Ethics and Information Technology (2025) 27:32
https://doi.org/10.1007/s10676-025-09843-4

ORIGINAL PAPER

A critique of current approaches to privacy in machine learning

Florian van Daalen[1,2] · Marine Jacquemin[1] · Johan van Soest[1,5] · Nina Stahl[6] · David Townend[3,6] · Andre Dekker[1] · Inigo Bermejo[1,4]

Published online: 20 June 2025

Abstract
Access to large datasets, the rise of the Internet of Things (IoT) and the ease of collecting personal data, have led to significant breakthroughs in machine learning. However, they have also raised new concerns about privacy data protection. Controversies like the Facebook-Cambridge Analytica scandal highlight unethical practices in today's digital landscape. Historical privacy incidents have led to the development of technical and legal solutions to protect data subjects' right to privacy. However, within machine learning, these problems have largely been approached from a mathematical point of view, ignoring the larger context in which privacy is relevant. This technical approach has benefited data-controllers and failed to protect individuals adequately. Moreover, it has aligned with Big Tech organizations' interests and allowed them to further push the discussion in a direction that is favorable to their interests. This paper reflects on current privacy approaches in machine learning and explores how various big organizations guide the public discourse, and how this harms data subjects. It also critiques the current data protection regulations, as they allow superficial compliance without addressing deeper ethical issues. Finally, it argues that redefining privacy to focus on harm to data subjects rather than on data breaches would benefit data subjects as well as society at large.

Keywords Privacy · Ethics · Machine learning · Privacy preserving

Background

The promise to deliver innovation in fields as diverse as healthcare, transportation and education has made it difficult to ignore the appeal of collecting and processing vast amounts of personal data. Access to large datasets, the rise of the Internet of Things (IoT), and the ease of collecting personal data, have led to significant breakthroughs in machine learning (Kairouz et al. 2021; Xu et al. 2021; Hagen 2021; Truong et al. 2021). However, they have also raised concerns about privacy data protection (Bak et al. 2024). Awareness of privacy issues in the era of Big Data is growing, fuelled by recent controversies such as the Facebook-Cambridge Analytica scandal (Hinds et al. 2020; ur Rehman

✉ Florian van Daalen
f.vandaalen@maastrichtuniversity.nl

Marine Jacquemin
marine.jacquemin@maastro.nl

Johan van Soest
j.vansoest@maastrichtuniversity.nl

Nina Stahl
n.stahl@maastrichtuniversity.nl

David Townend
d.townend@maastrichtuniversity.nl

Andre Dekker
andre.dekker@maastro.nl

Inigo Bermejo
i.bermejo@maastrichtuniversity.nl

[1] Radiation Oncology (MAASTRO) GROW School for Oncology and Reproduction, Maastricht University Medical Centre, Maastricht, Netherlands

[2] Department of Health Promotion, Care and Public Health Research Institute (CAPHRI), Maastricht University, Maastricht, Netherlands

[3] City Law School, University of London, London, United Kingdom

[4] Data Science Institute, Hasselt University, Hasselt, Belgium

[5] Brightlands Institute for Smart Society (BISS), Faculty of Science and Engineering, Maastricht University, Maastricht, Netherlands

[6] Department of Health, Ethics and Society (HES), Faculty of Health, University Maastricht, Maastricht, Netherlands

2019), Clear View AI (Ahmed 2023), and reports from privacy watchdogs like the Mozilla Foundation (Foundation 2023), which highlighted the unethical practices that have become commonplace in today's digital landscape. Personal data is extremely valuable (Palmer 2005) and often harvested without the knowledge or consent of individuals, leading to potentially negative consequences, not only for them, but also for society as a whole.

Partially in response to these concerns, European lawmakers adopted the General Data Protection Regulation (GDPR) in 2016 (GDPR 2016). In the US, the State of California soon followed suit, implementing the Californian Consumer Privacy Act (CCPA) in California Code Civil Code (2023), soon amended by the California Privacy Right Act (CPRA). These two acts however only apply to Californians, and there is no federal-level data protection regulation in the United States outside of the Health Insurance Portability and Accountability Act (HIPAA), which only applies to health data. Both the GDPR and CCPA provide enforceable rights to data subjects and clearly define the notion of lawful data processing, with real repercussions in case of non-compliance (EDPB 2023; GDPR Enforcement Tracker 2024; Register 2024; GDPR 2016; Euronews 2022; Cromack 2021; Euronews 2022).[1] This formed the perfect context to encourage the further development of so-called "privacy-preserving" data analysis solutions enabling machine learning models to be trained without compromising privacy and therefore avoiding data protection related fines. Various metrics, such as k-anonymity Sweeney (2002), sensitivity and ϵ-differential privacy Dwork and Roth (2015) and Dwork et al. (2006), have been established, advertised as ways to measure such privacy-preservation in an objective and generalizable manner.

This paper aims to critically reflect on the current approaches to privacy in machine learning. First, we will briefly introduce the concept of privacy as understood within social science and law. We will argue that privacy has increasingly been approached as a mathematical concept, explaining how this technical approach, while beneficial for data-controllers, fails to protect the interests of data subjects. Next, we will consider the role of Big Tech in defining what should or should not be considered private, and how their influence significantly impacts the social understanding of privacy. Finally, we will discuss how the current situation might be improved to benefit individuals and society as large, by arguing for a shift from privacy-preserving machine learning towards an approach focused on risk assessment and harm mitigation.

[1] See GDPR, Chapter 8. Specifically, GDPR, Article 83 (4–6), and Article 84.

What even is privacy?

To meaningfully assess the question of privacy preservation, we must first understand the concept of privacy itself. This concept is complex, multi-faceted and context-dependent, having been explored extensively among disciplines such as social science, law, and computer science. This section will focus on the first two, offering the knowledge basis necessary to understand how they differ from the understanding of privacy common within data science and machine learning. One of the first known definitions of privacy in law was given in 1890 by Warren and Brandeis Warren and Brandeis (1890), presenting it as "the right to be left alone", as a response to the increasingly intrusive behavior of the newspapers and paparazzi of the time. Their work catalyzed a broader conversation around individual control over access to one's personal life. Since then, layered and evolving understandings of privacy have been developed within social science. Different disciplinary understandings of privacy' can be summarized as frameworks developed to present the key nuances of privacy as a concept, and how it was shaped by our contemporary society. Indeed, privacy being a flexible and evolving concept Tavani (2008), it evades a clearly set definition.

That privacy is conceptually contested between and within disciplines needs further explanation. There is little agreement about the limits of privacy.[2] This can be explained by the dynamic nature of privacy; by its nature it is personal, defines the world of the individual claiming the right, and is defined to a very large extent by the person claiming the right. We agree the first line, "we have a human right to privacy", but there is very little agreement about the granularity of privacy thereafter—there is no agreed second line, "this is your privacy".[3] When one examines the case law on privacy of, for example, the European Court of Justice, privacy disputes are resolved, but this is arguably a dispute resolution between two parties where privacy is identified when it is seen, rather than creating a canon of law that prescribes the granular detail of privacy as a normative template to use moving forward. European Court of Human Rights (2022) Fundamentally, is "privacy"—a right to a private life—a right that can or should be defined, or constrained? Privacy is about, Laurie argues, one's space to be who one wishes to be, within broad (and not narrow)

[2] For a contemporary discussion of the range of definitions of privacy, see Roessler, Beate and Judith DeCew, "Privacy", The Stanford Encyclopedia of Philosophy (Winter 2023 Edition), Edward N. Zalta & Uri Nodelman (eds.) Roessler and DeCew (2023).

[3] This is explored in more detail in Townend, D (2024) "Big Data Research: can confidentiality and fiduciary duties fill the gaps in privacy and data protection?" Townend (2025).

societal constraints.[4] Laurie argues that privacy is the space where we negotiate our relationship with others in society. Townend has argued that the way of negotiating that space should be a claim, by the person seeking to challenge an individual's privacy, to a necessity so to do in the public interest. He argues that one's privacy can be 'breached' (as privacy in the international instruments is not an absolute right) where doing so is in a supervening public interest. This is not necessarily simply a Utilitarian approach. It can be argued as a question of reasonableness in Rawls, or an operation of the Categorical Imperative in making a claim to privacy without instrumentalising others.[5]

It is necessary to apply this conceptual perspective to more practical issues. One widely accepted perspective, closely tied to Warren & Brandeis' initial definition, is the idea of privacy as a form of control (Moor 2006; Moore 2008; Macnish 2018), specifically over access to one's body, communications, decisions, and personal data.[6] Furthermore, one's understanding of what is or should be private may differ based on context. For example, sharing personal health information with a doctor would not be considered an attack on privacy, whereas that same information being accessed by financial institutions would. This idea is the central thesis of Nissenbaum Nissenbaum (2009) and Roessler and DeCew (2023), who argues, through a concept known as contextual integrity, that our privacy expectations are shaped by the norms governing different social contexts, such as the doctor-patient relationship in our example. Privacy has also been closely linked to autonomy and freedom (Dean 2003; Bundesverfassungsgericht 1983; Solove 2002; Peloquin et al. 2020; Cohen 2012), with authors suggesting that attacks on privacy undermine individuals' capacity to think and act independently. This concern is particularly relevant in the digital era, when information is increasingly being concentrated in only a few powerful data-driven institutions. Individuals being consistently monitored, profiled, or influenced based on personal information about them may self-censor or alter their behavior. Privacy can thus also be understood as a matter of power. Control over personal data gives power over not only these individuals, but also entire populations. This dynamic stands at the core of what Zuboff describes as "surveillance capitalism" Zuboff (2023), where personal information is commodified and used to predict ad manipulate behavior at a large scale. It highlights the importance of privacy as a way to maintain power, given that erosion of that right results in individual and systemic harm.

Social science and legal understanding of privacy have historically mutually shaped one another Roessler and DeCew (2023). European law explicitly recognizes privacy as a fundamental right, enshrining it in article 8 of the European Convention on Human Rights, which states "everyone has the right to respect for his or her private and family life, home and communications". However, a precise definition of what constitutes such "privacy" from a legal perspective is lacking. In contrast, the closely entangled concept of "personal data" is clearly defined as data that can be linked back to an individual. As such, it forms the basis of European data protection regulation. This framework is often understood as a means to protect or safeguard privacy, despite data protection and privacy being often conflated but existing in the European Union as distinct fundamental rights (Fuster and Gellert 2012; Fuster 2014; European Union 2007). Indeed, it is entirely possible to abide by data protection regulation without respecting one's right to privacy (De Hert and Gutwirth 2006; Nair and Tyagi 2021), one such scenarios include claiming a legitimate interest for data processing [art. 6 (1)(f) of the GDPR] and following simple procedural rules such as informing the data subject that their data will be monitored or collected. In well-argued cases of legitimate interest, obtaining consent is not required Data Protection Working Party (2014). In such instances, formal compliance can obscure significant breaches of privacy Cohen (2012). Understanding privacy through a strict data protection lens can therefore lead to a narrow approach that neglects the broader ethical considerations mentioned earlier in this section, such as control, autonomy and asymmetrical concentration of power.

Privacy as a mathematical concept

One final understanding of "privacy" is its interpretation within the realm of machine learning and data science. Repercussions for non-compliance with the legal requirements set forth in the GDPR—such as ensuring the lawfulness of data processing, purpose limitation, data minimization and transparency obligations—can be severe. As a result, the regulation has sometimes had the unintended consequence of hindering data-sharing across institutions and EU member states, even for research purposes (Peloquin et al. 2020). Yet the appeal of conducting research based off large amounts of data processing has not diminished. As a response, technological solutions have been developed to reduce privacy leaks and thus data processing becomes more

[4] See Graeme Laurie's challenging recasting of privacy in Laurie, G. Genetic Privacy: A Challenge to Medico-Legal Norms (Cambridge University Press 2002), Chapter 5 Laurie (2002).

[5] Townend, D. (2004) "Overriding Data Subjects' Rights in the Public Interest." In Beyleveld, D., Townend, D., Rouillé-Mirza, S., and Wright, J. (Eds) The Data Protection Directive and Medical Research Across Europe. Farnham: Ashgate Publishing Ltd, pp 89–101. Beyleveld et al. (2004).

[6] See, for example, Anita Allen, "Genetic Privacy: Emerging Concepts and Values", in Mark Rothstein (ed.) Genetic Secrets: Protecting Privacy and Confidentiality in the Genetic Era Rothstein (1997).

compliant with existing legislation. A few of these solutions include Multiparty Computation (MPC) (Yao 1982), Federated Learning (FL) (Dwork and Roth 2015), homomorphic encryption (Parmar et al. 2014), and synthetic data (van Daalen et al. 2024; Gregor et al. 2015; Denton et al. 2015; Eigenschink et al. 2023) to replace real data when training models. Finally, attempts to measure privacy in a concrete mathematical manner have been developed. In this context, privacy is approached mainly using the following three nonconflicting methods: (a) by setting and utilizing privacy thresholds, (b) by focusing on limiting data breaches, which we will discuss in "The misplaced focus on preventing data leaks and its consequences" section, and (c) through the use of so-called "privacy-preserving" technologies.

Setting and utilizing privacy thresholds

Whether it is hypothesis testing with p-values (Amrhein et al. 2019), creating a sufficient level of privacy with a privacy budget using schemes like k-anonymity or ϵ-differential privacy, or deciding if a model's predictions are accurate enough, statistical measures use specific thresholds as cut-off points to determine if the scenario passes a test. However, these thresholds are often set based on historic precedent rather than any truly objective reasoning. While these thresholds can be informative to a certain extent, the focus on these historic precedents causes researchers to be mostly concerned with simply passing this threshold, which has resulted in several important problems.

First, researchers are often not aware of how and why these thresholds were set (Mitchell et al. 2010). This is especially true for researchers who are not statisticians themselves. For example, most researchers working with quantitative data know about p-values, but probably would not be able to explain why the common threshold used to indicate statistical significance was set to 0.05. Yet, they will still accept or dismiss research based on this threshold. Second, this can also tempt said researchers to tweak their experiments in various ways to pass this test, which might mislead research findings (Head et al. 2015; Bouter et al. 2016; John et al. 2012). In the context of privacy this may mean that a researcher may mindlessly accept a privacy solution because a statistical test shows that with $p < 0.05$ no data is leaked. Lastly, while these measures are often effective at ranking different scenarios, it can be extremely difficult to meaningfully explain the practical differences between a 'good" and a "bad" score. Combined with the arbitrary nature of the threshold this makes it very difficult to explain why a solution is dismissed as 'bad", other than a simple "computer says no". While there is occasionally pushback against this blind reliance on arbitrary thresholds, but it is still a problem that can commonly be observed.

Privacy-preserving technologies

Approaching privacy as a technical problem has inevitably led to attempts to solve it technologically. In recent years, privacy-preserving or enhancing technologies have been developed to minimize the risk of data leakage and data reidentification. These solutions enable organizations to undertake multi-institution research projects (Scheenstra et al. 2022). They have notably been used to develop various commercial products, such as personalized advertisements, predictive text models for mobile phones, and recommender systems based on users' profiles and purchase history, but also to improve public services. Hospitals have used Federated Learning to combine patient data in a privacy-preserving manner to train machine learning models for disease diagnosis, which in turn improves healthcare offerings (Kairouz et al. 2021; Li et al. 2020; Yang et al. 2019, 2020; Vatsalan et al. 2017; Truong et al. 2021; Liu et al. 2020; Banabilah et al. 2022).

The progress made in developing these privacy enhancing tools is undeniable. However, current literature is primarily focused on the technical aspects of privacy and ignores other important issues. The focus on technical definitions often means the general public has difficulty understanding what is actually offered to them. Additionally, while mathematical measures of privacy may allow solutions to be ranked easily, this ranking is largely a theoretical exercise, and it may be difficult to determine the exact practical differences between two competing solutions. Lastly, reducing privacy to a purely mathematical problem gives it an "objective" veneer, which can be used to whitewash a project. Additionally, large tech organizations may push their preferred metric in an effort to shape the discussion on privacy in ways that benefit their business model. In the following sections we will elaborate these topics.

The misplaced focus on preventing data leaks and its consequences

Privacy is important in 4 aspects of the development and implementation of data-driven projects. These aspects are: (1) the training of the model, (2) the use of the model, (3) the technique or technology deployed, (4) the aim and application of the project. It is only when privacy is accounted for in all 4 of these aspects that such a project can be considered "privacy-preserving". These aspects tend to compete for researchers' attention. For instance, problems that arise during training (1) might include technical challenges such as data leaks or poisoning attacks by malicious parties. Similarly, issues falling under aspect (2) are primarily technical, such as concerns about model inversion. Additionally, there may be practical problems that may need to be solved

regarding the model use, for example where is the model hosted and how it accesses new data. However, the use of the resulting analyses (4) introduces more social and ethical considerations. For example, could the resulting model lead to discrimination, or reinforce existing biases (Verma 2019; Dutch 2021)? Answers to these concerns are often less straight-forward.

Likewise, ensuring the proper technique is applied to a specific problem (3) is relatively straightforward to establish and control. For example, if a project requires zero trust then it is trivial to establish that techniques that rely on a trusted third party are inappropriate. However, determining whether these techniques are implemented in an ethically responsible manner, and will not be abused in the future, for example after a change of leadership, is considerably more difficult. Ethically assessing an algorithm is challenging; but the manner in which it is used deserves attention. As a result, it is common to focus purely on the technical aspects, ignoring the ethical, legal and societal aspects (ELSA), of which privacy is part. Technical data leaks usually result in damage to the data-controller, revealing industry secrets and/or causing organizations to lose their commercial advantage, as well as lead to significant reputational damage and/or fines to the data-controller. This has led to zero-trust policies and complicates large cooperative projects, as sharing data is often deemed too risky or complex to execute safely.

However, these technical leaks do not necessarily lead to real harm for the data subjects. For example, the data leaked might not be directly identifiable without the use of additional information that is only available to the data-controller. While it may be technically possible to combine and cross-reference various external data sources to identify individual data subjects, this is unlikely to be feasible and the risk should be weighed against the effect and probability of successful attack. Additionally, the more sensitive the data, the harder cross-referencing becomes; sensitive data is not only better protected, but also harder to acquire (GDPR 2016).[7] This greatly limits the real harm done to the data-subjects. Finally, the step from a data leak to personal harm or damages of an individual often requires an active and conscious act of someone, which may not be the case. While it is established that there are multiple dimensions to privacy and we can distinguish between model-controller, model-user and data-subject privacy (Domingo-Ferrer 2007), this distinction is often either ignored or left implicit in technical papers.

To illustrate this, let us look at one of the most famous examples of data reidentification, the Netflix competition of 2006, in which researchers used a freely available public IMDB dataset to re-identify the records contained in the Netflix dataset (Narayanan and Shmatikov 2007). This was despite the fact that the Netflix dataset was considered not to contain any sensitive identifiable data. Additionally, had the IMDB dataset not existed, it would have been possible to create a reference database via phishing attempts, such as using seemingly harmless quizzes on the internet (Parsons 2020), or by abusing data-leaks. In this instance, the risk of reidentification was very high, thanks to publicly available reference datasets, as is clear in hindsight. However, it is important to note that the real harm to data-subjects from this leak was minimal or non-existent as it contained little to no new information compared to the already public IMDB database, and knowing who makes which comments on which movie has very limited risk of damage to that reviewer anyway.

In contrast, a medical dataset does not have an easily accessible public database that could be used to re-identify individuals. It's also much less likely one would be able to successfully employ phishing websites to gain access to a dataset to cross-reference. No matter how personal the medical data may be, they are, in practice, extremely unlikely to lead to patient reidentification and subsequent harm and damages.

Another example of how minor leaks are often presented as major problems can be seen in the following paper. Slokom et al. claim synthetic data is not privacy preserving because they devised an attack which revealed sensitive data (Slokom et al. 2022). However, they overlook key contextual aspects. Most notably, that the leaked sensitive information is limited and often does not meaningfully improve upon baseline prediction of sensitive attributes; in some scenarios, it even performs worse. Even when successful, the attack only slightly exceeds random performance, achieving about 60% accuracy on a sensitive binary attribute. This high level of uncertainty means that such an attack cannot be deemed a serious privacy threat in this scenario. If attacks with such high levels of uncertainty are deemed major breaches of privacy, publishing any analysis would be impossible, as even basic analyses reveal information (Dwork and Naor 2010; Shokri et al. 2017). While the attack vector is relevant and a potential concern in specific scenarios, Slokom et al. do not address its practical limitations.

[7] See GDPR, Article 9. Processing of special categories of personal data. See GDPR, recital 51. further clarifying the protection of sensitive personal data, lifting the restriction on processing in cases were explicit consent is provided by the data subject, or: '[...] for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller.'. See GDPR, recital 52. which further derogates the processing prohibition of special data for the public interest: 'Such a derogation may be made for health purposes, including public health and the management of health-care services, especially in order to ensure the quality and cost-effectiveness of the procedures used for settling claims for benefits and services in the health insurance system, or for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes'.

These examples are illustrative of the broader problem within privacy-preserving literature caused by the focus on technical research questions and failure to consider contextual practical implications and limitations. These studies usually only consider the worst-case scenario where the attacker has practically unlimited resources, or it is assumed that a reference dataset exists to identify individuals. Each record is presupposed to always be unique enough to be identifiable, even if best practices show that such outliers are uninformative and should be removed from your dataset during preprocessing, thus providing a minimum level of privacy via k-anonymity (Sweeney 2002). Additionally, the impact of an attack is based on the amount of data revealed, not on the contextual harm it can do to the data subject. For example, if an attacker attacks two image recognition models, one trained on faces, one trained on MRI images, and in both cases retrieves an image, and nothing more, this is treated as an equivalent leak with equivalent damage in both scenarios. This ignores the fact that one image may be easier to identify but contains relatively limited sensitive information, while the other image contains a lot of sensitive information but may be more difficult to identify. The potential for harm towards the data subject is vastly different in the two scenarios.

In summary, the real impact on the data subjects in the given context is rarely considered, and neither are their preferences. This leads to a focus on secrecy over privacy. Which in turn also leads to researchers ignoring other important, and often connected, ethical aspects such as the risk of biases harming the data subjects. It also leads to researchers overlooking alternative solutions, such as legal solutions. Lastly, because of this focus on secrecy and the relatively short-term goal of protecting the data-controllers' interests, researchers and engineers often ignore the long-term implications of a project for the data-subjects.

The focus should instead lie on how the different stakeholders involved are affected, with a strong focus on the data subjects, by potential leaks, as well as how "normal" use of the data would affect them. Additionally, researchers and policy makers should rely less on generic definitions of the risks involved, instead the risks should be estimated on a project-by-project basis. It is important to note that this is not a trivial problem that can easily be automated. Discussing the potential risks and solutions for a particular project is complicated, and will require considerable effort each time. Lastly researchers should acknowledge and actively push for alternative solutions. They should not be allowing privacy preserving technologies to be used to whitewash questionable projects. We will further discuss this practice of whitewashing in the next section.

The role of big tech in defining what is, should or should not be private

In order to create and maintain a situation where they benefit, Big Tech companies have successfully pushed their own agenda, by influencing our understanding of privacy. This paper has introduced 4 aspects of AI privacy in "The misplaced focus on preventing data leaks and its consequences" section: (1) the training of the model, (2) the use of the model, (3) the technique or technology deployed, and (4) the aim and application of the project. Big Tech companies, however, almost entirely focus their privacy preserving efforts on the first 3 aspects. Indeed, those are the phases that allow them to focus on "objective" technical mathematical problems, for which easily demonstrable solutions can be found. Aspect 4, however, is neglected, as it is more likely to raise questions of a more ethical nature that cannot be addressed straightforwardly. While there are legal frameworks that need to be followed, an application being legal is no guarantee that it is ethical and Big Tech prefers to avoid questions on this topic as much as possible. This attitude is in alignment with the general practices highlighted in "The misplaced focus on preventing data leaks and its consequences" section. This section goes further with that observation, arguing that Big Tech is directly involved in maintaining a reductive understanding of privacy as an issue that can be fixed technologically. This allows them to circumvent deeper questions about their business model. Rather than to rethink the way that they collect and process data, they instead (a) created services advertised as "free" but that users in fact pay for by giving away personal data, (b) hide under the promise of "privacy-preserving" techniques, (c) use these techniques to justify targeted advertising and (d) eliminate dissent and competition through brain draining and lobbying.

The expectation of "free" online services

Meta, Amazon, Alphabet: all offer services—such as online shopping apps, search engines, and social media apps—that appear free and are so convenient that their use has become the norm. That those services are not 'free' for use but paid by the trade-in of personal information is not often clear to users, although awareness has been rising. In the EU, lawmakers have deployed efforts to protect the rights of individuals to their personal data with the General Data Protection Regulation. While the GDPR has, at times, constituted a minor setback or annoyance for these companies, it did not result in them rethinking their incredibly profitable business model. Instead, they opted for privacy-washing and complying in ways that could be

qualified as questionable. For example, back in October 2023, Meta announced that it would give its European users the choice to use their platform without being shown relevant ads Facebook and Instagram (2023) and Roush (2023), if they agreed to pay a monthly premium of 9,99€ for the web versions, and 12,99€ for the apps. This was a direct response to the European Data Protection Board (EDPB) warning Meta that they could not force their users to consent to their data being extracted by making them leave the platform if they wished to preserve their data Facebook and Instagram (2023). Yet, the effect was the same: Meta users were greeted with a long wall of text and the choice to tick either one box or the other, deciding whether they wanted to keep using the platform for free, or pay a subscription. This tactic, qualified "pay or okay" Roush (2023), was harshly criticized by the EDPB. It deemed that consent obtained under circumstances where data users are forced to choose between two options—either to allow their data to be used for targeted avertising, or refuse for their data to be processed and thus must pay a fee to keep using the service—without being offered any other alternative cannot be equated to "freely given" consent European Data Protection Board (2024).

Technological privacy preservation

Aside from offering convenient, attractive and "free" services to their users, Big Tech companies would also suggest that their processing of your personal data is entirely safe and private. While the development of privacy preserving solutions has by no means been limited to just Big Tech, they are heavily involved in the development and promotion of privacy-preserving technologies, in order to sell their point of view. Given that their business model relies on the use of vast amounts of personal data, they have a clear stake in the development of such technologies, as well as their perception by legislative authorities and the public. One of such technologies is federated learning, a term coined by Google, which heavily relied on this technique to develop personalized text prediction in the Gboard, the virtual keyboard with auto-correct and text prediction functionality (Banabilah et al. 2022). Many promotional materials can be found to sing the praises of this learning method, all the while obscuring the fact that the company, although indirectly so, is accessing the contents of our emails, messages, and all other text input processed using Google's software. Is personalized text prediction worth such an invasion of privacy? The default on our machines would suggest that the anwser is yes.

Given that personal data have become highly commodified and profitable goods, companies naturally try to accumulate as much of it as possible and allow their data scientists to run numerous analyses on it. Hiding behind the promise of privacy-preservation enables this data behaviour: after all, if the data used is anonymous (GDPR 2016),[8] why should consumers or legislators be worried[9]? However, this anonymity claim is flawed: while privacy preserving techniques (PPTs) can guarantee a certain layer of security for a single analysis, one could run multiple queries concurrently in order to reveal private information. Just because a technology makes data processing safer, does not mean safety is guaranteed. On the contrary, being truly concerned with privacy would mean implementing queries that are predefined and limited in scope for specific high-level functionalities: for instance, building a model. Such an approach would align more closely with EU legislation and its guidelines on ethical and trustworthy AI Proposal for a Regulation of the European (2021), which promote data minimization, human oversight, and prevention of harm. Big tech companies' tendency to greedily accumulate data directly contradicts these principles.

Rather than fundamentally rethink their unethical business practices, Big Tech has instead focused on advertising technological fixes to the issue of privacy, using objective numbers to solve an issue that is, in fact, societal. As such, the mathematical conception of privacy entirely benefits their agenda. It is much easier to develop new technology to remain under a set privacy threshold rather than to think more deeply about the ethical and safety concerns associated with their processing of personal data. Furthermore, this focus on secrecy and preventing data leaks presents a clear commercial advantage. To ensure that their vision is widely disseminated, Big Tech has funneled a significant amount of energy and funds into research that aligns with their agenda and promotes the use and effectiveness of PPTs (West 2019; Papaevangelou 2024; Zhao et al. 2019; Ochigame 2019). Consequently, views commonly held in privacy-preserving literature tend to align with the interests of these companies and organizations, a point this paper will elaborate on further in Sect. 5.4.

Privacy preservation and audience targeting

While PPTs afford a higher level of safety in gathering and processing personal data, their use (aspect 4) can result in models which would constitute an intrusion in one's private

[8] See GDPR, Article 2(1), respecting Article 4(1), and recital 26. Truly anonymous data, as explained in recital 26, does not fall within the material scope of the GDPR [Art.2(1)].

[9] Personal data protection legislation largely does not bite on anonymous—truly de-identified—data; research ethics has always seen consent and anonymisation as the gold standards of protection of the individual. This leaves open other dignity breaches in relation to de-identified data.

life. A blatant example of this is targeted advertising (Srivastava et al. 2023; Antonio et al. 2022; Radesky et al. 2020; Dogruel 2019), which is usually activated by default, and is not easily disabled. Furthermore, allowing your data to be collected is often the condition to access or use most services provided online. Companies claim that they are able to perform such a service while preserving user privacy (Reznichenko and Francis 2014), by using state of the art PPTs. However, it can easily be argued that targeted ads themselves present a breach of privacy. They reflect a user's browsing history, past purchases, etc. Who hasn't had the experience of searching something up, only for ads for that very item being plastered all over the next website we visited? Other examples below are empirical evidence of these practices. Expectant mothers are likely to engage with posts and hashtags related to pregnancy, leading to the ads being showed to them being very baby-centric. This could lead to potentially inadvertently revealing pregnancies if one uses their personal device in view of someone else, but it can also lead to personal harm. In 2018, a woman reported that Facebook would continuously show her ads for baby products while she was grieving the loss of her unborn child (Brockell 2018). The company had successfully detected her pregnancy, but not that it had resulted in a stillbirth. With Roe v. Wade being overturned by the Supreme Court in the US in 2022, additional concerns are rising regarding the right to privacy about one's pregnancy status (Kelly and Habib 2023). The popularity of period-tracking apps is leading to fears that such data might end in the hands of prosecutors trying to enforce the criminalization of abortion. A growing body of literature discusses the many risks associated with Big Tech's access to information about female reproductive health (Mehrnezhad and Almeida 2023; Shipp and Blasco 2020; De and Imine 2020; Mehrnezhad et al. 2022; Healy 2021). The Cambridge-analytica scandal (Hinds et al. 2020) has also shown that ads can be used to successfully influence democratic elections by targeting the individuals most likely to be swayed (Albright 2016), invading user privacy and using the information collected in the process to manipulate them.

Biased research and lobbying

A few powerful companies hide behind the label "Big Tech": arguably the most important are Meta, Alphabet, Microsoft, Amazon and Apple (Verdegem 2024; Affeldt and Kesler 2021). Together, they form an oligarchy, dominating the IT market worldwide. But this influence does not stop there. Recent studies have shown that Big Tech financially backs a large proportion of academics in the field of AI (Project 2023; Abdalla and Abdalla 2021; Menn and Nix 2023; Ochigame 2019). In doing so, they ensure that research aligns with their interest. An example of this is how research is currently being conducted on the topic of "fairly" rewarding data-contributors (Lyu et al. 2020). This would involve rewarding them proportionally to the value of the data they contributed, in a privacy-preserving manner, to a project. Such a system would entirely benefit these oligarchies, as they are the legal custodians of the largest amount of data, and would be completely irrelevant as far as the individual data subjects are concerned as individual subjects will never provide enough data to receive meaningful rewards under such a scheme. It would disproportionately hurt marginalized communities, which already benefit the least from improvements in AI, while suffering the most from its side-effects (Feuerriegel et al. 2020; Arora et al. 2023; Xenophobic 2021). In funding such research, oligopolies gain credibility and build trust by claiming to implement "fair" and "privacy preserving" AI, even though they are clearly serving their own interest, even at the cost of harming others. Consequently, government bodies, who rely on academia for guidance on how to regulate AI, are likely to be influenced by this agenda too Abdalla and Abdalla (2021).

When they are not actively draining brains from academia or from promising startups (Goldman 2024), Big Tech companies actively challenge what little is left of their competition through lobbying. It is interesting to consider the difference in the western public perception of TikTok and Meta, two companies that offer fundamentally similar services with the same modus operandi. That TikTok is facing being banned in the US whereas Meta is allowed to thrive is therefore puzzling, until one becomes aware of the role that the latter played in that situation. It was indeed reported that Meta had spent millions on lobbying for such a ban (Shaw 2024), insisting that the Chinese-owned TikTok represented not only a threat to the privacy of its users, but also to national security, helped in the process by US lawmakers' long history of sinophobia, which the COVID-19 pandemic only worsened (Siu and Chun 2020). In the EU, the app was banned from the official devices of government personnel for similar reasons (Maheshwari and Holpuch 2024), while the use of US-owned social media platforms remains allowed. This raises the suspicion that Meta rid itself of its biggest competitor on the market, by presenting itself as less of a threat than its Chinese counterpart (Lorenz and Harwell 2022).

Discussion

The way the issue of privacy is understood, approached and "solved" within the field of machine learning is currently flawed. Rather than serving the interests of the people whose privacy is at risk, arguably it aligns with those of Big Tech companies. They are able to profit from constantly exploiting our personal sphere, extracting as much data as possible,

and selling this data to third parties that use this information to either sell us goods and services or, more worryingly, influence our beliefs and behaviour. This situation needs to change. This section goes over potential solutions to some of the problems highlighted above.

First, the understanding of privacy within ML needs to be improved. Too often, invasion of privacy is conflated with data breach. A great deal of effort is deployed in protecting data against such breaches, even when they do not harm data subjects in any tangible way. On the other hand, anonymization techniques are often considered amply sufficient to preserve privacy, despite them not always preventing group or societal harms. As of now, there is a disproportionate focus on data leaks, which is more likely to benefit data-controllers, their trade secrets, and their position of power on the (AI) market, rather than data subjects themselves. Placing further importance on the likelihood of harm and the preservation of data subjects' control and autonomy is needed to complete the existing vision of privacy preservation in ML. Additionally, acknowledging the fact that privacy itself is a nuanced concept, which cannot entirely be solved through technological means, would be a more honest approach. Acknowledging this limitation would help counter the false sense of security that one's data can be kept 100% private even when harvested by Big Tech: a false sense of security that these companies will happily use.

Under the current definition of privacy within machine learning, a data leak comprising information about a random unknown patient's blood type that could very difficultly be linked back to them would be considered a privacy issue, whereas a specific individual woman being labelled as "pregnant" even when she did not voluntarily share that information about herself, might not be considered a privacy issue within the understanding of "privacy" held in machine learning (even though, of course, at law this would be a matter of privacy). This situation is nonsensical and is not beneficial to data subjects.

Secondly, when it comes to European regulation, while steps have been taken in the right direction, they are insufficient to generate change. For example, although consent is often presented as a central mechanism under the GDPR for legitimizing personal data processing, there is much debate around what constitutes freely obtained consent. Indeed, online consent is often obtained through pop-up, sometimes pre-ticked consent boxes to websites and apps that collect personal information (Jablonowska and Michatowicz 2020). Furthermore, as discussed earlier in this article, "pay or okay" subscription models have been developed by companies such as Meta to give users the illusion of agency—although this practice was deemed incompatible with providing a genuine and free choice for users by the EDPB (European Data Protection Board 2024). Still, some controllers may attempt to invoke legitimate interest to do away with consent requirements for scraping personal data for AI model training. Finally, in practice, users often have limited visibility over whether their information was used to train AI models, how to opt out of this process (Grinkevičius 2024; Mauran 2024; Fitzgerald 2024; Hoe Meta Gegevens Gebruikt 2024; Shah 2024), and it is thus difficult for them to become aware of malpractice and enforce their data protection rights. The vision of the EU is short-sighted and fails to address the deeper issues caused by the very business model of these companies. This failure to address the deeper issues, combined with lacking enforcement, has allowed these companies to cleverly avoid making any real change to their unethical practices. Focusing on user consent as a legal basis for data processing does not seem to be the answer when many data subjects have grown accustomed to using free apps and services in exchange for data or feel like they have lost complete control over their personal information. Additional steps in the regulation of AI have been taken, comprising the Digital Market Act, the Digital Services Act and the AI Act, however they are likely to have a similarly limited effect, although perhaps the observed shift (Kusche 2024), in the AI act, to focusing on possible harm and risks to humans, which is one of the 7 guiding principles of trustworthy AI according to EU, could lead to some more promising results. As discussed, data science projects often focus on the training and implementation stages of their models, while neglecting their possible real-life consequences. Becoming more conscious of these long-term implications would make it easier to identify potential risks further down the line. This would not only be an improvement on a project-by-project basis, but also create a healthier culture, where data scientists look beyond their direct responsibilities toward potential future problems. Furthermore, a clearer understanding of the real impact of data leaks onto data subjects, rather than data-controllers, would be beneficial. Current research focusses on theoretical impacts, rather than realistic impact. Identifiable data may be of extremely limited value to attackers, and reidentification itself might not necessarily lead to harm to the data subjects.

Thirdly, while shifting to a risk or harm-based approach rather than a consent-based or technologically private approach would be an improvement, another step in the right direction would be to give more importance to the purpose of data processing. While exceptions to certain GDPR obligations are in place for data processing that is undertaken for research or common good purposes, it is still frequently the cause of much confusion and frustration for many who perform such work. Research has suffered time and time again from attempting to comply with rules for which they were not the primary target. At the same time, EU regulation has failed to significantly hinder unethical and invasive large scale data collection and manipulation by Big Tech companies. Purposes such as targeted advertising and online

profiling should be more strictly regulated, especially when they have proven to be harmful. Without a bolder regulatory approach, EU data subjects will continue to see their data being extracted from them and sold to third parties that may not serve their interests. Furthermore, as this is often done without the subjects' knowledge, existing legal recourse becomes virtually inapplicable: without information about who is accessing or processing one's data, and for what purpose, how would one request that they cease to do so? Lack of transparency about data handling has led to the inability for people to effectively exercise their rights to control their personal information.

We acknowledge that these are not easy solutions to implement, and will encounter significant opposition, especially from for-profit organizations who do not want to see the status quo change. While it is a bitter pill to swallow, privacy is a complex, context-dependent topic. It is impossible to fully solve within a simple, scalable, generic solution. Accepting the limitations of privacy preserving technologies, and pushing back against those who claim to can solve complex ethical questions with a silver bullet, is the only way forward.

Conclusion

The machine-learning field has attempted to reduce the complex notion of "privacy" to a purely mathematical, technically solvable problem. This has led to several issues: the creation of privacy thresholds that hold little meaning, the claim that certain technologies will guarantee that personal data will remain private, and an overall focus on the development of such privacy-preserving techniques while completely neglecting longer-term effects of large-scale data-intensive projects. Treating privacy like a simple, solvable issue has allowed the Big Tech oligarchy to continue profiting from harvesting data from millions of users without having to pay sufficient attention to the potential harm caused to data subjects by their activities, justifying their behaviour by advertising their technologically robust privacy-preservation techniques. The steps taken in data protection law so far have not had a significant impact and led to new issues for actors that process data for purposes should instead have been facilitated, such as research. We acknowledge that our critiques clash with the wish for scalable, generic solutions and will invite pushback. Additionally, we acknowledge that our suggested solutions have their own practical limitations. A complex, context-dependent, topic such as privacy simply will not have easy answers."

It is our hope that this article will spark new discussions surrounding the role of Big Tech, and researchers themselves, in defining privacy not only within the machine-learning field, but also in policy-making and public discourse. These could in turn help reshape the privacy protection framework so that it focuses on those who truly should be protected: the data subjects.

Acknowledgements This research received funding from the Netherlands Organization for Scientific Research (NWO): Coronary ARtery disease: Risk estimations and Interventions for prevention and EaRly detection (CARRIER): project nr. 628.011.212.

References

Affeldt, P., & Kesler, R. (2021). Big tech acquisitions–Towards empirical evidence †. *Journal of European Competition Law and Practice, 12*(6), 471–478. https://doi.org/10.1093/jeclap/lpab025. Accessed 2024-08-16.

Bak, M., Madai, V. I., Celi, L. A., Kaissis, G. A., Cornet, R., Maris, M., Rueckert, D., Buyx, A., & McLennan, S. (2024). Federated learning is not a cure-all for data ethics. *Nature Machine Intelligence, 6*(4), 370–372. https://doi.org/10.1038/s42256-024-00813-x. Publisher: Nature Publishing Group. Accessed 2024-07-09.

Bouter, L. M., Tijdink, J., Axelsen, N., Martinson, B. C., & Ter Riet, G. (2016). Ranking major and minor research misbehaviors: Results from a survey among participants of four World Conferences on Research Integrity. *Research Integrity and Peer Review, 1*(1), 17. https://doi.org/10.1186/s41073-016-0024-5. Accessed 2024-07-08.

Cohen, J. E. (2012). What privacy is for. *Harvard Law Review, 126*, 1904.

Daalen, F., Ippel, L., Dekker, A., & Bermejo, I. (2024). VertiBayes: Learning Bayesian network parameters from vertically partitioned data with missing values. *Complex and Intelligent Systems*. https://doi.org/10.1007/s40747-024-01424-0. Accessed 2024-05-21.

De, S. J., & Imine, A. (2020). Consent for targeted advertising: The case of Facebook. *AI & SOCIETY, 35*(4), 1055–1064. https://doi.org/10.1007/s00146-020-00981-5. Accessed 2024-07-09.

Eigenschink, P., Reutterer, T., Vamosi, S., Vamosi, R., Sun, C., & Kalcher, K. (2023). Deep generative models for synthetic data: A survey. *IEEE Access, 11*, 47304–47320. https://doi.org/10.1109/ACCESS.2023.3275134. Conference Name: IEEE Access. Accessed 2024-07-08.

Feuerriegel, S., Dolata, M., & Schwabe, G. (2020). Fair AI. *Business and Information Systems Engineering, 62*(4), 379–384. https://doi.org/10.1007/s12599-020-00650-3. Accessed 2024-07-09.

Fuster, G. G. (2014). *The emergence of personal data protection as a fundamental right of the EU* (Vol. 16). Springer.

Fuster, G. G., & Gellert, R. (2012). The fundamental right of data protection in the european union: In search of an uncharted right.

International Review of Law, Computers and Technology, 26(1), 73–82.

Head, M. L., Holman, L., Lanfear, R., Kahn, A. T., & Jennions, M. D. (2015). The extent and consequences of P-hacking in science. *PLOS Biology, 13*(3), 1002106. https://doi.org/10.1371/journal.pbio.1002106. Publisher: Public Library of Science. Accessed 2023-08-01.

Healy, R. L. (2021). Zuckerberg, get out of my uterus! An examination of fertility apps, data-sharing and remaking the female body as a digitalized reproductive subject. *Journal of Gender Studies, 30*(4), 406–416. https://doi.org/10.1080/09589236.2020.1845628. Accessed 2023-09-12.

John, L. K., Loewenstein, G., & Prelec, D. (2012). Measuring the prevalence of questionable research practices with incentives for truth telling. *Psychological Science, 23*(5), 524–532. https://doi.org/10.1177/0956797611430953. Publisher: SAGE Publications Inc., Accessed 2024-08-14.

Laurie, G. (2002). *Genetic privacy: A challenge to medico-legal norms*. Cambridge University Press.

Macnish, K. (2018). Government surveillance and why defining privacy matters in a post-snowden world. *Journal of Applied Philosophy, 35*(2), 417–432.

Mitchell, M. S., Yu, M. C., & Whiteside, T. L. (2010). The tyranny of statistics in medicine: A critique of unthinking adherence to an arbitrary p value. *Cancer Immunology, Immunotherapy, 59*(8), 1137–1140. https://doi.org/10.1007/s00262-010-0859-4. Accessed 2023-07-13.

Moore, A. D. (2008). Defining privacy. *Journal of Social Philosophy, 39*(3), 411–428.

Nair, M. M., & Tyagi, A. K. (2021). Privacy: History, statistics, policy, laws, preservation and threat analysis. *Journal of Information Assurance & Security, 16*(1), 1.

Nissenbaum, H. (2009). Privacy in context: Technology, policy, and the integrity of social life. *Stanford University Press*. https://doi.org/10.1515/9780804772891. Accessed 2024-07-08.

Parmar, P. V., Padhar, S. B., Patel, S. N., Bhatt, N. I., & Jhaveri, R. H. (2014). Survey of various homomorphic encryption algorithms and schemes. *International Journal of Computer Applications, 91*(8), 26–32. https://doi.org/10.5120/15902-5081

Peloquin, D., DiMaio, M., Bierer, B., & Barnes, M. (2020). Disruptive and avoidable: GDPR challenges to secondary research uses of data. *European Journal of Human Genetics, 28*(6), 697–705. https://doi.org/10.1038/s41431-020-0596-x. Publisher: Nature Publishing Group. Accessed 2024-07-08.

Radesky, J., Chassiakos, Y. L. R., Ameenuddin, N., & Navsaria, D. (2020). Council on communication and media: Digital advertising to children. *Pediatrics, 146*(1), 20201681. https://doi.org/10.1542/peds.2020-1681. Accessed 2024-07-09.

Rothstein, M. A. (1997). *Genetic secrets: Protecting privacy and confidentiality in the genetic era*. Yale University Press.

Shipp, L., & Blasco, J. (2020). How private is your period?: A systematic analysis of menstrual app privacy policies. *Proceedings on Privacy Enhancing Technologies, 2020*(4), 491–510. https://doi.org/10.2478/popets-2020-0083. Accessed 2023-09-13.

Solove, D. J. (2002). Conceptualizing privacy. *California Law Review, 90*(4), 1087–1155. https://doi.org/10.2307/3481326. Publisher: California Law Review Inc, Accessed 2024-07-17.

Sweeney, L. (2002). k-Anonymity: A model for protecting privacy. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 10*(05), 557–570. https://doi.org/10.1142/S0218488502001648. Publisher: World Scientific Publishing Co., Accessed 2022-08-22.

Verdegem, P. (2024). Dismantling ai capitalism: The commons as an alternative to the power concentration of big tech. *AI & society, 39*(2), 727–737.

Warren, S., & Brandeis, L. (1890). The right to privacy. *Harvard Law Review, 4*(5), 193–220.

West, S. M. (2019). Data capitalism: Redefining the logics of surveillance and privacy. *Business and Society, 58*(1), 20–41. https://doi.org/10.1177/0007650317718185. Publisher: SAGE Publications Inc., Accessed 2024-07-08.

Zhao, R., Li, X., Liang, Z., & Li, D. (2019). Development strategy and collaboration preference in S &T of enterprises based on funded papers: A case study of Google. *Scientometrics, 121*(1), 323–347. https://doi.org/10.1007/s11192-019-03182-0. Accessed 2024-07-08.

Abdalla, M., & Abdalla, M. (2021) The grey hoodie project: Big Tobacco, Big Tech, and the Threat on Academic Integrity. In *Proceedings of the 2021 AAAI/ACM conference on AI, ethics, and society. AIES'21* (pp. 287–297). Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3461702.3462563. Accessed 2024-07-09.

Ahmed, I (2023). Aclu V. Clearview AI, Inc. *DePaul Journal of Art, Technology & Intellectual Property Law*, **33**(1), 4

Albright, J. (2016). *How Trump's campaign used the new data-industrial complex to win the election*. https://blogs.lse.ac.uk/usappblog/2016/11/26/how-trumps-campaign-used-the-new-data-industrial-complex-to-win-the-election/. Accessed 2024-07-17.

Amrhein, V., Greenland, S., & McShane, B. (2019). Scientists rise up against statistical significance. *Nature,567*(7748), 305–307. https://doi.org/10.1038/d41586-019-00857-9. Bandiera_abtest: a Cg_type: Comment Publisher: Nature Publishing Group Subject_term: Research data, Research management. Accessed 2024-07-08

Antonio, B. A., Cruz, K. L. D., Jimenez, A. I., & Pantoja, E. (2022). Invasion or personalization: An overview on user attitudes towards the privacy issues in targeted advertising in NCR and its effect in consumer purchase behavior. *Journal of Business and Management Studies*, *4*(2), 38–47 https://doi.org/10.32996/jbms.2022.4.2.4. Number: 2. Accessed 2024-07-09.

Arora, A., Barrett, M., Lee, E., Oborn, E., & Prince, K. (2023). Risk and the future of AI: Algorithmic bias, data colonialism, and marginalization. *Information and Organization,33*(3), Article 100478. https://doi.org/10.1016/j.infoandorg.2023.100478. Accessed 2024-07-09.

Banabilah, S., Aloqaily, M., Alsayed, E., Malik, N., & Jararweh, Y. (2022). Federated learning review: Fundamentals, enabling technologies, and future applications. *Information Processing and Management,59*(6), Article 103061. https://doi.org/10.1016/j.ipm.2022.103061. Accessed 2023-07-31.

Beyleveld, D., Townend, D., Rouillé-Mirza, S., & Wright, J (2004) *The data protection directive and medical research across Europe*.

Brockell, G. (2018). Perspective | Dear tech companies, I don't want to see pregnancy ads after my child was stillborn. Washington Post. Accessed 2023-08-01.

Bundesverfassungsgericht, S. (1983) *Bundesverfassungsgericht-Decisions–Decision on the constitutionality of the 1983 Census Act*. Archive Location: de Publisher: Bundesverfassungsgericht. https://www.bundesverfassungsgericht.de/SharedDocs/Entscheidungen/EN/1983/12/rs19831215_1bvr020983en.html. Accessed 2024-07-17.

California Code Civil Code-CIV DIVISION 3-OBLIGATIONS PART 4-Obligations arising from particular transactions title 1.81.5-California Consumer Privacy Act of 2018 Section 1798.192. Universal Citation: CA Civ Code §1798.192 (2023).

Cromack, J (2021) *Why Amazon's GDPR fine really matters: Consent in marketing*. https://www.dataguard.co.uk/blog/why-amazons-gdpr-fine-really-matters-its-time-to-be-open-with-consent-for-marketing. Accessed 2024-08-14.

Data Protection Working Party: Article 29: Opinion 06/2014 on the notion of legitimate interests of the data controller under Article 7 of Directive 95/46/EC (2014). https://ec.europa.eu/justice/article-29/press-material/public-consultation/notion-legitimate-interests/files/20141126_overview_relating_to_consultation_on_opinion_legitimate_interest_.pdf

De Hert, P., & Gutwirth, S. (2006). Privacy, data protection and law enforcement. Opacity of the individual and transparency of power. Privacy and the criminal law 61–104

Dean, J. (2003). Review of regulating intimacy: A new legal paradigm. *Journal of Law and Society,30*(3), 453–458. Publisher: [Cardiff University, Wiley]. Accessed 2024-07-17.

Denton, E. L., Chintala, S., Szlam, A., & Fergus, R. (2015). Deep generative image models using a laplacian pyramid of adversarial networks. In *Advances in neural information processing systems* (Vol. 28). Curran Associates, Inc. https://proceedings.neurips.cc/paper/2015/hash/aa169b49b583a2b5af89203c2b78c67c-Abstract.html. Accessed 2024-07-08.

Dogruel, L. (2019). Too much information!? Examining the impact of different levels of transparency on consumers' evaluations of targeted advertising. *Communication Research Reports,36*(5), 383–392. https://doi.org/10.1080/08824096.2019.1684253. Publisher: Routledge_eprint: Accessed 2024-07-09.

Domingo-Ferrer, J (2007) A three-dimensional conceptual framework for database privacy. In *Secure data management: 4th VLDB workshop, SDM 2007, Vienna, Austria, September 23–24, 2007. Proceedings 4* (pp. 193–202). Springer.

Dutch Childcare Benefit Scandal an Urgent Wake-Up Call to Ban Racist Algorithms. (2021). https://www.amnesty.org/en/latest/news/2021/10/xenophobic-machines-dutch-child-benefit-scandal/. Accessed 2024-07-08.

Dwork, C., McSherry, F., Nissim, K., spsampsps Smith, A (2006) Calibrating noise to sensitivity in private data analysis. In S. Halevi, T. Rabin (Eds.) Theory of cryptography. Lecture notes in computer science (pp. 265–284). Springer. https://doi.org/10.1007/11681878_14.

Dwork, C., & Naor, M. (2010). On the difficulties of disclosure prevention in statistical databases or the case for differential privacy. *Journal of Privacy and Confidentiality*.https://doi.org/10.29012/jpc.v2i1.585. Number: 1. Accessed 2024-07-09.

Dwork, C., & Roth, A. (2015). The algorithmic foundations of differential privacy. *Foundations and Trends®in Theoretical Computer Science*, 9(3–4), 211–407 https://doi.org/10.1561/0400000042. Accessed 2021-06-15.

EDPB: 1.2 billion euro fine for Facebook as a result of EDPB binding decision | European Data Protection Board (2023). https://www.edpb.europa.eu/news/news/2023/12-billion-euro-fine-facebook-result-edpb-binding-decision_en. Accessed 2024-07-17.

EU General Data Protection Regulation (GDPR): 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), OJ 2016 L 119/1. (2016). https://gdpr-info.eu/

Euronews: Meta hit with €265 million fine by Irish regulators for breaking Europe's data protection law | Euronews (2022). https://www.euronews.com/next/2022/11/28/meta-hit-with-265-million-fine-by-irish-regulators-for-breaking-europes-data-protection-la. Accessed 2024-08-14.

European Court of Human Rights: Guide on Article 8 of the European Convention on Human Rights Council of Europe (2022). https://ks.echr.coe.int/documents/d/echr-ks/guide_art_8_eng

European Data Protection Board: Opinion 08/2024 on Valid Consent in the Context of Consent or Pay Models Implemented by Large Online Platforms. (2024). https://www.edpb.europa.eu/system/files/2024-04/edpb_opinion_202408_consentorpay_en.pdf

European Union: Charter of Fundamental Rights. (2007). https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:C:2007:303:FULL.

Facebook and Instagram to Offer Subscription for No Ads in Europe. (2023). https://about.fb.com/news/2023/10/facebook-and-instagram-to-offer-subscription-for-no-ads-in-europe/. Accessed 2024-07-09.

Fitzgerald, C. (2024). *Facebook will soon use your photos, posts and other info to train its AI*. You can opt out (but it's complicated). https://www.thejournal.ie/facebook-data-ai-6391876-May2024/. Accessed 2024-07-09.

Foundation, M. (2023). 'Privacy nightmare on wheels': Every car brand reviewed by Mozilla—Including Ford, Volkswagen and Toyota—Flunks Privacy Test. https://foundation.mozilla.org/en/blog/privacy-nightmare-on-wheels-every-car-brand-reviewed-by-mozilla-including-ford-volkswagen-and-toyota-flunks-privacy-test/. Accessed 2024-07-08.

GDPR Enforcement Tracker-List of GDPR fines. https://www.enforcementtracker.com. Accessed 2024-07-17.

Goldman, S. (2024). *The 'Meta AI mafia' brain drain continues with 3 more major departures*. https://fortune.com/2024/04/02/mark-zuckerberg-ai-jobs-meta-brain-drain-erik-meijer/. Accessed 2024-07-09.

Gregor, K., Danihelka, I., Graves, A., Rezende, D., & Wierstra, D (2015) DRAW: A recurrent neural network for image generation. In *Proceedings of the 32nd international conference on machine learning* (pp. 1462–1471). PMLR. ISSN: 1938-7228. https://proceedings.mlr.press/v37/gregor15.html. Accessed 2024-07-08.

Grinkevičius, P. (2024). *Meta uses your data to train AI, and it doesn't want you to opt out*. https://cybernews.com/tech/meta-trains-ai-on-your-data/. Accessed 2024-07-09.

Hagen, L. (2021). *Privacy preserving machine learning: Maintaining confidentiality and preserving trust*. https://www.microsoft.com/en-us/research/blog/privacy-preserving-machine-learning-maintaining-confidentiality-and-preserving-trust/. Accessed 2023-07-25.

Hinds, J., Williams, E. J., & Joinson, A. N. (2020). "It wouldn't happen to me": Privacy concerns and perspectives following the Cambridge Analytica scandal. *International Journal of Human-Computer Studies,143*, Article 102498. https://doi.org/10.1016/j.ijhcs.2020.102498. Accessed 2024-07-08.

Hoe Meta Gegevens Gebruikt voor generatieve AI-modellen (2024). https://www.facebook.com/privacy/genai/. Accessed 2024-07-09.

Jablonowska, A., & Michatowicz, A. (2020). Planet49: Pre-ticked checkboxes are not sufficient to convey user's consent to the storage of cookies case notes. *European Data Protection Law Review (EDPL)*, 6(1), 137–142. Accessed 2024-07-09.

Kairouz, P., McMahan, H. B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A. N., Bonawitz, K., Charles, Z., Cormode, G., Cummings, R., D'Oliveira, R. G. L., Eichner, H., Rouayheb, S. E., Evans, D., Gardner, J., Garrett, Z., Gascón, A., Ghazi, B., Gibbons, P. B., Gruteser, M., Harchaoui, Z., He, C., He, L., Huo, Z., Hutchinson, B., Hsu, J., Jaggi, M., Javidi, T., Joshi, G., Khodak, M., Konecný, J., Korolova, A., Koushanfar, F., Koyejo, S., Lepoint, T., Liu, Y., Mittal, P., Mohri, M., Nock, R., Özgür, A., Pagh, R., Qi, H., Ramage, D., Raskar, R., Raykova, M., Song, D., Song, W., Stich, S. U., Sun, Z., Suresh, A. T., Tramèr, F., Vepakomma, P., Wang, J., Xiong, L., Xu, Z., Yang, Q., Yu, F. X., Yu, H., & Zhao, S. (2021). Advances and open problems in federated learning. *Foundations and Trends® in Machine Learning*, 14(1–2), 1–210 https://doi.org/10.1561/2200000083. Publisher: Now Publishers, Inc. Accessed 2023-01-20.

Kelly, B. G., & Habib, M. (2023). Missed period? The significance of period-tracking applications in a post-Roe America. *Sexual and*

Reproductive Health Matters,31(4), 2238940. https://doi.org/10.1080/26410397.2023.2238940. Publisher: Taylor Francis_eprint: Accessed 2024-07-09.

Kusche, I. (2024). Possible harms of artificial intelligence and the EU AI act: Fundamental rights and risk. *Journal of Risk Research,1–14*,. https://doi.org/10.1080/13669877.2024.2350720. Publisher: Routledge_eprint: Accessed 2024-07-09

Li, L., Fan, Y., Tse, M., & Lin, K.-Y. (2020). A review of applications in federated learning. *Computers and Industrial Engineering,149*, Article 106854. https://doi.org/10.1016/j.cie.2020.106854. Accessed 2021-03-03.

Liu, B., Ding, M., Shaham, S., Rahayu, W., Farokhi, F., & Lin, Z (2020) When machine learning meets privacy: A survey and outlook. arXiv:2011.11819 [chttps://doi.org/10.48550/arXiv.2011.11819. Accessed 2023-07-25.

Lorenz, T., & Harwell, D. (2022). *Facebook paid GOP firm to malign TikTok*. Washington Post. Accessed 2024-07-17.

Lyu, L., Yu, J., Nandakumar, K., Li, Y., Ma, X., Jin, J., Yu, H., & Ng, K. S. (2020). towards fair and privacy-preserving federated deep models. *IEEE Transactions on Parallel and Distributed Systems*, 31(11), 2524–2541 https://doi.org/10.1109/TPDS.2020.2996273. Conference Name: IEEE Transactions on Parallel and Distributed Systems. Accessed 2024-07-09.

Maheshwari, S., & Holpuch, A. (2024). Why the U.S. is forcing TikTok to be sold or banned. https://www.nytimes.com/article/tiktok-ban.html. Accessed 2024-07-09.

Mauran, C. (2024). *Meta is using your posts to train AI. It's not easy to opt out*. Section: Tech. https://mashable.com/article/meta-using-posts-train-ai-opt-out. Accessed 2024-07-09.

Mehrnezhad, M., & Almeida, T. (2023). My sex-related data is more sensitive than my financial data and I want the same level of security and privacy: User risk perceptions and protective actions in female-oriented technologies. In *Proceedings of the 2023 European symposium on usable security. EuroUSEC'23* (pp. 1–14). Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3617072.3617100. Accessed 2024-07-09.

Mehrnezhad, M., Shipp, L., Almeida, T., & Toreini, E. (2022). Vision: Too little too late? Do the risks of femtech already outweigh the benefits? In *Proceedings of the 2022 European Symposium on Usable Security. EuroUSEC '22, pp. 145–150. Association for Computing Machinery, New York, NY, USA*. https://doi.org/10.1145/3549015.3554204. Accessed 2024-07-09.

Menn, J., & Nix, N. (2023). *Big tech funds the very people who are supposed to hold it accountable*. Washington Post. Accessed 2024-07-09.

Moor, J. H. (2006). *Using genetic information while protecting the privacy of the soul*. Ethics, Computing, and Genomics, Jones and Bartlett, Sudbury, MA (pp. 109–119).

Narayanan, A., & Shmatikov, V. (2007) *How to break anonymity of the netflix prize dataset*. arXiv:cs/0610105. https://doi.org/10.48550/arXiv.cs/0610105. Accessed 2024-07-08.

Ochigame, R. (2019). *How big tech manipulates academia to avoid regulation*. https://theintercept.com/2019/12/20/mit-ethical-ai-artificial-intelligence/. Accessed 2024-07-09.

Palmer, M. (2005). *Data is the new oil*. https://ana.blogs.com/maestros/2006/11/data_is_the_new.html. Accessed 2024-07-17.

Papaevangelou, C. (2024). Funding intermediaries: Google and Facebook's strategy to capture journalism. *Digital Journalism,12*(2), 234–255. https://doi.org/10.1080/21670811.2022.2155206. Publisher: Routledge_eprint: Accessed 2024-07-08.

Parsons, K. (2020). *Why taking Facebook quizzes is a really bad idea | CBC news*. https://www.cbc.ca/news/canada/nova-scotia/tech-columnist-warns-against-taking-social-media-quizzes-for-fun-1.5442282. Accessed 2023-07-31.

Project, T. T. (2023). *Zuckerberg and meta reach deep into academia*. https://www.techtransparencyproject.org/articles/zuckerberg-and-meta-reach-deep-into-academia. Accessed 2024-07-09.

Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union legislative Acts. COM/2021/206 final (2021). https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:52021PC0206

Register Berispingen | Autoriteit Persoonsgegevens. https://www.autoriteitpersoonsgegevens.nl/over-de-autoriteit-persoonsgegevens/register-berispingen. Accessed 2024-07-17.

Rehman, I. (2019). *Facebook-cambridge analytica data harvesting: What you need to know*. Library Philosophy and Practice, 1–11.

Reznichenko, A., & Francis, P. (2014) Private-by-design advertising meets the real world. In *Proceedings of the 2014 ACM SIGSAC conference on computer and communications security. CCS '14* (pp. 116–128). Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/2660267.2660305. Accessed 2024-07-09.

Roessler, B., DeCew, J (2023) Privacy. In E. N. Zalta, U. Nodelman (Eds.), The Stanford Encyclopedia of Philosophy, Winter 2023 edn. Metaphysics Research Lab, Stanford University. Accessed 2024-07-17.

Roush, T. (2023). Meta launching paid subscriptions to use facebook and instagram ad-free in Europe. Section: Business. https://www.forbes.com/sites/tylerroush/2023/10/30/meta-launching-paid-subscriptions-to-use-facebook-and-instagram-ad-free-in-europe/. Accessed 2024-07-09.

Scheenstra, B., Bruninx, A., Van Daalen, F., Stahl, N., Latuapon, E., Imkamp, M., Ippel, L., Duijsings-Mahangi, S., Smits, D., Townend, D., Bermejo, I., Dekker, A., Hochstenbach, L., Spreeuwenberg, M., Maessen, J., & Van 'T Hof, A., & Kietselaer, B,. (2022). Digital health solutions to reduce the burden of atherosclerotic cardiovascular disease proposed by the CARRIER consortium. *JMIR Cardio,6*(2), 37437. https://doi.org/10.2196/37437. Accessed 2024-07-08.

Shah, S. (2024). *How to opt out of Instagram and Facebook training AI on your photos*. Section: News. https://www.standard.co.uk/news/tech/instagram-facebook-train-ai-photos-posts-how-opt-out-b1162065.html. Accessed 2024-07-09.

Shaw, D. (2024). *Meta shatters lobbying record as house passes TikTok Ban*. https://readsludge.com/2024/04/23/meta-shatters-lobbying-record-as-house-passes-tiktok-ban/. Accessed 2024-07-09.

Shokri, R., Stronati, M., Song, C., & Shmatikov, V. (2017). Membership inference attacks against machine learning models. In 2017 IEEE symposium on security and privacy (SP) (pp. 3–18). https://doi.org/10.1109/SP.2017.41. ISSN: 2375-1207. https://ieeexplore.ieee.org/abstract/document/7958568. Accessed 2024-07-09.

Siu, L., & Chun, C. (2020). Yellow peril and techno-orientalism in the time of Covid-19: Racialized contagion, scientific espionage, and techno-economic warfare. *Journal of Asian American Studies*, 23(3), 421–440. Publisher: Johns Hopkins University Press. Accessed 2024-07-09.

Slokom, M., Wolf, P.-P., spsampsps Larson, M. (2022). When machine learning models leak: An exploration of synthetic training data. In J. Domingo-Ferrer, M. Laurent (Eds.), *Privacy in statistical databases*, pp. 283–296. Springer. https://doi.org/10.1007/978-3-031-13945-1_20.

Srivastava, S., Wilska, T.-A., & Nyrhinen, J. (2023). Awareness of digital commercial profiling among adolescents in Finland and their perspectives on online targeted advertisements. *Journal of Children and Media,17*(4), 559–578. https://doi.org/10.1080/17482798.2023.2257813. Publisher: Routledge_eprint: Accessed 2024-07-09.

Tavani, H. T. (2008). *Informational privacy: Concepts, theories, and controversies*. The handbook of information and computer ethics, 131–164

Townend, D. (2025). Big data research: Can confidentiality and fiduciary duties fill in the gaps in privacy and data protection? In *Confidentiality, privacy, and data protection in biomedicine* (pp. 81–101). Routledge.

Truong, N., Sun, K., Wang, S., Guitton, F., & Guo, Y. (2021). Privacy preservation in federated learning: An insightful survey from the GDPR perspective. *Computers and Security,110*, Article 102402. https://doi.org/10.1016/j.cose.2021.102402. Accessed 2022-11-07.

Vatsalan, D., Sehili, Z., Christen, P., spsampsps Rahm, E. (2017). Privacy-preserving record linkage for big data: Current approaches and research challenges. In A.Y. Zomaya, S. Sakr (Eds.) Handbook of Big Data Technologies (pp. 851–895). Springer.https://doi.org/10.1007/978-3-319-49340-4_25. Accessed 2021-07-07.

Verma, S. (2019). Weapons of math destruction: How big data increases inequality and threatens democracy. *Vikalpa: The Journal for Decision Makers*, 44(2), 97–98 https://doi.org/10.1177/0256090919853933. Accessed 2024-07-08.

Xenophobic Machines. (2021). *Discrimination through unregulated use of algorithms in the Dutch childcare benefits scandal*. https://www.amnesty.org/en/documents/eur35/4686/2021/en/. Accessed 2024-07-17.

Xu, R., Baracaldo, N., & Joshi, J. (2021) *Privacy-preserving machine learning: Methods, challenges and directions*. arXiv:2108.04417 [cs]. https://doi.org/10.48550/arXiv.2108.04417. Accessed 2023-07-25.

Yang, L., Tan, B., Zheng, V.W., Chen, K., spsampsps Yang, Q. (2020). Federated recommendation systems. In Q. Yang, L. Fan, H. Yu (Eds.) Federated learning: privacy and incentive. Lecture notes in computer science (pp. 225–239). Springer. https://doi.org/10.1007/978-3-030-63076-8_16. Accessed 2023-09-14.

Yang, Q., Liu, Y., Chen, T., & Tong, Y. (2019). *Federated machine learning: Concept and applications*. arXiv:1902.04885. Accessed 2021-05-18.

Yao, A. C. (1982). Protocols for secure computations. In *23rd Annual Symposium on Foundations of Computer Science (sfcs 1982)* (pp. 160–164). https://doi.org/10.1109/SFCS.1982.38. ISSN: 0272-5428.

Zuboff, S (2023) The age of surveillance capitalism. In *Social theory re-wired* (pp. 203–213). Routledge.