that existing approaches require further streamlining, as their complexity impacts privacy engineering practice.

Beyond privacy assessment methodologies, several reviews have examined anonymisation techniques and PETs. A mapping review conducted in 2020 [84] provides a comprehensive overview of software artefacts, conventional anonymisation techniques, and PETs. The review identifies several important research gaps, including the trade-off between achieving robust anonymisation and preserving data utility; the challenges of implementing robust anonymisation in the healthcare domain due to the high risk of re-identification; and the limited empirical validation of existing methods in real-world industrial settings. Similarly, the study [14] provides a comprehensive overview of the roles of various PETs and related techniques, such as private information retrieval (PIR) and private set intersection (PSI), in mitigating privacy and security risks, underscoring the significance of PETs in the biomedical domain.

[1] confidentiality integrity availability

The privacy trade-offs associated with synthetic data have also been evaluated across different implementation approaches [33]. These findings are broadly consistent with those reported by [48], which further emphasise the importance of PETs, including DP, while highlighting the complexity and multifaceted nature of biomedical data.

Beyond privacy and confidentiality implications, the study [79] also raises concerns about broader, unaddressed issues, such as potential re-identification through dataset sharing and cross-border data transfers (EU/non-EU and vice versa). The study systematically examines ethical challenges associated with OoC, emphasising the limited attention given to a) consent and re-consent measures, b) standardisation and governance issues, and c) environmental and sustainability impacts.

In summary, existing research has not yet produced an SLR that addresses privacy risks in the OoC domain within the context of the evolving EU regulations. Existing reviews primarily focus on biomedical privacy, risk management, privacy impact assessment methodologies, PETs, and ethical concerns in isolation. Thus, there remains a lack of a comprehensive synthesis integrating the perspectives discussed within OoC, highlighting the urgent need for the present study.

3 Systematic Literature Review Protocol

3.1 Method

To perform SLR, we follow the approach of Kitchenham [39] to collect, structure, and summarise related evidence and gaps in the previously referred domains. SLR aims to provide a rigorous and auditable methodology that could enable validation, review, and replication [41]. SLR aims to answer and map the information retrieved from selected datasets of studies on privacy risk and risk management for OoC in the EU, considering DP and GDPR, AIA, and the Digital Omnibus proposal.

3.2 Identifying relevant work

Following the guidelines [39] and recommendations [58] to ensure systematic and methodological rigour. The conducted SLR aims to map information from academic studies on privacy risks and mitigation strategies concerning OoC against the introduced legislation.

The initial phase of the SLR (**Fig.1**) consisted of defining search queries for the selected databases. Based on the study's objectives and research questions, the search was limited to the following domains: privacy risk, risk management, OoC, GDPR, AIA, and the Digital Omnibus Proposal. A two-part search query was developed to address both the OoC and privacy domains and was applied to author keywords, metadata, and all available searchable fields. These keywords were derived from the defined **RQs**, and the privacy-related key block was adopted from IEEE S&P 2026 topics of interest. The search strings and queries were tailored to accommodate the specific capabilities and interfaces of each database. In addition, search keywords before conducting the final run of searches were tested with various combinations across all targeted databases. Once the primary and final search queries were finalised, the search strings were structured into two parts: *Part 1* addressed OoC, while *Part 2* addressed privacy and related aspects. Consequently, the final search queries (**Table 1**) reflect the domain under study and the corresponding database-specific adaptations.

Beyond the OoC domain, specific terms were added, and the scope was broadened to include the various specific OoC chips elaborated in *Section 2*. SLR was conducted by searching and selecting

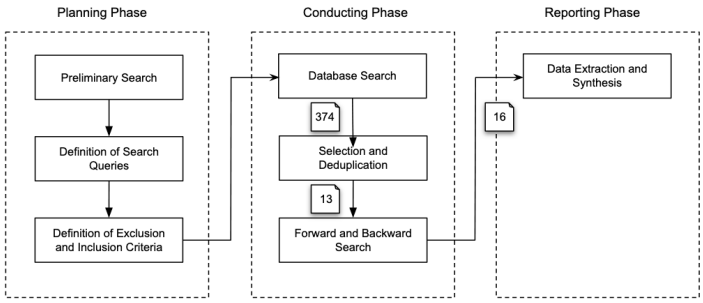

Fig. 1. SLR methodology overview adapted from [42]

Table 1. Summary of search criteria, databases, initial results, and final sample

Search criteria	Database	Initial	Final	Ref.
KEY: Part 1 AND ALL: Part 2	ACM Digital Library	2	–	–
FULL: Part 1 AND ALL: Part 2	IEEE Xplore	50	3	[81], [76], [8]
ALL: Part 1 AND ALL: Part 2	Web of Science	120	4	[79], [34], [29], [30]
KEY: Part 1 AND TITLE-ABS-KEY: Part 2	Scopus	105	3	[2], [64], [7]
Part 1 AND Part 2	SpringerLink	105	3	[63], [28], [78]

Note: **Part 1:** "organ?on?chip?" OR "organ?on?a?chip?" OR "microphysiological system?" OR "tissue chip?" OR "multiorgan-on-chip" OR "multi-organ-on-chip" OR "blood-brain barrier chip" OR "gut-on-chip" OR "lung-on-chip" OR "kidney-on-chip" OR "liver-on-chip" **Part 2:** "privacy" OR "data protection" OR "personal data" OR "health data" OR "patient data" OR "biomedical data" OR "risk assessment" OR "risk management" OR "GDPR" OR "AI Act" OR "privacy-enhancing technologies" OR "privacy enhancing technologies"

potentially relevant articles from the academic databases [2] using the search query (**Table 1**) and the complete SLR process shown in **Fig.1**. The search process, conducted in the second quarter of 2026, was designed to identify studies published from 2016 onward. Our initial search across the selected databases yielded **374** studies.

3.3 Filtering and Final Dataset

We screened all initial studies against our inclusion criteria and the scope of our analysis, as shown in **Fig.1**. The first iteration excluded **50** duplicates, potentially due to overlap across academic databases, resulting in **324** studies. For the remaining studies, we screened abstracts, including titles. Guided by our inclusion criteria, we aim to include any full-length academic study that (i) is written in English, (ii) covers topics related to the privacy and OOC, and (iii) utilises sensitive or personal data and is a primary study. The primary exclusion criteria are studies that do not address the research questions, duplicate titles, articles published as position papers or posters, articles unavailable through the university subscription, and articles not written in English. Furthermore, to ensure reliability in the filtering and selection process, the researcher utilised a literature management platform, such as Rayyan [55].

[2] https://dl.acm.org, https://ieeexplore.ieee.org, https://www.scopus.com, https://www.webofscience.com, https://link.springer.com

Each researcher independently and in a blinded manner coded the same set of **49** studies (*15%* of the dataset after deduplication), determining relevance by reviewing the title & abstract. The results were then compared and discussed until a consensus was achieved. The agreement between the researchers produced a Cohen's kappa value of **0.58**, while the observed percentage agreement was **87.8%**, indicating moderate inter-rater agreement. Subsequently, one researcher reviewed the final subset of studies and consulted another researcher only for studies with uncertain eligibility. This process resulted in the exclusion of **11** studies that did not meet the inclusion criteria after full-text assessment.

Following filtering and deduplication, an additional forward search[3] phase was conducted to identify further relevant studies. This process yielded **3** additional studies, bringing the final dataset to **16** studies. Final dataset consists of **16** studies, partly summarised in **Table 1**, as shown in **Fig. 2a** demonstrated the historical trend within the dataset. **Fig. 2b** report on the distribution of papers addressing privacy compared to the OoC. Academic contributions on the DP were underrepresented in the final dataset, with only two studies referencing this topic.

3.4 Dataset extraction and analysis

As discussed in **§2**, privacy challenges and risks within the OoC domain - including risk identification, mitigation, and the privacy–utility trade-off associated with PET implementation - remain insufficiently addressed. This issue is particularly significant given the increasing adoption of OoC technologies and the domain's primary emphasis on privacy risk governance, including the selective implementation of PETs, as noted in **§1**. Given that an OoC-specific privacy taxonomy and corresponding risk categories are not well established, this review develops provisional privacy risk categories that can be utilised in the current review and further examined for broader transferability. Since the primary objective of this study is to examine privacy risks in the OoC domain, a systematic and traceable process is required to develop an initial list of privacy risks and to ensure transparency across the domain. Developing these privacy risks requires substantial effort and may involve a degree of subjectivity, given the aims of the current study. This consideration motivates the adaptation of a qualitative hybrid *best-fit framework synthesis* [12], supplemented by descriptive quantitative mapping of reported evidence based on established aspects and concepts of privacy risk categories (rather than adapting and tailoring privacy risks to the OoC domain), as detailed below. We argue that the approach selected in the current study provides a more transparent and structured overview of high-level privacy risk categories in the OoC domain by considering multiple aspects and integrating multiple perspectives. It also provides a foundation for the subsequent refinement and further validation of the relevant privacy-risk categories. The selected approach is subject to the limitations discussed in **§7**.

As noted above, existing academic studies do not provide an OoC-dedicated privacy taxonomy or corresponding privacy risk/categories; therefore, the methodological design was informed by several complementary pieces of evidence. First, the study [44] demonstrates how AI privacy risk can be developed, including the adaptation of a chosen methodology for establishing AI risk taxonomies [69]. The demonstrated applicability of this methodology to the privacy domain motivated efforts to address the heterogeneity of privacy risks within the OoC domain. Second, in agreement with the CNIL [15], by informing with the explanation of privacy risk: ..."*What is a privacy risk? A risk is a hypothetical scenario that describes a feared event and all the threats that would allow this*

[3] Backward search was excluded mainly due to its recent emergence, applicability to the OoC domain, and limited academic contributions, particularly at the intersection of privacy and legislation

to occur..." and acknowledging the importance of privacy threat resources [67]. We opted for a hybrid best-fit framework synthesis [12] rather than a fully inductive thematic synthesis, which led to the development of high-level privacy-risk categories from heterogeneous evidence, applicable to the OoC domain (**Table 2**). *Note*: This approach was selected instead of directly adopting privacy risks from existing academic studies or mapping them to Solove's Taxonomy [70] due to the heterogeneity of privacy risks, the overlap among referenced risks, and the current study's previous unsuccessful attempts to apply these risks objectively across legal, technical, and biomedical domains.

An overview of the analysis protocol outlines its phases and relevant comments as applied to the final SLR dataset, which is shown in the *first phase* (**Table 3**). The development and analysis of the high-level privacy-risk categories consisted of several phases and are discussed further below.

Phase 2 focused on developing the *a priori* source-layer framework (§10, **Table 3**), identifying the fixed analytical function for each source layer, and extracting the analytical outputs contributed by each layer. This process aligns with the best-fit framework synthesis, in which related frameworks, concepts, and theories are identified and represented by reduced key elements in the a *priori* framework. *Priori* framework includes six source layers: [**L1**], LINDDUN [80] served as the primary technical taxonomy of privacy threats, clarifying concepts such as linking, identifying, detecting, information disclosure, unawareness, and non-compliance. *Note: Non-repudiation is omitted.* [**L2**], WP29 Opinion 05/2014 [5] guided the EU identifiability dimensions of singling out, linkability, and inference. *Note: concepts highlighted are anonymisation/identifiability evaluation criteria.* [**L3**], GDPR rights-based risks included transparency, purpose limitation, accountability, data-subject rights, and residual risk. [**L4**], AIA provisions addressed AI-related risks across the AI lifecycle and data governance. [**L5**], Digital Omnibus proposal utilised as an emerging legal context for uncertainty regarding the scope of personal data and the classification of pseudonymised or derived data. [**L6**], Privacy attacks and associated DP mitigation considerations informed the layer, including membership inference, reconstruction, model inversion, and statistical leakage. *Note: DP may mitigate the referred attacks.* The CEN/CENELEC OoC roadmap [13] was treated as a cross-cutting relevance lens and utilised as domain justification for applying these categories to OoC contexts involving human biological materials, associated data, data sharing, and multi-actor governance, rather than as a privacy-risk taxonomy. The referenced concepts were then utilised in *Phase 3* to develop provisional candidate privacy risk concerns.

Phase 3 is devoted to the development and retention of provisional risk codes (§10, **Table 5**); in simple terms, it converts analytical outputs from the six source layers and OoC/domain context into candidate privacy concerns and groups them by shared privacy-risk mechanisms. This phase makes a considerable contribution to the development of risk codes (*R1-R10*), which includes several steps: risk candidate concerns, grouping mechanism, retention rule, traceability, and finally harmonisation into high-level risk categories. We aim to keep (retain) the privacy risk concern only when three elements can be specified. It means satisfying each of the following elements (risk mechanism, source anchor/evidence and OoC pathway). Traceability rule applies following this trace: *source anchor → risk mechanism → OoC pathway → privacy consequence → provisional risk code. Note*: Closely related terms were grouped when they shared the same privacy-risk mechanism. We utilise the term *risk mechanism* [16] in a limited analytical sense, informed by mechanism-oriented reasoning in realist evaluation, to mean the privacy-relevant pathway through which an OoC data-processing context can create a privacy exposure, harm, or legal risk. The present review does not claim to conduct a full realist evaluation; rather, the term is used to explain how a privacy risk arises in context. We also use the term *source evidence/anchor* as an operational term to denote the recognised concept, category, legal criterion, or technical attack model that grounds a

Table 2. Privacy-risk categories (R1-R10) traceability matrix

Risk	[L1], **LINDDUN**	[L2], **WP29**	[L3], **GDPR**	[L4], **AIA**	[L5], **Digital Omnibus**	[L6], **DP**	**OoC relevance**
R1	✓	✓	✓	◑	✓	◑	✓
R2	✓	✓	✓	◑	✓	◑	✓
R3	◑	✓	✓	✓	◑	✓	✓
R4	✓	◑	✓	◑	◑	◑	✓
R5	✓	◑	◑	◑	◑	✓	✓
R6	◑	✓	✓	◑	◑	✓	✓
R7	◑	◑	✓	◑	◑	—	✓
R8	✓	—	✓	◑	◑	—	✓
R9	✓	—	✓	✓	◑	—	✓
R10	◑	✓	✓	◑	✓	◑	✓

Legend: ✓ = direct coverage; ◑ = partial, indirect, or conditional coverage; — = not identified as a distinct category

Note: This matrix inform the source-layer support for the provisional *R1-R10* codes after the risk-code development and retention step. It is a traceability table, not a second coding scheme. The columns should **not** be interpreted as equivalent taxonomies: LINDDUN provides privacy threat types, WP29 provides anonymisation and identifiability criteria, GDPR provides rights-and-freedoms and accountability relevance, the AIA provides AI lifecycle and data-governance relevance, Digital Omnibus proposal - provide emerging personal/non-personal data-scope context, DP informs technical privacy-attack mechanisms, and OoC relevance provides domain plausibility

retained privacy-risk category in *a priori* framework. The *priori* layers were connected through *risk mechanisms* rather than through word similarity. For example, **R1** was retained because LINDDUN *identifying*, WP29 *singling out*, GDPR *personal-data scope*, Digital Omnibus *scope uncertainty*, and OoC donor-specific data all concern the mechanism by which a person, donor, or sample can be distinguished or re-identified.

Phase 4 provides traceability between the retained *R1–R10* categories and the layers of the a *priori* source-layer framework. The retained R1-R10 privacy-risk categories developed in *Phase 3* are mapped back to the source layers in the traceability matrix (**Table 2**). Each retained risk category had to have: (i) a clear privacy-relevant mechanism, (ii) at least one source of evidence/anchor from *[L1]-[L6]* within an established privacy, legal, or technical framework, and (iii) plausible applicability to the OoC domain. *Phase 5* details the pilot coding process conducted on the test dataset (n=4) representing approximately *25%* of the final corpus. This phase assesses the applicability of *R1-R10* across the selected evidence. The pilot sample includes studies related to the OoC, privacy and the EU legislation domains. After the pilot sample corpus coded against *R1-R10*, the *Moderate* (**M**) option was included in the final coding phase because it more accurately represents the realistic characteristics of high-level risk categories. Additionally, *R5* demonstrated moderate applicability and was selected for further refinement. The review was conducted by a single reviewer; ambiguous cases were revisited after a time delay rather than being treated as independently adjudicated. *Phase 6* is devoted to coding the final SLR dataset against the high-level risk categories after the pilot coding and refinement of *Phase 5*. Additional quality control measures were implemented; however, these adjustments were minor and did not modify the risk categories established in *Phase 4*. The operational codebook for coding is presented in **Table 7**. *Note:* An Excel workbook served as the primary instrument for recording the coding decisions. Overlap rules require further clarification. *Phase 7*, as reported in §4, presents the distribution of the identified risk categories, their allocation by study ID and the general demographics of the included studies. The following sections provide further details on these aspects.

Table 3. Overview of the multilayer coding framework development process adapted from [12]

Phase	Name	Outcome	Comments / Methodological flow
1	Dataset overview	§3.1 **Table 1**	search → screening → eligibility → final included studies
2	*A priori* source-layer framework	§9, **Table 5**	L1–L6 source layers → deductive concepts Cross-cutting domain lens: OoC/CEN-CENELEC relevance → domain plausibility
3	Provisional risk-code development and retention	§9, **Table 6**	Source concepts + OoC contexts → candidate concerns → mechanisms → provisional codes
4	Multilayer traceability matrix	§2, **Table 2**	R1–R10 → mapped back to L1–L6 + OoC relevance
5	Pilot coding and framework refinement	§9, **Table 7**	Provisional codebook → pilot subset → refinement
6	Full coding and quality control	§9, **Table 7**	Estabilised R1–R10 codebook → final coding
7	Results synthesis and reporting	§4	Coded dataset → frequencies + qualitative interpretation

4 Results

Privacy risks in the OoC domain have received limited explicit attention compared to the ethical aspects in the reviewed academic literature. However, several emerging privacy-related concerns were identified. The limited focus on privacy risks appears to be partly attributable to the predominantly research-oriented context in which OoC technologies are currently developed and applied.

Firstly, the reviewed studies generally aimed to address OoC applications in controlled research environments for research purposes. Where an exemption from GDPR is due to the *Art. 9(2)* are aimed at being established.

The processing and sharing of sensitive data, biological samples, experimental outputs, and associated (meta)data beyond the original research setting were discussed only to a limited extent. When data sharing, biobank integration, or dataset combination was reported, these activities typically occurred within institutional research settings or under specific agreements between participating entities.

Second, the reviewed literature, in addition to technical OoC implications, identified the volume, complexity, and heterogeneity of OoC-related data as technical barriers to the prompt distribution, integration, and analysis of data across institutions. Inconsistent approaches to data labelling, storage, metadata description, documentation, and experimental reporting were also observed. These limitations reduced the interoperability and reusability of data across diverse stakeholders. Simultaneously, limited interoperability appeared to restrict linking datasets from different sources and could therefore indirectly reduce certain privacy risks.

The findings also demonstrate that the absence of standardisation and interoperability efforts may have a dual effect. Greater consistency in data collection, labelling, storage, and metadata can improve data quality, reproducibility, and reuse. However, such consistency may also facilitate the combination, comparison, and linkage of sensitive data and related datasets, such as biobanks. This may increase the likelihood of identifying, profiling, or inferring information about data subjects when OoC data remain linked to donor, clinical, genomic, or other sensitive information.

Inconsistencies in data standardisation were found to impact the development of tailored and personalised medicines and computational models used in drug development and validation. Limited interoperability reduced the transferability and reuse of data between laboratories and made

Table 4. Risk categories matrix

Study ID	Year	Ref.	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	Total
P01	2019	[81]	✓	✓	✓	✓	—	✓	—	—	—	✓	**6**
P02	2025	[63]	—	M	M	—	—	—	M	—	M	—	**0**
P03	2025	[79]	✓	✓	✓	✓	—	—	✓	✓	✓	✓	**8**
P04	2025	[34]	—	M	M	✓	—	M	M	M	M	M	**1**
P05	2025	[29]	—	M	M	✓	—	M	M	M	M	—	**1**
P06	2022	[2]	—	M	M	—	—	M	—	—	M	—	**0**
P07	2020	[30]	—	M	M	—	—	M	—	—	M	—	**0**
P08	2025	[64]	—	✓	✓	M	—	M	✓	✓	M	—	**4**
P09	2026	[7]	—	✓	M	✓	—	M	M	✓	✓	M	**4**
P10	2025	[76]	—	M	M	M	—	M	—	M	✓	—	**1**
P11	2024	[8]	—	M	M	M	—	—	M	M	M	M	**0**
P12	2023	[28]	—	✓	M	✓	—	—	—	—	M	M	**2**
P13	2019	[78]	—	✓	✓	—	—	M	—	—	✓	—	**3**
P14	2026	[51]	M	✓	✓	✓	—	M	—	—	—	—	**3**
P15	2026	[61]	✓	✓	✓	✓	—	—	—	✓	✓	✓	**7**
P16	2026	[66]	✓	✓	✓	M	—	—	—	✓	✓	—	**5**
Total			**4**	**9**	**7**	**8**	**0**	**1**	**2**	**5**	**6**	**3**	

Note: (✓) indicates that the risk category was included; (—) indicates excluded; (M) denotes partial or indirect coverage. Counts are provided exclusively for the (✓) category

cross-dataset linkage more difficult. The examined studies, therefore, revealed a tension between enhancing the scientific value of OoC data and limiting the privacy risks associated with broader data integration.

Applications involving (exp)AI, ML, DL, virtual organs or dedicated OoC systems, and related computational technologies, including algorithmic artefacts, were commonly described as developed or evaluated within bounded research environments. Privacy implications were less frequently addressed regarding the subsequent deployment, external distribution, or service-based availability of these artefacts, systems and OoC devices. The potential privacy risks associated with the transition from R&D development to wider implementation were therefore not examined systematically in the reviewed literature. This might also result from a multistage experimental environment with multiple phases, specifically applicable to the biomedical and OoC domains.

Only a limited number of the reviewed studies explicitly addressed data-governance mechanisms or privacy-preserving computational approaches. Federated learning was identified as a prominent utilisation of PETs, as well as the governance model applied to sensitive data. However, the reviewed studies provided limited discussion of the additional safeguards required to protect model updates, outputs, and other information exchanged during federated learning.

DP applications in the OoC domain remain limited and unrepresentative. Analysis indicates that DP utilisation in the OoC domain remains largely unexplored, owing to the importance of preserving biological fidelity and to uncertainties regarding the data-utility trade-off in this domain. These factors may partly explain the limited utilisation of DP in the reviewed studies. The absence of formal privacy-preserving mechanisms was particularly evident in scenarios involving cross-institutional data sharing, model distribution, and the potential public release of research data.

Finally, the primary regulatory focus of the reviewed studies was on legislation tailored to medical devices and other sector-specific or federal laws applicable to the OoC domain, including

aspects related to toxicity and FAIR principles. Limited attention is given to EU legislative norms that underscore data subjects' rights, due to a) taxonomy issues applied to the OoC domain, b) uncertain mechanisms to track data subjects' rights in a research environment and c) the necessity of data processors to demonstrate fulfilment of requirements where applicable.

4.1 SLR overview

Across all studies, nine of the ten risk categories were directly represented in at least one study, while risk category *R5* was not identified in any study. *R2* was the most frequently included risk category (n=9, 56.3%), followed by *R4* (n=8, 50.0%) and *R3* (n=7, 43.8%). Across the 160 possible study–risk combinations, 45 (28.1%) were coded as directly included (✓) and 65 (40.6%) as excluded. The (*M*) indication accounted for 31.3% of all coding decisions and 52.6% of the 95 instances in which a risk category was represented, either partly or indirectly. Based exclusively on the directly included categories, studies contained an average of 2.81 risk categories out of 10, corresponding to an average direct coverage of 28.1%. The analysis identified a persistent gap in the *R5* risk category, as reflected in the study-level coding matrix, and limited direct coverage of R6 (n=1, 6.3%). The high proportion of M classifications indicates that privacy risk categories were frequently acknowledged only indirectly or incompletely. All studies were published after the GDPR came into force in May 2018. Publications from 2025 and 2026 comprised 62.5% of the assessed study corpus.

Overall, applying the best-fit framework resulted in a consistent coding pattern. However, the substantial number of (*M*) indications for *R6* demonstrates that reporting of the privacy-risk category remains fragmented and frequently implicit, or shows high overlap between categories due to the high-level perspective. Refining the codebook against a pilot dataset representing (n=4, 25%) of the final corpus may not have captured all relative changes across the complete dataset. Alternatively, it is unclear whether the best-fit synthesis framework is more efficient when applied to a larger final corpus or more capable of performing analysis with fewer layers. The pilot codebook was also applied to excluded papers, as described in **§3.4** (*negative scenario*), to assess broader applicability and fulfil the framework requirements. The results demonstrated that the developed risk categories are applicable beyond the final included corpus. The best-fit synthesis framework appeared suitable for developing high-level risk categories across various EU legislation, with the anchors *L1-L6* specified in **Table 2**. However, no suitable method was identified to assess the quality of the selected approach or to determine representative metrics for integrating the referenced *L1-L6*. Additionally, challenges may arise in adopting the best-fit synthesis framework unless additional researchers are involved to validate uncertain cases, provide supplementary input to confirm results, or assist in resolving uncertainties. The inclusion of additional experts with biomedical and legal backgrounds in OoC could improve the organisation of high-level risk categories. Nevertheless, the granularity of these categories remains complex to define.

4.2 RQ1 overview

Across the included studies in the analysis, ten high-level categories were developed that are explicitly relevant for the OoC domain. Given (n=160, study risk combinations), only 45 combinations (*28.1%*) demonstrate direct coverage, and 64 (*40.0%*) showed no identifiable coverage. R2 (*Linkability*) is the most frequently represented risk category, followed by R4 (*Data disclosure/unauthorised access*) and R3 (*Sensitive inference*). Nine of the ten risk categories were directly represented in at

least one study, whereas R5 (*Membership inference/detectability*) was not registered as previously declared (*§3.4*). Considering possible overlap among *R1/R2*, *R3/R6*, and *R7/R8* underscores the importance of re-identification risks, data-reuse uncertainties, and consent issues. R2 and R3 risk categories reflect widespread concern about merging donor metadata, omics data, experimental data, and outputs, particularly for inferring health status, genotype, phenotype, tailored treatment response, or toxicity information. Although nine risk categories were directly identified, privacy risks were rarely analysed as independent research problems. Instead, privacy risks were commonly debated within ethics, governance, technical validation, and regulatory compliance, without explicit reference to the GDPR, AIA, or the Digital Omnibus proposal. These debates primarily emerged with the application of OoC technologies in tailored drug development and diseases such as cancer and its variants.

The absence of R5 indicates a gap between technically established privacy attacks and the risks explicitly considered in OoC research. Additionally, the indirect coverage of R9 (*Governance and responsibility ambiguity*) demonstrates ongoing uncertainty among diverse stakeholders, including researchers, indicating significant gaps that require further investigation even when governance measures are implemented. Overall, the findings demonstrate fragmented and predominantly indirect engagement with developed privacy risks categories. Nevertheless, the recorded data partly demonstrate that the introduced high-risk categories depend on the chosen frameworks, the layers *L1-L6*, and the assigned analytical functions for *R1-R10* (**Table 2**). Thus, it potentially explains the fragmented analysis presented and should be considered in future studies or study designs.

A combined perspective on the GDPR, the AIA, and the Digital Omnibus proposal as applied to the OoC domain, informed by analysis, demonstrates both the practicality of these frameworks and the complexity of implementation. This analysis leads to three principal observations. Firstly, the prevalence of privacy risk categories demonstrates that linkage, inference, and data disclosure sustain residual identification pathways. As a result, re-identification risk persists and should be addressed as an emergent concern. Secondly, technical and organisational measures depend on specific risk categories and should be aligned accordingly during the design phase. Thirdly, determining OoC roles and responsibilities remains complex, even when these roles are explicitly assigned.

4.3 RQ2 overview

The recorded evidence provides limited confirmation of the utilisation of PETs such as DP, despite the use of Federated learning. Partially confirmed by the frequency of the following risks: R6 (*Reconstruction/model inversion/statistical leakage*) informed by only one study [?], R10 presented only in three studies [?], [?], [?], and R5 (M*embership inference/detectability*) was not identified in any study, indicating limited consideration of established model-based privacy attacks. Rather than systematically employing PETs, most studies primarily addressed privacy risks through governance and organisational measures.

The reported evidence included the use of controlled-access repositories [?], access restrictions [?], data-sharing agreements [?], and governance mechanisms for managing sensitive biomedical data [?]. Also, FL is considered an alternative data-sharing method rather than a privacy-preserving approach. There was no evidence supporting the implementation of DP in any study, whether applied directly to the OoC dataset, during model training, or in analytical outputs. Thus, there is insufficient evidence to support the applicability of DP to OoC-specific applications. While PETs are indicated as relevant components of privacy risk mitigation, no explicit confirmation of their use was identified.

5 Discussion

The findings demonstrate that privacy risks in the OoC domain are largely unrecognised, only partially addressed, and often discussed alongside ethics, governance, interpretability, and technical development. The inconsistent identification of high-risk categories suggests that privacy implications within the OoC domain have not been sufficiently examined. Informed standardisation and interoperability issues have a dual effect. Enhancing reproducibility, data registration, data quality, and cross-laboratory collaboration may facilitate the integration of donor, clinical output, and experimental data. Furthermore, following recommendations and removal of direct identifiers may not fully mitigate re-identification risks when additional datasets are linked. The combined application of the GDPR, the AIA, and the Digital Omnibus proposal underscores both the significance and complexity of EU regulation in the OoC context. However, challenges persist in implementing technical and organisational measures, as well as in the legal interpretation of personal data.

6 Limitations

First, the scope of database coverage, including the selection of specific medical databases and the restriction to English-language studies, may have influenced the range of evidence collected. The review primarily considers academic studies, thereby excluding grey literature, position papers, white papers, opinion papers, book chapters, and industry reports, which could provide further evidence on current OoC practices and the stated RQ. The analytical approach and study design introduce further limitations. Study selection, data extraction, codebook development,codebook processing, codebook granularity, and final coding were conducted primarily by a single researcher, which introduces additional risks of selection and interpretation bias. The limited contributions of academics and experts in the medical, biomedical, policy, and legal domains may have affected the interpretation of domain-specific evidence and regulatory requirements. The analysis employed high-level privacy-risk categories, which enable comparability but may have obscured distinctions between overlapping or context-dependent privacy risks. Also, the selection of analytical functions and layers (L1-L6) may have introduced impacts that are difficult to observe. In addition, the risk categories overlap, and interpretation may affect the evidence reported. The selected analytical methodology influenced both the construction and interpretation of these categories. The Digital Omnibus was considered only in its proposal-stage form at the time of analysis; its provisions and legal implications may change following adoption or amendment.

Additional limitations should be considered. Issues related to the novelty and limited availability of research studies, as well as established taxonomies at the intersection of the OoC and privacy domains, also present challenges. Some relevant topics may be under submission or not yet released due to limitations in publishing, sharing datasets, and handling sensitive information.

7 Conclusion

The study examined privacy risks and PETs utilisation in the OoC domain employing the SLR and a best-fit framework synthesis. Integrating legal, technical and privacy-related aspects through layers, the study developed and applied ten high-level privacy risk categories to the selected study corpus. RQ1 - the evidence demonstrates that privacy risks in the OoC domain are primarily associated with linkability, sensitive inference, unauthorised disclosure, and, to a lesser extent, residual

indetifiability, secondary reuse, and ambiguity concerning governance and stakeholder responsibilities. Even though nine out of ten risk categories were recorded, privacy risks in studies are generally addressed primarily through debates in ethics, governance, technical OoC artefact validation, and regulatory compliance. Furthermore, the absence of R5, along with the limited representation of R6, suggests that privacy implications and associated attack vectors remain insufficiently examined. Evidence for RQ2 indicates that there are limited evaluations or utilisation of PETs in the OoC domain. PET's utility in mitigating OoC-related privacy risk cannot be established from the evidence collected. The finding further indicates that sanitising and removing direct identifiers may not eliminate re-identification risks when a given further linkable capability is available. Moreover, the best-fit framework synthesis demonstrates practical utility as a valuable tool for assessing privacy risks in emerging domains, particularly when tailored to specific legislative, international, or national requirements.

Overall, the study highlights an underdeveloped privacy risk landscape in the OoC domain, where ongoing technical and algorithmic advancements, as well as developments in personalised drug development and virtual organs, introduce new and emerging privacy risks. In parallel with advancements in the OoC domain, greater attention to responsibilities, legislative controls, and the taxonomy evolution of genetic data is demanded.

8 Acknowledgments

Funded by the European Union

This work is funded by the European Union and Germany (Grant Agreement Nos. 101140192 and 16MEE0484). Views and opinions expressed are however those of the author only and do not necessarily reflect those of the European Union, Germany, Chips Joint Undertaking or Bundesministerium für Forschung, Technologie und Raumfahrt. Neither the European Union, Germany nor the granting authorities can be held responsible for them. AI-based language assistance was utilized exclusively to improve grammar, punctuation, and readability.

Declaration of interests The authors report there are no competing interests to declare.

9 Appendix

Table 5. *Priori* framework overview: analytical functions, source layers and key source concepts

Layer	Source layer / *Type*	Fixed analytical function	Key source concepts extracted	CEN/CENELEC relevance
L1	LINDDUN / *Privacy threat framework*	Provides the privacy threat-modelling vocabulary used to identify privacy-relevant threat types	Linking; Identifying; Detecting; Disclosure of Information; Unawareness; Non-compliance	Relevant to translate privacy-threat concepts into OoC-relevant context involving data sharing, traceability, access, and governance across actors
L2	WP29 Opinion 05/2014 / *EU Regulatory guidance*	Provides the EU anonymisation and identifiability assessment lens.	Singling out; linkability; inference; residual-risk; anonymisation and pseudonymisation limits; contextual identifiability	Relevant because OoC workflows may involve donor-derived biological materials, associated metadata, omics outputs, and derived datasets whose identifiability must be assessed contextually
L3	GDPR / *EU legislation*	Provides the legal-risk and rights-and-freedoms layer for assessing personal-data scope, accountability, transparency, and high-risk processing	Personal-data scope; special-category data; transparency; purpose limitation; accountability; consent; DPIA high-risk logic; data-subject rights; residual risk; confidentiality	Relevant because the OoC data may involve human biological materials, omics, imaging, metadata, clinical linkage, datasets scope, model outputs, associated data, consent scope, sharing scope
L4	AIA / *EU legislation*	Provides the AI lifecycle, risk-management, and data-governance layer where OoC data are used for AI-enabled analysis or AI-ready datasets.	High-risk AI systems; risk management; data governance; training, validation, and testing data; model-output; bias detection; documentation; logging; transparency scope; AI lifecycle scope; governance;	Relevant where OoC data, imaging, sensor streams, omics, or computational models are used to develop or validate AI-supported systems, predictive models, or regulatory evidence
L5	Digital Omnibus proposal / *upcoming EU legislation*	Provides legal-context on clarifications of the personal/non-personal data boundary and classification of pseudonymised, anonymised, or derived data, datasets;	Personal/non-personal boundary; identifiability scope; pseudonymised-data scope; anonymisation scope; derived-data uncertainty; responsibility between actors; research reuse;	Relevant because OoC data, workflows may move through multi-actor ecosystems and produce derived outputs, making classification as personal, pseudonymised, anonymised, or non-personal context-dependent
L6	DP / *Privacy attacks*	Provides the technical safeguard, leakage mechanisms relevant to model outputs, statistical releases, synthetic data, and privacy-preserving analytics	Privacy-utility trade-off; membership inference; attribute inference; reconstruction; model inversion; data extraction; statistical leakage	Relevant where OoC data may be sensitive; donor-derived; datasets, aggregate outputs, synthetic data, omics summaries, or AI/ML models are shared, queried, or reused across actors

Note: This table demonstrates the source layers (& type) and their fixed analytical function in constructing the *a priori* framework before provisional high-level privacy risks categories are being developed. The aim is to identify each layer's function and underlying concepts, and to consider potential limitations. The analytical role of each source layer and the key source concepts extracted from each are identified. The extracted concepts are subsequently employed in phase three to generate provisional privacy-risk categories

Table 6. Phase 3, Provisional Risk-code development and retention logic

Candidate concern group	**Shared risk mechanism**	**Main source anchors**	**OoC pathway**	**Retention decision**	***Provisional* risk code**, *category*
Identifiability; re-identification; singling out; donor uniqueness; anonymisation failure; pseudonymisation limits	A person, donor, sample, or cell source becomes distinguishable, identifiable, or re-identifiable from data, metadata, or outputs	[L1] Identifying; [L2] singling out; [L3] personal-data scope; [L5] - related scope uncertainty	Donor-derived cells, rare-disease models, patient-specific iPSC lines, omics profiles, metadata, or clinical linkage may make a donor or sample distinguishable	Retained because the mechanism is source-grounded, plausible in OoC data processing, and operationally codable	**R1** *Identifiability / re-identification / singling out*
Linking; linkage; data combination; metadata linkage; provenance linkage; cross-dataset linkage.	Records, samples, datasets, metadata, or outputs can be connected across contexts, even without directly naming the person	[L1] Linking; [L2] linkability; [L3] pseudonymisation and anonymisation limits	OoC workflows may connect cell provenance, donor metadata, omics, imaging, sensor streams, biobank records, and clinical records	Retained because linkage is a recognised privacy mechanism and is plausible in multi-source OoC data ecosystems	**R2** *Linkability*
Genetic inference; disease inference; attribute inference; family risk; group risk; drug-response inference; toxicity-response inference	Sensitive health, genetic, disease, family, treatment-response, toxicity-response, or group information can be inferred from data or outputs	[L2] inference; [L3] special-category data; [L1] Linking/Identifying; [L6] attribute inference and model-inversion; [L4]] relevance where AI is used	OoC readouts may reveal disease status, genotype, drug response, toxicity sensitivity, family implications, or group-level health risks	Retained because the mechanism concerns inference of sensitive information and is directly relevant to donor-derived and disease-specific OoC contexts	**R3** *Sensitive inference*
Breach; leakage; unauthorised access; confidentiality failure; excessive access; third-party misuse	Personal data, special-category data, or sensitive research data become accessible beyond the intended, authorised, or necessary boundary	[L1] Disclosure of Information; [L3] security and confidentiality duties; DPIA risk logic; [L4] data-governance relevance where AI is involved	Cloud storage, CROs, pharma partners, AI vendors, platform providers, or data-sharing infrastructures may expose donor-linked data or derived outputs	Retained because confidentiality and access-boundary failures are core privacy and data-protection risks in multi-actor OoC workflows	**R4** *Data disclosure / unauthorised access*
Membership inference; inclusion inference; detectability; participation inference.	An observer can determine whether a person, donor, sample, cell line, or record was included in a dataset, model, query, or analysis	[L1] Detecting; [L6] membership-inference; [L3] sensitive-inclusion risk; [L4] model-output risk where AI is used	An AI model, query system, or statistical release trained on donor-derived OoC data may reveal whether a donor or sample participated	Retained because inclusion detectability is a recognised technical privacy mechanism and may reveal disease status, genetic status, or study participation	**R5** *Membership inference / detectability*
Reconstruction; model inversion; data extraction; statistical leakage; synthetic-data leakage; model-output leakage	Outputs, aggregate statistics, synthetic data, query responses, or trained models leak or reconstruct information about underlying donors, samples, or records	[L6] reconstruction, model-inversion, data-extraction, and statistical-leakage; [L2] inference; [L1] Disclosure/Identifying	Aggregate OoC outputs, omics summaries, synthetic datasets, query systems, or trained models may reveal donor- or sample-level information	Retained because output-based leakage is a distinct technical mechanism not reducible to ordinary unauthorised access	**R6** *Reconstruction / model inversion / statistical leakage*
Secondary use; downstream reuse; future unspecified use; AI training reuse; commercial reuse; regulatory reuse; cross-study reuse	Data, biological-material-derived outputs, or models are reused beyond the original purpose, consent, expectation, or governance context	[L3] purpose limitation, lawful basis, consent, and transparency principles; [L1] Non-compliance/Unawareness; [L4] data-governance relevance where AI reuse occurs	OoC data collected for one study may later be reused for AI training, pharma research, commercial benchmarking, regulatory evidence, or future unspecified research	Retained because downstream reuse is a central data-protection concern in research ecosystems involving human biological materials and derived data	**R7** *Secondary (re)use / purpose drift*
Unclear consent; poor transparency; donor misunderstanding; unclear downstream use; withdrawal difficulty; lack of control	Donors or data subjects do not understand, expect, or control downstream processing of their cells, data, derived outputs, or AI-ready datasets	[L1] Unawareness; [L3] transparency, consent, information duties, and data-subject rights	Donors, patients may not understand downstream use of cells, omics data, derived outputs, third-party sharing, AI model training, or commercial reuse	Retained because the mechanism concerns awareness, expectation, and control, which are central to data-protection and research-ethics governance	**R8** *Transparency / consent mismatch*
Controller/processor ambiguity; joint controllership; shared responsibility; vendor responsibility; third-party accountability; DPIA ownership; audit gaps	Responsibility for safeguards, data classification, access control, DPIA, downstream reuse, AI governance, or PET implementation is unclear	[L1] Non-compliance; [L3] accountability and controller-processor logic, compliance, roles [L4] provider/deployer and risk-management roles	OoC ecosystems may involve hospitals, biobanks, academic labs, CROs, pharma partners, cloud providers, AI vendors, platform developers, and regulators	Retained because multi-actor OoC workflows can create accountability gaps that affect privacy safeguards and data-subject rights	**R9** *Governance and responsibility ambiguity*
Anonymisation failure; pseudonymisation limits; de-identification limits; synthetic-data claims; DP/PET residual risk; non-personal-data claims	Data are treated as anonymous, non-personal, or sufficiently protected despite residual identifiability, linkability, inference, or model-leakage risk	[L2] singling out, linkability, and inference; [L3] personal-data boundary and residual-risk logic; [L5] - related scope debate; DP residual-risk reasoning	Derived OoC data, omics summaries, pseudonymised donor-linked datasets, synthetic data, or DP-protected outputs may be wrongly treated as non-personal	Retained because classification errors can affect the applicability of data-protection obligations and the adequacy of safeguards	**R10** *Misclassification as anonymous/non-personal or residual risk after safeguards*

Note: Table informs how candidate privacy concerns were keeped and harmonised into the provisional high-level privacy categories R1-R10 codes. A candidate concern was retained only where three elements could be specified: a privacy-relevant risk mechanism, at least one source anchor, a plausible OoC data-processing pathway. The table demonstrate how key source concepts were transformed into an a *priori*, OoC-adapted codebook before coding the final SLR dataset. Main source anchors aimed to have one framework or concepts involved

Table 7. R1-R10 privacy-risk codebook

Risk category	Include when	Exclude when
R1 *Identifiability / re-identification / singling out*	Re-identification; identifiability; singling out; donor, sample, or cell-source uniqueness; rare-disease identifiability; anonymisation failure; pseudonymisation limits; or contextual information that may make a person, donor, sample, or cell source distinguishable	The study merely uses human cells, human biological material, or human data without describing or implying an identifiability, re-identification, or singling-out pathway
R2 *Linkability*	Linking of records, datasets, donor metadata, cell provenance, omics, imaging, sensor streams, biobank records, clinical records, cross-institutional records, or other data sources; metadata linkage; provenance linkage; or data combination that may connect information across contexts	The study only refers to data sharing, collaboration, interoperability, or data integration in general terms without describing or implying linkage, combination, provenance, or cross-context connection risk
R3 *Sensitive inference*	Inference of disease status, genotype, phenotype, family risk, group risk, drug response, toxicity response, treatment response, health status, or other sensitive attributes from data, metadata, model outputs, biological readouts, or derived results	The study discusses prediction, modelling, classification, or analytical accuracy only as a scientific or technical objective, without describing or implying privacy-relevant inference of sensitive information
R4 *Data disclosure / unauthorised access*	Breach; leakage; unauthorised access; confidentiality failure; excessive access; third-party misuse; cloud exposure; weak access control; inappropriate sharing; or access to personal, special-category, donor-linked, or sensitive research data beyond the intended, authorised, or necessary boundary	The issue is limited to general cybersecurity, system robustness, availability, or technical performance without a connection to personal data, sensitive data, confidentiality, access control, or privacy-relevant disclosure
R5 *Membership inference / detectability*	Whether a person, donor, sample, cell line, record, or group was included in a dataset, training set, model, query system, study, experiment, or analysis; membership inference; inclusion inference; participation inference; or detectability of presence in a dataset or model	The study refers only to general privacy risk, re-identification, or data leakage without describing or implying inclusion, membership, participation, or detectability of presence in a dataset, model, query, or analysis
R6 *Reconstruction / model inversion / statistical leakage*	Reconstruction; model inversion; data extraction; leakage from aggregate statistics, query outputs, synthetic data, trained models, model parameters, gradients, or analytical outputs; or statistical disclosure of underlying donor-, sample-, or record-level information	The study discusses AI, modelling, synthetic data, aggregate analysis, or data sharing generally, but does not describe or imply leakage from outputs, models, statistics, query systems, synthetic data, or released analytical results
R7 *Secondary (re)use / purpose drift*	Reuse beyond the original purpose, consent, expectation, or governance context; downstream reuse; future unspecified research; AI training reuse; commercial reuse; regulatory reuse; cross-study reuse; secondary analysis; or data/material-derived output sharing for new purposes	Reuse remains clearly within the original stated purpose and no privacy, consent, governance, lawful-basis, transparency, or purpose-limitation concern is described or implied
R8 *Transparency / consent mismatch*	Unclear consent; inadequate notice; donor or data-subject misunderstanding; unclear downstream use; uncertainty about future use; withdrawal difficulty; lack of control; limited awareness of third-party sharing, AI training, commercial reuse, or derived-output processing	The study merely states that consent, ethics approval, or institutional approval was obtained without discussing limitations, uncertainty, downstream processing, donor understanding, withdrawal, transparency, or control
R9 *Governance and responsibility ambiguity*	Unclear controller, processor, joint-controller, provider, deployer, vendor, platform, biobank, CRO, pharma, cloud, or AI-provider responsibility; unclear DPIA ownership; unclear safeguard allocation; third-party accountability gaps; cross-border governance issues; audit gaps; or unclear responsibility for PET implementation	The study merely lists multiple actors, collaborations, institutions, platforms, or data recipients without describing or implying responsibility ambiguity, accountability gaps, governance uncertainty, or unclear allocation of safeguards
R10 *Misclassification as anonymous/non-personal or residual risk after safeguards*	Anonymisation, pseudonymisation, de-identification, synthetic data, DP, PETs, safe-data claims, non-personal-data claims, residual risk after safeguards, personal/non-personal data classification, or claims that data are sufficiently preserved despite possible identifiability, linkability, inference, or model-leakage risk	The study uses anonymisation, pseudonymisation, de-identification, DP, PET, or synthetic-data terminology only as a routine technical description, without connecting it to residual risk, legal classification, personal/non-personal data status, or privacy-risk assessment

Note: This table operationalises the final R1–R10 review codes for coding the included studies. A code is assigned where the study directly discusses, or indirectly implies, the corresponding privacy-risk mechanism. The categories are not mutually exclusive; multiple codes may be assigned where a passage or study describes more than one mechanism. Inclusion and exclusion rules were used to reduce subjective interpretation and to distinguish closely related risks, such as R1/R2, R3/R6, R7/R8, and R9/R10.

References

1. Al Zaabi, M., Alhashmi, S.M.: Big data security and privacy in healthcare: A systematic review and future research directions. Information Development **42**(2), 907–921 (2026)
2. Alciati, A., Reggiani, A., Caldirola, D., Perna, G.: Human-induced pluripotent stem cell technology: toward the future of personalized psychiatry. Journal of personalized medicine **12**(8), 1340 (2022)
3. Alver, C.G., Drabbe, E., Ishahak, M., Agarwal, A.: Roadblocks confronting widespread dissemination and deployment of organs on chips. Nature communications **15**(1), 5118 (2024)
4. Arasteh, S.T., Ziller, A., Kuhl, C., Makowski, M., Nebelung, S., Braren, R., Rueckert, D., Truhn, D., Kaissis, G.: Private, fair and accurate: Training large-scale, privacy-preserving ai models in medical imaging. arXiv preprint arXiv:2302.01622 (2023)
5. Article 29 Data Protection Working Party: Opinion 05/2014 on Anonymisation Techniques, `https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf`
6. Atreya, R.V., Smith, J.C., McCoy, A.B., Malin, B., Miller, R.A.: Reducing patient re-identification risk for laboratory results within research datasets. Journal of the American Medical Informatics Association **20**(1), 95–101 (2013)
7. Bai, L., Su, J.: Artificial intelligence virtual organoids (aivos). Bioactive Materials **59**, 45–68 (2026)
8. Baig, M.A.A., Li, S.: Ai-powered human-on-chip: Redefining biomedical frontiers. In: 2024 8th Asian Conference on Artificial Intelligence Technology (ACAIT). pp. 297–306. IEEE (2024)
9. Becker, R., Dove, E.S.: The eu gdpr and secondary use of health and genetic data for research support purposes. International Data Privacy Law p. ipag001 (2026)
10. Bernier, A., Liu, H., Knoppers, B.M.: Computational tools for genomic data de-identification: facilitating data protection law compliance. Nature Communications **12**(1), 6949 (2021)
11. Brauneck, A., Schmalhorst, L., Kazemi Majdabadi, M.M., Bakhtiari, M., Völker, U., Baumbach, J., Baumbach, L., Buchholtz, G.: Federated machine learning, privacy-enhancing technologies, and data protection laws in medical research: scoping review. Journal of medical Internet research **25**, e41588 (2023)
12. Carroll, C., Booth, A., Leaviss, J., Rick, J.: "best fit" framework synthesis: refining the method. BMC medical research methodology **13**(1), 37 (2013)
13. CEN/CENELEC Focus Group Organ-on-Chip: Focus Group Organ-on-Chip Standardization Roadmap. Roadmap, CEN/CENELEC (Jul 2024), `https://www.cencenelec.eu/media/publication-july-2024-fg-ooc-roadmap.pdf`, secretariat: NEN
14. Cho, H., Froelicher, D., Dokmai, N., Nandi, A., Sadhuka, S., Hong, M.M., Berger, B.: Privacy-enhancing technologies in biomedical data science. Annual review of biomedical data science **7**(1), 317–343 (2024)
15. Commission Nationale de l'Informatique et des Libertés: Privacy Impact Assessment: Methodology. Commission Nationale de l'Informatique et des Libertés, Paris, France, february 2018 edition edn. (Feb 2018), `https://www.cnil.fr/sites/cnil/files/atoms/files/cnil-pia-1-en-methodology.pdf`, pIA methodology guide
16. Dalkin, S.M., Greenhalgh, J., Jones, D., Cunningham, B., Lhussier, M.: What's in a mechanism? development of a key concept in realist evaluation. Implementation science **10**(1), 49 (2015)
17. Deznabi, I., Mobayen, M., Jafari, N., Tastan, O., Ayday, E.: An inference attack on genomic data using kinship, complex correlations, and phenotype information. IEEE/ACM transactions on computational biology and bioinformatics **15**(4), 1333–1343 (2017)
18. Dwork, C., Roth, A.: The algorithmic foundations of differential privacy. Foundations and trends® in theoretical computer science **9**(3-4), 211–487 (2014)
19. Eliazar, A., Brown, J.T., Cinamon, S., Kantarcioglu, M., Malin, B.: Re-identification risk for common privacy preserving patient matching strategies when shared with de-identified demographics. Journal of the American Medical Informatics Association **33**(2), 336–346 (2026)

20. European Commission: Proposal for a Regulation of the European Parliament and of the Council amending Regulations (EU) 2016/679, (EU) 2018/1724, (EU) 2018/1725, (EU) 2023/2854 and Directives 2002/58/EC, (EU) 2022/2555 and (EU) 2022/2557 as regards the simplification of the digital legislative framework, and repealing Regulations (EU) 2018/1807, (EU) 2019/1150, (EU) 2022/868, and Directive (EU) 2019/1024 (Digital Omnibus). COM(2025) 837 final, 2025/0360(COD) (2025), https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52025PC0837
21. European Data Protection Board: EDPB Document on response to the request from the European Commission for clarifications on the consistent application of the GDPR, focusing on health research, https://www.edpb.europa.eu/system/files/documents/files/file1/edpb_replyec_questionnaireresearch_final.pdf
22. European Data Protection Board: Opinion 28/2024 on certain data protection aspects related to the processing of personal data in the context of AI models, https://www.edpb.europa.eu/system/files/documents/2024-12/edpb_opinion_202428_ai-models_en.pdf
23. European Data Protection Board: Guidelines 01/2025 on pseudonymisation. https://www.edpb.europa.eu/system/files/2025-01/edpb_guidelines_202501_pseudonymisation_en.pdf (2025), adopted 16 January 2025, accessed 5 July 2026
24. European Data Protection Board, European Data Protection Supervisor: EDPB-EDPS Joint Opinion 2/2026 on the Proposal for a Regulation as regards the simplification of the digital legislative framework (Digital Omnibus) (Feb 2026), https://www.edps.europa.eu/system/files/2026-02/edpb_edps_jointopinion_202602_digitalomnibus_en.pdf, adopted on 10 February 2026; accessed 11 July 2026
25. European Parliament and Council of the European Union: Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679 (2016), official Journal of the European Union, L 119, 4 May 2016
26. European Parliament and Council of the European Union: Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act). https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng (2024), official Journal of the European Union, L, 2024/1689, 12 July 2024
27. European Union Agency for Cybersecurity: Deploying Pseudonymisation Techniques: The Case of the Health Sector, https://www.enisa.europa.eu/sites/default/files/publications/WP2021%20-%20O.2.3%20Pseudonymisation%20Healthcare%20.pdf
28. Franklin, I.B., Bhuvaneswari, R., Vasanthi, V., Jerald, M.P.A.: Replica controlled sensor enabled architecture for management of electronic health records. International Journal of Information Technology **15**(7), 3643–3653 (2023)
29. Hendriks, H.: Advancing oncology drug development: Innovative approaches to enhance success rates while reducing animal testing. Biochimica et Biophysica Acta (BBA)-Reviews on Cancer p. 189467 (2025)
30. Herland, A., Maoz, B.M., Das, D., Somayaji, M.R., Prantil-Baun, R., Novak, R., Cronce, M., Huffstater, T., Jeanty, S.S., Ingram, M., et al.: Quantitative prediction of human drug pharmacokinetic responses using multiple vascularized organ chips coupled by fluid transfer. Nature biomedical engineering **4**(4), 421 (2020)
31. Huh, D., Matthews, B.D., Mammoto, A., Montoya-Zavala, M., Hsin, H.Y., Ingber, D.E.: Reconstituting organ-level lung functions on a chip. Science **328**(5986), 1662–1668 (2010)
32. Humbert, M., Ayday, E., Hubaux, J.P., Telenti, A.: Addressing the concerns of the lacks family: quantification of kin genomic privacy. In: Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security. pp. 1141–1152 (2013)
33. Hyrup, T., Lautrup, A.D., Zimek, A., Schneider-Kamp, P.: A systematic review of privacy-preserving techniques for synthetic tabular health data. Discover Data **3**(1), 5 (2025)

34. Ibrahim, M.S., Kim, M.: Artificial intelligence-aided microfluidic cell culture systems. Biosensors **16**(1), 16 (2025)
35. Ingber, D.E.: Is it time for reviewer 3 to request human organ chip experiments instead of animal validation studies? Advanced Science **7**(22), 2002030 (2020)
36. Ingber, D.E.: Human organs-on-chips for disease modelling, drug development and personalized medicine. Nature Reviews Genetics **23**(8), 467–491 (2022)
37. Iwaya, L.H., Alaqra, A.S., Hansen, M., Fischer-Hübner, S.: Privacy impact assessments in the wild: A scoping review. Array **23**, 100356 (2024)
38. Jiang, L., Li, Q., Liang, W., Du, X., Yang, Y., Zhang, Z., Xu, L., Zhang, J., Li, J., Chen, Z., et al.: Organ-on-a-chip database revealed—achieving the human avatar in silicon. Bioengineering **9**(11), 685 (2022)
39. Keele, S., et al.: Guidelines for performing systematic literature reviews in software engineering. Tech. rep., Technical report, ver. 2.3 ebse technical report. ebse (2007)
40. Khan, F., Shah, S.A., Tahir, S., Yasin Ghadi, Y., Shah, S.I., Zahid, A., Ahmad, J., Hussain Abbasi, Q.: Privacy enhancing technologies for intelligent healthcare: Research challenges and opportunities. ACM Computing Surveys **58**(7), 1–27 (2026)
41. Kitchenham, B., et al.: Procedures for performing systematic reviews. Keele, UK, Keele University **33**(2004), 1–26 (2004)
42. Klymenko, A., Meisenbacher, S., Favaro, L., Matthes, F.: Supporting the integration of privacy-enhancing technologies into the software development life cycle. SN Computer Science **6**(6), 592 (2025)
43. Kujala, V.J., Pasqualini, F.S., Goss, J.A., Nawroth, J.C., Parker, K.K.: Laminar ventricular myocardium on a microelectrode array-based chip. Journal of Materials Chemistry B **4**(20), 3534–3543 (2016)
44. Lee, H.P., Yang, Y.J., Von Davier, T.S., Forlizzi, J., Das, S.: Deepfakes, phrenology, surveillance, and more! a taxonomy of ai privacy risks. In: Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems. pp. 1–19 (2024)
45. Lin, D., McAuliffe, M., Pruitt, K.D., Gururaj, A., Melchior, C., Schmitt, C., Wright, S.N.: Biomedical data repository concepts and management principles. Scientific Data **11**(1), 622 (2024)
46. Liu, C., Chakraborty, S., Mittal, P.: Differential privacy under dependent tuples. NDSS, Dependence makes you vulnberable (2016)
47. Liu, W., Zhang, Y., Yang, H., Meng, Q.: A survey on differential privacy for medical data analysis. Annals of Data Science **11**(2), 733–747 (2024)
48. Liu, Y., Acharya, U.R., Tan, J.H.: Preserving privacy in healthcare: A systematic review of deep learning approaches for synthetic data generation. Computer Methods and Programs in Biomedicine **260**, 108571 (2025)
49. Low, L.A., Mummery, C., Berridge, B.R., Austin, C.P., Tagle, D.A.: Organs-on-chips: into the next decade. Nature Reviews Drug Discovery **20**(5), 345–361 (2021)
50. Marx, U., Andersson, T.B., Bahinski, A., Beilmann, M., Beken, S., Cassee, F.R., Cirit, M., Daneshian, M., Fitzpatrick, S., Frey, O., et al.: Biology-inspired microphysiological system approaches to solve the prediction dilemma of substance testing. Altex **33**(3), 272 (2016)
51. Moghimi, N., Rezaeian, M., Kohandel, M.: Precision oncology paradigm: Integrating tumor-on-chip platforms, mathematical modeling, and ai for personalized cancer therapeutics. Journal of Drug Delivery Science and Technology p. 108722 (2026)
52. Molnár-Gábor, F., Korbel, J.O.: Genomic data sharing in europe is stumbling—could a code of conduct prevent its fall? EMBO molecular medicine **12**(3), EMMM201911421 (2020)
53. Niemiec, E., Davis, P.A.E., Hauglid, M.: Will the eu ai act help to eliminate dataset bias in medical ai? Available at SSRN 5045561 (2024)
54. Ohno-Machado, L., Jiang, X., Kuo, T.T., Tao, S., Chen, L., Ram, P.M., Zhang, G.Q., Xu, H.: A hierarchical strategy to minimize privacy risk when linking "de-identified" data in biomedical research consortia. Journal of biomedical informatics **139**, 104322 (2023)

55. Ouzzani, M., Hammady, H., Fedorowicz, Z., Elmagarmid, A.: Rayyan—a web and mobile app for systematic reviews. Systematic reviews **5**(1), 210 (2016)
56. Pati, S., Kumar, S., Varma, A., Edwards, B., Lu, C., Qu, L., Wang, J.J., Lakshminarayanan, A., Wang, S.h., Sheller, M.J., et al.: Privacy preservation for federated learning in health care. Patterns **5**(7) (2024)
57. Pediaditakis, I., Kodella, K.R., Manatakis, D.V., Le, C.Y., Hinojosa, C.D., Tien-Street, W., Manolakos, E.S., Vekrellis, K., Hamilton, G.A., Ewart, L., et al.: Modeling alpha-synuclein pathology in a human brain-chip to assess blood-brain barrier disruption. Nature communications **12**(1), 5907 (2021)
58. Petersen, K., Vakkalanka, S., Kuzniarz, L.: Guidelines for conducting systematic mapping studies in software engineering: An update. Information and software technology **64**, 1–18 (2015)
59. Piergiovanni, M., Leite, S.B., Corvi, R., Whelan, M.: Standardisation needs for organ on chip devices. Lab on a Chip **21**(15), 2857–2868 (2021)
60. Rocher, L., Hendrickx, J.M., De Montjoye, Y.A.: Estimating the success of re-identifications in incomplete datasets using generative models. Nature communications **10**(1), 3069 (2019)
61. Saini, R., Thakur, B., Basaba, B.K., Satapathy, M.K.: Artificial intelligence–enabled organoid platforms for precision medicine: Integrating multi-omics, digital twins, and microphysiological systems. Organoids **5**(3), 20 (2026)
62. Scheibner, J., Raisaro, J.L., Troncoso-Pastoriza, J.R., Ienca, M., Fellay, J., Vayena, E., Hubaux, J.P.: Revolutionizing medical data sharing using advanced privacy-enhancing technologies: technical, legal, and ethical synthesis. Journal of medical Internet research **23**(2), e25120 (2021)
63. Sheng, Q.S., Liu, B., Wang, X., Hua, L., Zhao, S.C., Sun, X.Z., Li, M.Y., Zhang, X.Y., Wang, J.X., Hu, P.L.: Revolutionizing toxicological risk assessment: integrative advances in new approach methodologies (nams) and precision toxicology. Archives of Toxicology **99**(12), 4697–4707 (2025)
64. Sillé, F., Smirnova, L., Hartung, T.: Microphysiological systems as a pillar of the human exposome project. Journal of Biological Chemistry p. 110782 (2025)
65. da Silva, R.G.L., Blasimme, A.: Organ chip research in europe: players, initiatives, and policies. Frontiers in bioengineering and biotechnology **11**, 1237561 (2023)
66. Singh, J., Singh, P.: Synergizing multi-omics and artificial intelligence to unravel gut microbiome complexity: A review. Innovative Medicines & Omics p. 025490068 (2026)
67. Sion, L., Van Landuyt, D., Joosen, W.: Leveraging the domain experts: specializing privacy threat knowledge. In: European Symposium on Research in Computer Security. pp. 534–541. Springer (2024)
68. Skardal, A., Devarasetty, M., Forsythe, S., Atala, A., Soker, S.: A reductionist metastasis-on-a-chip platform for in vitro tumor progression modeling and drug screening. Biotechnology and bioengineering **113**(9), 2020–2032 (2016)
69. Slattery, P., Saeri, A.K., Grundy, E.A., Graham, J., Noetel, M., Uuk, R., Dao, J., Pour, S., Casper, S., Thompson, N.: The ai risk repository: A comprehensive meta-review, database, and taxonomy of risks from artificial intelligence. SuperIntelligence-Robotics-Safety & Alignment **1**(1) (2024)
70. Solove, D.J.: A taxonomy of privacy. U. Pa. l. Rev. **154**, 477 (2005)
71. Staunton, C., Slokenberga, S., Parziale, A., Mascalzoni, D.: Appropriate safeguards and article 89 of the gdpr: considerations for biobank, databank and genetic research. Frontiers in genetics **13**, 719317 (2022)
72. Tagle, D.: Organs-on-chips: into the next decade. Nature Reviews Drug Discovery (2020)
73. Thakar, R.G., Fenton, K.N.: Bioethical implications of organ-on-a-chip on modernizing drug development. Artificial organs **47**(10), 1553–1558 (2023)
74. Thomas, M., Mackes, N., Preuss-Dodhy, A., Wieland, T., Bundschus, M.: Assessing privacy vulnerabilities in genetic data sets: scoping review. JMIR Bioinformatics and Biotechnology **5**(1), e54332 (2024)
75. Van Den Berg, A., Mummery, C.L., Passier, R., Van der Meer, A.D.: Personalised organs-on-chips: functional testing for precision medicine. Lab on a Chip **19**(2), 198–205 (2019)
76. Vijayaragavan, V., Vijayakumar, M., Dharshini, P., Karunakaran, K., Kesavan, K., Raj, R.D.S.: Ai and automation in organ-on-a-chip technology: Advancing personalized medicine and disease modeling. In: 2025 IEEE 6th International Conference in Robotics and Manufacturing Automation (ROMA). pp. 40–45. IEEE (2025)

77. Wairimu, S., Iwaya, L.H., Fritsch, L., Lindskog, S.: On the evaluation of privacy impact assessment and privacy risk assessment methodologies: A systematic literature review. IEEE Access **12**, 19625–19650 (2024)
78. Walker, M.J., Bourke, J., Hutchison, K.: Evidence for personalised medicine: mechanisms, correlation, and new kinds of black box: Mj walker et al. Theoretical medicine and bioethics **40**(2), 103–121 (2019)
79. Weidema, J., De Vries, M., Mummery, C., De Graeff, N.: The ethical aspects of human organ-on-chip models: a mapping review. Stem Cell Reports (2025)
80. Wuyts, K., Joosen, W.: Linddun privacy threat modeling: a tutorial. CW Reports (2015)
81. Yan, X., Cui, B., Xu, Y., Shi, P., Wang, Z.: A method of information protection for collaborative deep learning under gan model attack. IEEE/ACM Transactions on Computational Biology and Bioinformatics **18**(3), 871–881 (2019)
82. Yu, W.: Organs-on-chips: Discussing the technology and ethical implications of biomimetic microfluidic chips. A Letter From the Editors p. 29
83. Zhang, S., Singh, M.M.: Privacy and security in health big data: A nist-guided systematic review of technologies, challenges, and future directions. Information **17**(2), 148 (2026)
84. Zuo, Z., Watson, M., Budgen, D., Hall, R., Kennelly, C., Al Moubayed, N.: Data anonymization for pervasive health care: systematic literature mapping study. JMIR medical informatics **9**(10), e29871 (2021)