Scientific Research Publishing

Beijing Law Review, 2026, 17(2), 523-545
https://www.scirp.org/journal/blr
ISSN Online: 2159-4635
ISSN Print: 2159-4627

AI Training with Personal Data on Social Media: Regulatory Response in Brazil and in the European Union

Beatriz Graziano Chow

Department of Law, Fundação Getulio Vargas, São Paulo, Brazil
Email: beatrizchow2@gmail.com

How to cite this paper: Chow, B. G. (2026). AI Training with Personal Data on Social Media: Regulatory Response in Brazil and in the European Union. *Beijing Law Review, 17*, 523-545.
https://doi.org/10.4236/blr.2026.172028

Received: April 6, 2026
Accepted: June 12, 2026
Published: June 15, 2026

Abstract

The widespread use of personal data found in social media platforms for artificial intelligence (AI) training purposes raises significant legal concerns, particularly regarding data protection and privacy. This paper critically examines the legal boundaries and possibilities involved in the use of personal data extracted from social media platforms for AI training, with a focus on the Brazilian General Data Protection Law (LGPD) and the General Data Protection Regulation (GDPR). It investigates the legal foundations for processing such data through the analysis of concrete cases and the challenges arising from them, including the need to justify further processing of personal data, as well as the lawfulness of legal bases arising from this context. To this end, current industry practices, decisions from data protection authorities, and regulatory initiatives are examined, considering the approval and enforcement of the EU AI Act. As a result, the investigation unfolds that the current baseline for personal data protection presents significant gaps in light of the complexity of the digital context emerging around increasing adoption of AI. Therefore, it is proposed that legal guidelines must be established within robust AI governance, in line with technical solutions, to ensure that personal data processing for the purpose of generative AI training is compatible with data protection and privacy.

Keywords

Data Protection, Privacy, Artificial Intelligence, Generative Artificial Intelligence, Social Media Platforms, GDPR, LGPD, EU AI Act

1. Introduction

A new paradigm has begun to unfold considering the advance of artificial intelli-

gence (AI), which has been increasingly integrated into society since the launch of ChatGPT in November 2022. This new form of interaction has raised profound social and regulatory repercussions.

As exciting as the state of the art in AI may be with the spread of the so-called LLMs, which is a type of generative AI, there is an evident issue of large volumes of data that must be collected and used for the training of a model in the development stage and eventually when the AI is deployed. This layered complexity inherent to generative AI gives rise to risks and potential consequences associated with the use of the technology.

AI governance can be highlighted within this complex scenario, serving as an essential foundation for successful implementation within organizations and as the primary driving force that enhances its effectiveness. Without mature governance of both AI and data, AI models are hindered in delivering responsible AI to their users.

Data used to train AI, which is in extremely high demand, impacts how data protection rights should be preserved, for example in the social media ecosystem, assuming that users' personal data found in the platform is used for the training process of a generative AI model, including for cases in which the very platform is creating its own proprietary model.

Given the complexity of how technology travels in this era, this article will analyze the regulatory applicability that can be identified in this specific context of social media platforms making efforts to detain their own AI models/systems.

2. Treasure-Trove: Personal Data in Social Media

It is a fact that social media influences users' daily lives, affecting habits, tastes, opinions, helping shape popular culture and trends, and above all, having the ability to modify how society interact in society and in their intimate sphere.

The greatest competitive advantage of social media companies to acquire data is that every image, post, interaction, video, comment, among others, are easily accessible sources, since the environment and database are managed by the platforms themselves. These data, which evidently are composed of personal data and therefore could be seen as a vivid reflection of an individual, directly impact their social relationships, which justifies the need for specific guardrails focused on the autonomy of personal data protection in the face of challenges posed by AI.

In search of data, including high quality data, one can state that social media platforms have significant leverage in this competitive scenario to extract a large amount of data and use it for generative AI training purposes. So, it is not a coincidence that popular social media companies are positioning themselves fiercely in the AI race, as follows.

2.1. Meta's Use of Personal Data for AI Training

In June 2024, Meta announced updates to its privacy policy, with the prerogative of using personal data for AI training, including improving its proprietary Llama

model. Thus, all data made available on its platforms, whether on Facebook, Instagram, or Threads, could be used for such purposes.

In Europe, as a response to this announcement, the non-profit association co-founded by Max Schrems, Noyb, a non-profit organization acting in defense of digital rights and privacy in Europe, called out Meta's intent to process user data for AI training without proper consent before 11 European data protection authorities (Nyob, 2024a). In the complaint, they argued that this practice would violate GDPR (European Union, 2016) principles by not offering an opt-in option and clearly pointed to the existence of a dark pattern when applying the opt-out.

The stance of Meta also caught the attention of the Data Protection Commission (DPC), Ireland's national data protection authority, where Meta's headquarters is located. The authority and the company reached an agreement that suspended Meta's plans. Meta stated that its objective was to train their AI models with European's users data who were over 18 years old and found as publicly available information (i.e., personal data that a data subject has made accessible to an unrestricted audience on the platform), as opposed to private information, such as private profiles and messages.

Similarly, in Brazil, the country's National Data Protection Authority (ANPD) had to intervene by issuing a suspension of personal data processing for Meta's AI training of Brazilian users in vote n. 11/2024/DIR-MW/CD (Autoridade Nacional de Proteção de Dados, 2024a). The suspension is based on the lack of reasonable expectations for data processing for a specific purpose (i.e., what a user would foresee as a likely use of their data according to the disclosed purposes found in the platform's terms and conditions and privacy policy), with penalties such as fines for non-compliance. The vote pointed out to the following: i) absence of applicable legal basis; ii) lack of transparency in disclosing information to data subjects; iii) non-adherence to data subjects' rights; and iv) processing of personal data of minors under 18 years old without appropriate safeguards.

Unlike Europe, the response came after the published privacy policy version in July 2024 due to the fact that the announcement about the privacy policy update in Brazil was not announced. In the European Union (EU), Meta platform users were previously informed by email and app notifications of the company's plans, which provided greater transparency regarding the privacy policy in contrast to the experience of Brazilian users.

According to the new privacy policy, Meta would collect personal data publicly made available by Brazilians on Instagram and Facebook. This included various types of media, such as images, audio, and posts, aiming to train and improve their generative AI. Third parties whose images are disclosed or who are mentioned indirectly on Meta platforms are included in the scope of the training data.

In the vote, the lawfulness of the processing was also a major point of concern, as the legitimate interest was apparently unsuited for the intended purpose. The requirement of the data subject's reasonable expectation was not met. Within the scope of training data, there is also the matter of processing special category of

data—sensitive data (i.e., data revealing racial or ethnic origin, religious belief, political opinion, union membership, health or sex life, genetic or biometric data, or other special category information under GDPR and the broadly equivalent of sensitive data under LGPD), which requires a new legal basis framing as per art. 11, LGPD (Brazil, 2018) that did not occur in this specific case. Another relevant factor is that the principles of necessity and purpose limitation were not observed.

Moreover, Meta did not provide plain information about the consequences of processing data for the development of generative AI models. In the Privacy Channel, the company only generically mentions future benefits without detailing risks to data subjects and third parties.

The opt-out mechanism for users to object to AI training was not presented to Brazilian citizens in a facilitated manner. Several steps and actions were necessary to oppose to the processing of their personal data, such as locating hidden pages to access an extremely detailed opt-out form that demanded a lot of details on why the user wanted to do so. All these mechanisms are evidently designed to burden users into creating unnecessary hurdles to meet their rights. In contrast, for EU data subjects, the opt-out link was sent directly by email.

Therefore, the opt-out model had an underlying dark pattern for Brazil, since users needed to follow an extensive path to get to the appropriate page to submit the request and therefore be able to exercise the rights provided for in art. 18 of the LGPD.

In August 2024, the ANPD let Meta resume the processing of personal data following Meta's cooperation and efforts to the compliance plan proposed by the Authority in Decision n. 23/2024/DIR-JR/CD (Autoridade Nacional de Proteção de Dados, 2024b). The measures would then promote transparency while allowing users to opt-out of using their personal data for the purpose of AI training by Meta. However, the suspension was maintained for minors.

In response to the authority's requirements, the company did indeed implement measures to promote greater transparency and compliance in the processing of personal data. However, at the conclusion of this case, the ANPD recognized that dealing with sensitive data posed a complex challenge with respect to the lawfulness of processing, especially considering the ongoing evaluation of such matter by other relevant data protection authorities across the world, especially in the EU.

Privacy policy documents on the Meta website vary according to its location. In the case of Brazil, there were no changes in the privacy policy when mentioning "artificial intelligence" in the versions published on June 26, 2024, October 9, 2024, and November 14, 2024 (Meta, 2024a). There was a removal of AI training references in the privacy policy published on July 9, 2024 (Meta, 2024b), due to the ANPD's decision in the first vote abovementioned.

The AI references in most versions are limited to informing that: i) user data is used to "enable the creation of content such as text, audio, images, and videos, including through the artificial intelligence technology we provide"; and ii) to

support "research in areas such as artificial intelligence and machine learning". The most relevant information for users to understand how and what information is being used for Meta's generative AI training was added later through more detailed privacy notices for Brazil.

There is a page linked in the privacy notice titled "How Meta uses information for generative AI models and features" (Meta, 2025) that even though it has been referenced in the Brazilian privacy notice, is only available in English, which creates a language barrier for Brazilian users. The page promotes necessary clarifications for users about Meta's AI purposes, regarding the following points: i) user's publicly available data; ii) pillars of the company's privacy protection strategy; and iii) opt-out mechanism, thus promoting the exercise of data subjects' rights.

In the specific case of Brazil, there was additional information provided that made the intended purpose and use of AI in the concrete case much more coherent, and a provisional attribution of a legal basis, followed by solid measures frequently monitored by the Authority. The ANPD ensured that Brazil succeeded in improving transparency for data subjects with the allocation of responsibilities under the LGPD.

It is worth highlighting that a company with global presence will often adopt locally especially considering different targeted outcomes, which weakens the wide adoption of the right to data protection and privacy throughout all operations.

On May 14, 2025, the debate reignited when Meta announced plans to resume personal data from Instagram and Facebook users in the EU to train its new AI systems. Nyob once again reacted and sent a cease-and-desist letter to the platform (Nyob, 2025), reaffirming the need to request consent through opt-in, as opposed to Meta's claims of using legitimate interest to collect all user data lawfully.

The platform faced significant pressures to stop using legitimate interest as a legal basis, and to move away from the opt-out system for AI training purposes. Additionally, there is the possibility of a collective action in the EU against the company, alleging consumer damages, which could result in billions.

Thus, Meta bet on the limits of the legal basis of legitimate interest to train AI with public user data and even controversially with private data, a fact denied by the company, attracting evident media and public attention. Despite measures taken regarding user transparency and promotion of improvements in the opt-out feature, which was adopted later, and efforts in dialogue with various data protection authorities, doubts remain towards the lawfulness of processing around sensitive data.

Finally, in this case, the company's intent to build its own AI project by developing and improving proprietary models (e.g., Llama) is evident. The platform attempted to build an ecosystem in favor of its AI projects, without specific safeguards in place and tried to convert user data into competitive advantage while reducing potential need to rely on third parties for acquiring data.

2.2. Personal Data Use in X for AI Training

The case of X, previously known as Twitter, and currently a xAI company, was the subject of repercussions regarding the use of personal data for training the company's AI, the Grok model. The company silently introduced in its platform the option to use personal data from its users for AI and machine learning models, as provided in its privacy policy without going into further details. In a supplementary document to the privacy policy, the social media platform declared that they acted according to their legitimate interest as the legal basis to support all data processing.

In Europe, when it was revealed that the practice had been going on for months, the case attracted the attention of various authorities, starting with the DPC. It initiated an urgent action, citing serious concerns about X's use of personal data from public posts of millions of EU/EEE users to train Grok.

In August 2024, the authority stated that the users were still having their data processed without prior knowledge and their posts were being used to integrate X's AI training, without any appropriate safeguards and in non-compliance with the GDPR. X denied engaging in any unlawful conduct. After negotiations took place, the DPC understood that X introduced the necessary mitigatory measures in response, such as the insertion of an opt-out mechanism for users, which was not previously made available, which is a basic right of data subjects.

Similar to the Meta case, Nyob filed complaints about X's practices before 9 data protection authorities in Europe to protect the rights of European citizens, alleging that the appropriate legal basis for this case would be consent, with the insertion of an opt-in mechanism instead of opt-out, which was subsequently incorporated (Nyob, 2024b). The association criticized the DPC's conduct directly, as it was considered very lenient in its response to X, as it apparently was acting in favor of the company.

Among the main points raised in the complaint, the following were cited: i) X did not inform its users when it changed the privacy policy to accommodate the provision of using personal data for AI training; ii) X did not provide an opt-out option for data subjects in advance; iii) in July 2024, without prior announcement, X activated a default setting (Milmo, 2024) that allowed users to have their data processed for AI training purposes; iv) when offering opt-out to users, the steps to do so were not understandable, requiring an active search by the user to find the privacy setting; and v) X did not provide specific details about which personal data it would process, whether that data would be separated between sensitive and non-sensitive data, or if there would be distinction by data processing per location, etc.

Following up of the case, in August 2024, the DPC entered into an agreement with X to suspend the use of personal data from public posts of EU/EEE users collected between May 7 and August 1, 2024, for the Grok training (Data Protection Commission, 2024a). After a month, in September 2024, X agreed to definitively comply with the terms of the commitment made in August. This resulted in

the conclusion of ongoing legal proceedings, and the case was formally closed (Data Protection Commission, 2024b). However, regarding the data that had been previously incorporated into the model (Bray, 2024) the issue persisted including whether there would be a possibility of deletion.

During this process, the DPC requested a statement from the European Data Protection Board (EDPB), in accordance with Article 64(2) of the GDPR, with the intent of promoting greater clarity and consensus about the main legal and regulatory issues related to the use of personal data in AI training. This consultation covers themes such as data processing and applicable legal foundations. The EDPB did indeed return with this request as discussed in more detail in item 3.3.

In Brazil, the ANPD (Causin, 2024) issued a decision order in December 2024, to investigate X's practices on this data processing in its platform for AI training, such as the disclosure of privacy policy and terms of use, to request mitigatory measures, and to evaluate overall compliance with the LGPD. As part of the measures, the ANPD asked for clarification in relation to the documentation presented as evidence of compliance (e.g., balancing test, which is the equivalent to the LIA—Legitimate Interests Assessment; and DPIA—Data Protection Impact Assessment) as to understand whether it would be applicable according to the local jurisdiction, since it only referenced legislation from the EU, European Free Trade Association (EFTA) countries, and the United Kingdom. In case it failed to comply with the LGPD, the company should present equivalent versions of such documentation that meet the local criteria.

Regarding the purpose of the data processing, the ANPD determined that X should delete the expression "for any purpose" from its terms of use. Additionally, considering the types of data used in generative AI training, X would have to be very explicit in the legal documentation on which data was used for the intended purpose. There would also need to be an inclusion of the scope of the training data, that is publicly accessible personal data, except for data from protected or private accounts and data in private messages between users.

Furthermore, X should immediately stop using personal data from accounts belonging to minors in Brazil for AI training purposes. Evidence of compliance with such order should be made available, signed by a legal representative or by the DPO (Data Protection Officer). Moreover, the company should include, in its privacy policy, or Help Center information about the non-existence of the processing of minor's data and disable any data sharing option aimed at generative AI training for these users.

Brazil's version of the privacy policy has not been updated so far (Twitter, 2020), unlike the global privacy policy (X, 2024), to convey with the ANPD's requirements, to justify the processing of data with clear purposes instead of vague expressions, such as "to improve AI models". The ANPD in Decision n. 29/2024/FIS/CGF (Autoridade Nacional de Proteção de Dados, 2024c) has not made any public statement after its decision in its official channels; therefore, the company's con-

duct analysis was left inconclusive and to that extent, users in Brazil did not have their rights effectively protected.

Many countries reacted with investigations into X's practices, such as Switzerland. In March 2025, the FDPIC (Federal Data Protection and Information Commissioner), the country's data protection authority, concluded that users' opt-out was guaranteed and that data processing was in accordance with the FADP (Federal Act on Data Protection). In its position, the authority understood that users also have responsibility for the information they make public available (Eidgenössischer Datenschutz und Öffentlichkeitsbeauftragter, 2025).

This is an extremely relevant point, considering that users often do not measure the consequences of data sharing, including sensitive data. In November 2024, Elon Musk urged users to share medical exams with images (MRI, X-ray, etc.) when interacting with Grok (Passarella, 2024). Users voluntarily engaged with the idea, which led to a large volume of health data being shared and used in the training of the AI tool.

On April 11, 2025, the DPC resumed the X's AI training case in the EU and announced the start of an investigation into the processing of personal data of Europeans contained in publicly available posts in X for training generative AI models, particularly Grok. The investigation aimed at evaluating compliance with the GDPR, including legality and transparency in processing (Data Protection Commission, 2025).

X initiated its AI training with publicly available information found in the social media platform while sustaining the argument of having a legitimate interest to do so, with inefficient communication to users at first, as well as an absent opt-out feature, which resulted in interventions across the EU to modify the scenario and implement appropriate safeguards. The company's strategy led to reputation damage, as well as being left with a questionable legal basis to process personal data.

The company's intent to build its own AI project Grok was clear. The social media's vision was focused on transforming data shared on the platform to their X products, adding value to it with easily accessible resources. Additionally, in September 2025, the platform was once again criticized, as user conversations with Grok were leaked on search engines (McMahon, 2025) constituting an evident exposure of intimacy (e.g., sensitive mental health, relationship data, etc.).

2.3. Personal Data Use by LinkedIn for AI Training

Following the benchmark established by its competitors, LinkedIn did not fall behind. In September 2024, the company announced new changes in the privacy policy to use personal data for generative AI training (LinkedIn, 2024). The conduct was similar to X's case, with the implementation of an easy to find opt-out, new changes in the privacy policy to process users' personal data for generative AI training.

In this case, LinkedIn's silent action did not attract as much repercussion. There

is a reason for this. LinkedIn acted quickly so that the company's new business strategy would not draw too much attention. LinkedIn announced officially on its blog that the generative AI training would not include data from members of the EEA, Switzerland, and the United Kingdom (LinkedIn, 2024) Additionally, within the FAQ section (LinkedIn, 2025a), the social network declares that it does not process personal data from users in Canada, EU, EEA, United Kingdom, Switzerland, Hong Kong SAR, or China for AI training purposes.

In relation to LinkedIn's advance on generative AI training and development, relevant information on the company's strategy can be found in the FAQs, privacy policy (LinkedIn, 2025b), and in the Help Center website, where there is a specific section explaining how to control the opt-out feature (LinkedIn, 2025c). The information is clear regarding the use of personal data and the purpose of AI processing, as specified in the FAQ section. The legal bases highlighted for data processing by the company are legitimate interest and consent, with an emphasis on the mention of consent use with the possibility of revoking it.

This case is slightly different from the others, since LinkedIn did not focus on developing its own proprietary AI model, rather it aimed to leverage data for existing generative AI models in order to customize it for its own needs. This will be a very common practice in the market, since creating a completely new model is a high-cost investment. (Hacker & Holweg, 2025)

3. The State of the Art on Personal Data Processing in Social Media Platforms for AI Training Purposes

In the cases above, it became evident that there is an evolving intersection between personal data and AI advancement demonstrated through processing of data found in social media platforms, whether for training a proprietary model or to customize AI models available in the market. In these scenarios, there is already a pre-existing relationship between the data controller and data subject, and, consequently, a reasonable expectation of certain forms of processing. As mentioned, this rely on clear purposes to data processing established in the platform's terms and conditions and privacy policy, typically covering core social media services (e.g., content creation and publishing; user connection, messaging and interaction; content discovery and personalization; and identity/profile management), so repurposing this processing to AI training inadvertently causes a clear shift in expectations.

Therefore, this data processing activity can fall under the further processing scope (i.e., the subsequent use of personal data for a purpose other than the one for which it was initially collected), so the extent of applicability of both the LGPD and the GDPR must be analyzed.

Data protection is not a new regulation, and companies have had time and resources to adhere to the obligations set by the legislation for the past decade. This means there are enforceable measures for personal data processing, such as purpose limitation, legal basis, transparency mechanisms, necessary documentation

to attest for data subject's rights, etc. In the challenging context brought by AI, these measures should be reviewed and analyzed, as further discussed.

3.1. Legal Impositions under the LGPD and the GDPR

The purpose limitation principle provided in art. 6, LGPD reflects that the processing must be conducted for legitimate, specific, explicit purposes, informed to the data subject, without the possibility of further processing in a manner incompatible with those purposes. In the GDPR, the principle is defined in art. 5 and in short states that the processing must occur for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes. Both concepts are almost identical, which strengthens its universal applicability.

Essentially, this principle aims to inform the data subject about the ways in which their personal data will be primarily used, and not for secondary and hidden purposes, even when faced with the same processing context. In the case of data processing for social media users, the primary context is evident: to enter the so-called "community" established in the platforms. While navigating and integrating into such community, the users are subject to interactions with targeted advertising, recommendation algorithms, and many other forms that have been around for long, etc. However, using these data to feed a scalable product or service in light of new technologies is something entirely different. The user did not sign up for this nor could have envisioned such data processing.

The use of personal data for different purposes than their original one, in this case, for training a generative AI model resulting directly from the data controller's decision to turn personal data as inputs, is a form of further processing. This entails subsequent processing for a new or secondary purpose that differs from the purpose communicated at collection, in the sense of GDPR and the equivalent purpose limitation logic under LGPD. Therefore, framing AI training as further processing reinforces the need to demonstrate purpose compatibility, transparency, and an applicable legal basis for the secondary purpose.

The vague reasons portrayed by the social media platforms to justify such purposes do not ensure transparency to users. As previously analyzed in detail, the companies did not come forward with explanations and details in their official communications, documents, as it is simply stated that they'll be using the data to the improvement of services, to help in the training of their AI models/systems.

One of the allegations presented by Noyb in both Meta and X complaints, is that the companies failed to demonstrate grounds for alleging the legal basis of legitimate interest for further processing of personal data, including the requirement of proving that there was an actual legitimate interest in the first place, which would conflict with the purpose limitation principle. Thus, there was a violation of art. 5, (1) b), GDPR. It was argued that the further processing remains incompatible with the original purpose. There is an evident mismatch

between the initial processing purpose and the subsequent data use, which was not indeed justified.

Also, there is a second factor to be addressed within data protection legislation: the lawfulness of the processing that comes through the legal bases, which are the hypotheses that make the data processing activity lawful. The legal bases reflected in both the LGPD and GDPR are very similar, and their applications were designed for shared scenarios. Additionally, there is the special category of data provided in the GDPR, which configures sensitive data broadly. In this context, both laws tie the lawfulness in this case as something different, with its own set of legal bases for this type of data.

Often, individuals in their social media interactions may disclose sensitive data, for example, by revealing information about a health condition, a specific treatment, or even a third party's health information (e.g., to ask for blood donations, etc.). Additionally, it is common to see expressions of intimate beliefs about religion, different genders, gender, and political views. So, these types of sensitive data are constantly made available by users, compounding large-scale information found in social media.

Social media platforms, such as Meta argue that it is technically impossible to separate data between sensitive and non-sensitive personal data for generative AI training purposes, as stated in case C081-11 (Nyob, 2024c). Consequently, if this is not possible, the attribution of an appropriate legal basis remains impaired under arts. 7 and 11, LGPD and arts. 6 and 9, GDPR.

Both laws are very explicit in stating that when personal data processing results from an economic activity, a legal basis must be identified for that specific scenario. The purpose limitation principle states that the data subject must be aware of its purpose from its collection, so that a reasonable expectation about the processing can occur. It is the data controller responsibility to promote transparency measures from the beginning of their interaction and whenever relevant changes take place within the data processing scope.

In this context, there are two legal bases constantly addressed in this ongoing debate, envisioning the possibility of using personal data for generative AI training purposes: i) consent and ii) legitimate interests. It is worth noting that the purpose limitation principle is intrinsically linked to the application of these hypotheses.

3.2. Consent

The legal basis of consent must be freely given, specific, informed and unambiguous, applied specifically for a determined purpose through the data subject's manifestation of will. To this end, the data subject must be informed prior to any data processing activity to be considered valid.

The big issue with consent relies on how it manifests, since there needs to be certain conditions in place. Therefore, operational guarantees must often be implemented, such as the opt-in alternative. Opt-in is a mechanism that allows the

data subject to give their consent and acknowledgement for the processing of their personal data. On the other hand, the opt-out is not considered a valid alternative for expressing consent. Pre-checked boxes are also considered invalid as they do not represent effective opt-in measures, as well as opt-out solutions that require intervention by the data subject to express their giving consent (European Commission, 2018).

According to the complaints presented against Meta (C081-11) and X (C087-06) by Nyob, both in 2024, it was alleged that the opt-in resource was incorrectly disregarded as an alternative. It was argued that the companies used measures that prevented data subjects from exercising their right of choice, of consent, not being able to opt-in, rather, they only had a choice to opt-out.

Considering that consent must be freely given when faced with a choice: to consent or not to, without resulting in consequences, the same applies to the revoke of consent, which the data subject could rely on at any time if there is a change in the purpose limitation, as per art. 9, §2, LGPD. As of art. 7, 3, GDPR, consent can be revoked at any time and must be provided in a facilitated process. In this sense, it is necessary for the controller to manage consent so that the data subject can opt to withdraw it and the controller can meet this request.

In the discussed scenario, the situation becomes a bit stretching, since it demands keeping track of all consent that were made or revoked by millions of users. The controller must demonstrate compliance with their request; and so, a viable and effective method must be identified whereby the data subject receives the guarantee that their data will no longer be processed once consent has been revoked.

Also, in generative AI training using personal data from social media's users, the grounding of this legal basis becomes technically impractical, since the company will not be possible to erase data once used in the training, as stated in Nyob case C087-06 (Nyob, 2024d). As such, what would be feasible is to have personal data excluded from future LLM training cycles, given that the retroactive effect would not be possible here.

Another relevant factor in art. 7, §4, LGPD is the exemption from informed consent when data is made publicly available by the data subject, provided that their rights and LGPD principles are safeguarded. This makes way for the application of other legal bases in this context, as demonstrated with the application of legitimate interests.

For those reasons, the use of consent as a legal basis has been framed as a challenging alternative in large-scale social-media AI training scenarios, particularly where: i) the notice is not sufficiently granular; ii) an opt-in design is not realistically implemented; or iii) the control over consent is hard to operationalize in practice for the data controller. At the same time, enforcement practice is not uniform: some stakeholders (such as Noyb in the Meta and X cases discussed above) argue that opt-in consent should be the default, whereas controllers have preferred legitimate interest and some authorities have focused their interventions on trans-

parency, forms to perform the right to object and safeguards rather than requiring consent in every instance.

3.3. Legitimate Interests

One of the hypotheses deemed lawful for the processing of personal data is the pursuit of the legitimate interests of the controller or a third party, provided that the fundamental rights and freedoms of the data subject are not overridden.

For the legitimate interests to be a valid and applicable basis, the following criteria must be met: i) the interest must be legal, fair, clear, precise, real, and present; ii) the personal data to be processed must be necessary for the processing purpose; and; iii) the interests or fundamental freedoms and rights of the individuals in question must not be nullified by the use of legitimate interest. To demonstrate its applicability, this basis requires appropriate documentation to showcase that specific conditions were reached so that the data controller could lawfully rely on this basis.

Regarding the documentation and testing, one of these requirements is the evaluation of the necessity of the processing. To account for that, an important factor to be evaluated is whether data processing respected the data minimization principle, and if there was proportionality between the data collected and used for AI training. Considering generative AI development, answering this question becomes increasingly difficult.

To conduct a balancing test or LIA, for LGPD and GDPR compliance respectively, it is important to consider the impact on the data subjects, which can vary according to the type of personal data. This categorization is a key contributing factor to the assessment, and potential consequences arising from the AI model implementation context.

An essential fact to the applicability of the legitimate interest legal basis is within the data subject's reasonable expectation. The controller must evaluate and be able to demonstrate that personal data processing is expected by the data subject in the practical scenario (Autoridade Nacional de Proteção de Dados, 2024d). Other factors are imperative for fulfilling this requirement: i) established relationship between the controller and the data subject; ii) personal data collection: the source (website or derived from the service itself), the collection method (direct, by third parties, such as via scraping or from public sources), the context and date of collection; and iii) the purpose of data collection and its compatibility with legitimate interest application. In the specific context of GDPR, the data subject's reasonable expectation is impacted by the form data was made available, if made public or not, as demonstrated in the real cases highlighted earlier.

Therefore, legitimate interest is a legal basis that many platforms and authorities have treated as a trending alternative for lawful personal data processing for social media companies when training their own AI system/model (Solove, 2024). Considering the feasibility to apply this legal basis in practice and the continuous reliance on such legal basis, regulators' responses to date often turn on whether

the assessment is credibly documented (balancing test/LIA), and whether reasonable expectations, transparency, and valid opt-out are meaningfully operationalized.

However, there are two main obstacles when dealing with this approach. First, the argument that the purpose is to "serve the commercial interests" of the company, typically found in standard terms of use of platforms and presented in a generic manner. In this context, a significant part of information and clarifications are lacking about the purposes for which AI can be applied in the real world, even while in the development stage of the AI lifecycle. Second, there is the issue of relying on legitimate interests when applying it to sensitive data, even when data is made publicly available, when faced with the applicability of the LGPD. In the case of the GDPR, it is possible to adhere to this in a lawful manner, since there is an exemption in the law allowing the processing to take place when sensitive data that been made publicly available. For other cases, the GDPR presents the same rationale as the LGPD. There is still a need to comply with the obligations of presenting additional documentation and other requirements to legitimize such a basis in personal data processing in this scenario as per the GDPR (Solove, 2024).

Additionally, there is an inherent challenge in separating personal data from non-personal data. Taking that to another level, the separation between personal and sensitive personal data is yet more complicated. This becomes an obstacle to obtaining lawfulness in this context, since the types of data are intertwined. Moreover, the volume of this data only increases over time, especially considering how long social media has been around collecting user data.

In the opinion requested from EDPB as a response to the X case, mentioned above, they issued an official opinion (European Data Protection Board, 2024) on certain data protection aspects related to the processing of personal data in the context of AI models. The general topics discussed were: i) when AI models can be considered anonymous; ii) the application of legitimate interests as a legal basis; and iii) the impact of unlawful processing (in the development phase) on further processing of data.

The need for clarification becomes evident in aspects such as special data categories (art. 9, GDPR) and in the draft of DPIAs, which is another unparalleled documentation to ensure fundamental rights. There is emphasis on robust technical and organizational measures, detailed documentation on data subjects' rights and reasonable expectations.

The AI lifecycle stages are also addressed in the Opinion and were split into two stages: development and deployment, which can involve different legal bases in each stage with distinct purposes. So, an evaluation of each specific case is imperative to ensure compliance with the legislation. Cases may happen where a basis is applied at the end of the process in the deployment phase, but in the previous AI model development stage, it engaged in illicit practices, not meeting the expected legal basis requirement, for example. As such, the opinion imposes measures to enable the processing, such as applying corrective measures to mitigate the un-

lawful initial processing performed by data controllers (European Data Protection Board, 2024).

When it comes to the application of legitimate interest in practice, data protection authorities across the EU did leave room for interpretation. AP, the Dutch data protection authority and CNIL, the French data protection authority published materials about the application of legitimate interest when faced with specific circumstances, such as data scraping and AI development.

The AP released a guide for private entities and individuals performing scraping practices, in which it entailed data processing in social media as part of its scope. Furthermore, the authority indicates that scraping would ideally be justified on the legitimate interest legal basis (Digital Policy Alert, 2024).

CNIL discussed additional measures for applying legitimate interests, such as fulfilling the cumulative requirements for the applicability itself (documentation, LIA, DPIA, etc.), including for scraping (CNIL, 2026a) for AI system development, without delving into the legal basis application for sensitive data processing cases. The authority compiled a step-by-step page (CNIL, 2026b) to guide developers to follow the GDPR in this very context of application.

Similarly, in the case of ANPD in Brazil, as analyzed in the Meta case, the authority chose to proceed with the application of legitimate interest as a legal basis after the company presented the required documentation and supplementary clarifications on the data processing. However, the issue of sensitive data processing was not addressed—leaving room for interpretation.

Social media platforms were originally seen as a community for socializing, but the dynamics have changed over the years. Therefore, additional measures were required by law and as of today, companies can be held accountable for a lot of outcomes in real life. The data controller should act in accordance with the law, by absorbing inherent data protection principles and allocating an appropriate legal basis into their processing activities. Given the evolving scenario of data processing for the purpose of AI training in social media, it becomes evident that there is still a lot of uncertainty in terms of a widely adopted regulatory approach, as well as enforcement by authorities and the law.

3.4. Brief Comparison between the Regulatory Response in Brazil and in the European Union

Although the GDPR and the LGPD share a common normative foundation grounded in purpose limitation, transparency, and accountability, the social media AI training cases examined in this paper reveal both similarities and differences in how these principles are operationalized in practice. In both jurisdictions, the use of personal data originally collected for social media services to train generative AI models has been treated as a form of further processing, thereby requiring a renewed assessment of purpose compatibility, an applicable legal basis, and adequate safeguards. Regulatory scrutiny in both Brazil and the EU has focused not on the abstract permissibility of AI training as such, but on whether platforms

can demonstrate that such secondary use aligns with users' reasonable expectations and complies with core data protection practices.

With respect to legal basis, enforcement debates in the EU have largely revolved around the compatibility analysis on further processing and the conditions under which legitimate interest may support AI training, particularly when accompanied by robust documentation, balancing test/LIA, and effective objection mechanisms. In Brazil, the ANPD's approach in the Meta case similarly emphasized the coherence between the declared legal basis and the purposes communicated to users, while focusing on transparency failures and the (im)practical accessibility of opt-out mechanisms. While neither regime has categorically indicated that the legitimate interest is the most appropriate legal basis, the cases suggest a shared regulatory concern with preventing its use as a purely formal or abstract justification detached from users' expectations and concrete risk mitigation.

Processing of minors' data constitutes a peculiar aspect. Under the GDPR, children's data benefits from heightened protection, including the special rule on children's consent for information society services where consent is relied upon as the legal basis. Under the LGPD, the best interests are taken into account considering the processing of minor's data. In the Meta's response case, this resulted in the continued suspension of AI training involving minors' data even after adult data processing was conditionally resumed, illustrating a more categorical and outcome-oriented protective stance in the Brazilian context.

Transparency obligations also reveal differences in regulatory emphasis. In the EU the effectiveness of open communication with data subjects played a central role in assessing whether users could form a reasonable expectation regarding AI training. In Brazil, the ANPD attached particular importance to the asymmetry between the platform's internal data reuse strategies and the limited information effectively conveyed to users that led to dark patterns. In both regimes, however, transparency was treated not as a purely formal requirement, but as a substantive condition for the lawful reuse of personal data in AI development.

Finally, enforcement outcomes highlight some differences despite substantive common ground. EU enforcement has been characterized by coordinated yet fragmented interventions involving national authorities, with a focus on the DPC since it has more leverage in this situation from a strategic standpoint (headquarters location), often resulting in pauses, negotiated commitments, and iterative compliance adjustments. In contrast, in Brazil, the ANPD demonstrated a willingness to impose precautionary suspensions, negotiate conformity plans that envisioned compliance on concrete measures.

Overall, these developments indicate that neither the GDPR nor the LGPD treats generative AI training as per se unlawful, but both demand a justification for repurposing social media data. Across jurisdictions, enforcement is converging on a practical test: whether controllers can evidence purpose alignment, a lawful legal basis, meaningful transparency, and effective rights mechanisms, especially where are minors and sensitive data are concerned. As regulators continue

to refine expectations through iterative decisions, the direction points toward a common compliance baseline for AI training.

4. AI Governance in Organizations and the EU AI Act

Data protection legislation helps us to navigate the new landscape of cutting-edge technology spread all around the globe. However, the legislation is not strong enough to hold all ground considering new challenges brought by the advances in AI. Relying on different angles and complementary applications is a must, such as new developing AI regulations or other complementary approaches from data protection authorities worldwide, as examined in both Brazilian and European experiences, which currently are trying to find the best way to proceed in the AI era.

In order to obtain balance between innovation and regulation, where fundamental rights and safeguards are kept without hindering technological progress, an efficient approach in the governance field is necessary. Governance emerges as a viable and reliable solution, not limited to compliance with explicitly stated regulatory requirements, but extending to the operationalization and effective application of these legal instruments.

Legislators themselves must design sufficient instructions so that promising governance structures can emerge as an alternative, such as the EU Artificial Intelligence Act, the EU AI Act (European Union, 2024) with relevant requirements for AI governance (e.g., QMS—Quality Management System) and multiple additional relevant materials set in consultations and guidance that have been released to meet the EU's expectations for this new era, such as the Code of Practice (European Commission, 2025), which xAI has partially signed the chapter of Safety and Security.

There are some essential factors that those involved in AI development and training, and actors within the entire AI value chain, must aim to guarantee principles and best practices embedded in their products and services. Assuring data quality is a requirement brought by law in the EU AI Act, as well as a good practice observed in the market for years. Another evident practice is ensuring accurate documentation, in which is already part of the day-to-day activities of those involved in the technology field, but the law required some additional requirements that would become incorporated into technical documentation (Annex IV, EU AI Act), which would eventually be monitored and audited by authorities.

In the EU context, several of the EU AI Act obligations that are particularly relevant for the social media training scenarios described here include, but not limited to: i) documented data-governance and data-quality practices for training, validation and testing data; ii) technical documentation and record-keeping that supports traceability of design choices and risk controls; and iii) transparency duties that vary according to the role in the AI value chain.

Once again, governance serves the purpose of elucidating how to put legal requirements into practice, such as all obligations regarding assessments, whether

designed specifically for risk, bias, fundamental rights, etc. It is mandatory that companies continue to invest in IA governance (IAPP, 2025), into developing internal and proprietary governance frameworks. To avoid losing momentum, companies should precisely focus on centralizing resources to promote governance measures by having a firm and steady AI strategy.

The cases emerging from influential social media platforms discussed in this paper reflect this challenge in practice, setting a new threshold, especially regarding data protection, which integrates another layer of regulatory complexity.

4.1. Considerations about Social Media Platforms' Cases

It is possible to learn with some of the aspects arising from effective governance to contribute to the adherence of data protection in the AI space, as discussed in the social media cases. The points of concern that were not addressed properly timely were: i) silent opt-in features; ii) inconsistent official and publicly available instruments (privacy policies and terms of use); iii) diffuse information in AI strategies through official communication channels; iv) lack of transparency about further processing; and v) variation in the level of protection of data subjects' rights according to jurisdiction—a direct consequence to tougher regulatory context varying from one territory to another. These factors resulted in a lack of compliance with essential data protection legislation safeguards, such as purpose limitation principle, transparency, adequacy, data minimization, accountability and the lawfulness of a legal basis for processing activities entirely.

In the complaint presented against Meta it was alleged that the company did not implement technical and organizational measures to limit data processing or impact on data subjects' fundamental rights, therefore affecting the data minimization principle. They also failed to observe the necessity principle in the insufficiency of technical measures that would anonymize or pseudonymize personal data (Nyob, 2024c).

Despite the evident flawed conduct in social media platforms' when dealing with data protection, various data protection authorities acted promptly with regulatory and supervisory responses, which lead to the pause in the training of AI using European's personal data. Nevertheless, after these reported events, some inconsistencies began to appear, such as Meta resuming the training without approval from EU bodies and authorities. This raises the possibility that the regulators' efforts were insufficient to resolve the issue.

Regarding the AI regulatory scenario, the EU AI Act does not change the obligations of AI system developers/providers and deployers in their roles as data controllers or processors, nor does it impact data subjects' rights in any way. However, it materially affects how AI actors must organize compliance across the AI value chain. Platforms may simultaneously act as: i) data controllers for user content under data protection law; ii) providers (where they develop or fine-tune a foundation model using platform data for various purposes); and/or iii) deployers when integrating third-party models into the platform. These roles matter because

they will determine obligations and establish responsibilities (e.g., model documentation and transparency expectations, instructions for use, risk-management and post-market monitoring). As AI evolves, it is the responsibility of the AI value chain actors and data processors/controllers to promote transparency measures and provide information about personal data processing in AI projects and solutions, since the recurring obligations of both ends will overlap in a substantial manner.

This mirrors what will be expected from organizations going further in the development and deployment of AI models/systems, in which there is an increasing need to establish a strong baseline for internal governance as a backbone to protect fundamental rights, such as privacy and data protection. There must be efforts to ensure that those do not collide with AI, rather coexist.

4.2. Regulatory Interplay between Data Protection and AI

Using personal data as input for AI technology, given the influence and adherence that social media platforms have in the lives of billions of people worldwide, has become quite attractive for gold diggers. The analysis of the cases involving Meta, X, and LinkedIn highlighted fundamental aspects tied to this new context of processing personal data for the purposes of AI training. Transparency measures and data subjects' rights were conflicted and there was a quick response of data protection authorities in various jurisdictions. After pushing back, effects of this new strategy in the business driven by social media platforms were observed and analyzed with practical implications and conflicting rights.

Therefore, it is extremely relevant to reflect and seek to understand in practice how such cases would be evaluated within the fairly recent binding requirements of the EU AI Act, especially in regard to the AI value chain actors when faced with existing accountability criteria coming from data protection regulations.

The aftermath of the situation is evident, in which the fundamental capabilities to deal with new contexts brought by AI advances are lacking in the regulatory *status quo*. Alternative regulatory proposals are underway, such as the recent proposal and approval of the Digital Omnibus on AI (European Council, 2026) to insert a new specific legal basis to process sensitive personal data involving bias mitigation in AI training within the EU AI Act's scope. Brazil is also voting on their bill for AI regulation in Bill n. 2338/2023 (Brazil, 2023) but up until this stage the amendments do not suggest the inclusion of a new legal basis for AI training data in its scope. Therefore, it is necessary continue to follow new regulatory developments as they unfold to reflect on new possible outcomes to deal with this emerging situation.

Considering the responses examined in Brazil and the EU, this paper argues that social media platforms engaging in generative AI training with user data must adopt governance measures that go beyond compliance and instead operationalize core responsible AI and data protection principles throughout the AI lifecycle. Some interesting takeaways that platforms should adopt as part of their AI gov-

ernance, especially considering the outcomes of the real cases of social media companies and the EU AI Act's compliance logic are: i) platforms should ensure traceability between original collection purposes and AI training objectives through documented purposes for compatibility and lawful analysis; ii) implement robust controls that clearly limit the use of publicly available content while enforcing data minimization and traceability standards; iii) provide meaningful, timely, and intelligible transparency mechanisms that allow users to realistically understand and object to AI training uses of their data beforehand; and iv) adopt heightened safeguards for minors' data and processing of sensitive data, including default exclusions and technical risk-mitigation measures. The implementation of these measures gives an idea of how data protection law and emerging AI governance frameworks can be translated into concrete enterprise-wide obligations in large‑scale AI development.

5. Conclusion

This article sought to address the use of personal data found in social media platforms for generative AI training, highlighting the current challenges and legal implications of this practice. The critical analysis conducted demonstrated that while the use of personal data can provide significant advances in AI technology, it also raises important questions about data protection as part of individuals' fundamental rights concerning their privacy.

The data protection model encompasses significant gaps considering the complexity of the emerging digital ecosystem. Lack of transparency, difficulty in obtaining valid consent, failure to observe all conditions to prove legitimate interest and data subjects' reasonable expectation are some of the few obstacles in the AI era when using personal data for generative AI training. Additionally, the proportionality and purpose limitation principle when processing personal data must be carefully considered to ensure compliance with GDPR and LGPD.

The use of personal data on social media for generative AI training requires in essence a dynamic approach, with the implementation of appropriate regulation, robust governance, and an ethically rooted approach that considers both technological benefits and the protection of fundamental rights.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

Autoridade Nacional de Proteção de Dados (2024a). *Voto n.º 11/2024/DIR-MW/CD. Processo n.º 00261.004509/2024-36. Medida preventiva para evitar dano grave e irreparável ou de difícil reparação.* https://www.gov.br/anpd/pt-br/assuntos/noticias/anpd-determina-suspensao-cautelar-do-tratamento-de-dados-pessoais-para-treinamento-da-ia-da-meta/SEI_0130047_Voto_11.pdf

Autoridade Nacional de Proteção de Dados (2024b). *Despacho decisório n. o 23/2024/DIR-*

JR/CD. Processo n.o 00261.005116/2024-40. https://www.gov.br/anpd/pt-br/assuntos/noticias/SEI_0161130_Despacho_Decisorio_29.pdf

Autoridade Nacional de Proteção de Dados (2024c). *Voto n. o nº 29/2024/FIS/CGF. Processo n.o 00216.004529/2024-36. Medida preventiva. Pedido de reconsideração com efeito suspensivo*. https://www.gov.br/anpd/pt-br/assuntos/deliberacoes-do-conselho-diretor-1/circuitos-deliberativos-2024/cd-18-2024-votos.pdf

Autoridade Nacional de Proteção de Dados (2024d). *Guia orientativo: Hipóteses legais de tratamento de dados pessoais—Legítimo interesse*. https://www.gov.br/anpd/pt-br/centrais-de-conteudo/materiais-educativos-e-publicacoes/guia_legitimo_interesse.pdf

Bray, O. (2024). *X Suspends Personal Data Training of AI Chatbot Grok Following Irish DPC Pressure*. RPC. https://www.rpclegal.com/snapshots/data-protection/autumn-2024/x-suspends-training-of-ai-chatbot-grok-following-irish-dpc-pressure/

Brazil (2018). *Law No. 13.709*. https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm

Brazil (2023). *PL 2338/2023*. https://www.camara.leg.br/proposicoesWeb/fichadetramitacao?idProposicao=2487262

Causin, J. (2024). *X muda regras para usar dados de usuários no treinamento de IA e compartilhar com terceiros*. O Globo. https://oglobo.globo.com/economia/tecnologia/noticia/2024/10/18/x-muda-regras-para-usar-dados-de-usuarios-no-treinamento-de-ia-e-compartilhar-com-terceiros.ghtml

Commission Nationale Informatique & Libertés CNIL (2026a). *The Legal Basis of Legitimate Interests: Focus Sheet on Measures to Implement in Case of Data Collection by Web Scraping*. https://www.cnil.fr/en/legal-basis-legitimate-interests-focus-sheet-measures-implement-case-data-collection-web-scraping

Commission Nationale Informatique & Libertés CNIL (2026b). *The AI System Development: CNIL's Recommendations to Comply with the GDPR*. https://www.cnil.fr/en/ai-system-development-cnils-recommendations-comply-gdpr

Data Protection Commission (2024a). *The DPC Welcomes X's Agreement to Suspend Its Processing of Personal Data for the Purpose of Training AI Tool "Grok"*. https://www.dataprotection.ie/en/news-media/press-releases/dpc-welcomes-xs-agreement-suspend-its-processing-personal-data-purpose-training-ai-tool-grok

Data Protection Commission (2024b). *Data Protection Commission Welcomes Conclusion of Proceedings Relating to X's AI Tool "Grok"*. https://www.dataprotection.ie/en/news-media/press-releases/data-protection-commission-welcomes-conclusion-proceedings-relating-xs-ai-tool-grok

Data Protection Commission (2025). *Data Protection Commission Announces Commencement of Inquiry into X Internet Unlimited Company (XIUC)*. https://www.dataprotection.ie/en/news-media/latest-news/data-protection-commission-announces-commencement-inquiry-x-internet-unlimited-company-xiuc

Digital Policy Alert (2024). *Netherlands: Announced AP Guidelines for Scraping by Private Individuals and Private Organisations*. https://digitalpolicyalert.org/event/19595-announced-ap-guidelines-for-government-scraping

Eidgenössischer Datenschutz und Öffentlichkeitsbeauftragter (2025). *Conclusion of Preliminary Investigation X Formerly Twitter: Use of Personal Data for Training the AI Grok*. https://www.edoeb.admin.ch/en/conclusion-investigation-x-grok

European Commission (2018). *Article 29 Data Protection Working Party. Guidelines on Consent under Regulation 2016/679*. https://ec.europa.eu/newsroom/article29/items/623051

European Commission (2025). *The General-Purpose AI Code of Practice*. https://digital-strategy.ec.europa.eu/en/policies/contents-code-gpai

European Council (2026). Artificial Intelligence: Council and Parliament Agree to Simplify and Streamline Rules. https://www.consilium.europa.eu/en/press/press-releases/2026/05/07/artificial-intelligence-council-and-parliament-agree-to-simplify-and-streamline-rules/

European Data Protection Board (2024). *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/2024-12/edpb_opinion_202428_ai-models_en.pdf

European Union (2016). *Regulation (EU) 2016/679. General Data Protection Regulation*. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679

European Union (2024). *Regulation (EU) 2024/1689. Artificial Intelligence Act*. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L_202401689

Hacker, P., & Holweg, M. (2025) The Regulation of Fine-Tuning: Federated Compliance for Modified General-Purpose AI Models. *SSRN*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5289125

IAPP (2025). *AI Governance Profession Report*. https://iapp.org/media/pdf/resource_center/ai_governance_profession_report_2025.pdf

LinkedIn (2024). *Updates to LinkedIn's Terms of Service*. https://www.linkedin.com/blog/member/platform-information/updates-to-our-terms-of-service-2024

LinkedIn (2025a). *LinkedIn and Generative AI (GAI) FAQs*. https://www.linkedin.com/help/linkedin/answer/a5538339#:~:text=As%20with%20most%20features%20on,services%20(see%20Section%202)

LinkedIn (2025b). *Privacy Policy Effective November 3, 2025*. https://www.linkedin.com/legal/privacy-policy

LinkedIn (2025c). *Control Whether LinkedIn Uses Your Data to Train Generative AI Models That Are Used for Content Creation on LinkedIn*. https://www.linkedin.com/help/linkedin/answer/a6278444

McMahon, L. (2025). *Hundreds of Thousands of Grok Chats Exposed in Google Results*. BBC. https://www.bbc.com/news/articles/cdrkmk00jy0o

Meta (2024a). *What Is the Privacy Policy and What Does It Cover? Effective November 14, 2024*. https://www.facebook.com/privacy/policy/version/8810742435690564/

Meta (2024b). *What Is the Privacy Policy and What Does It Cover? Effective July 9, 2024*. https://www.facebook.com/privacy/policy/version/8002516216460537/

Meta (2025). *How Meta Uses Information for Generative AI Models and Features*. https://www.facebook.com/privacy/genai

Milmo, D. (2024). Elon Musk's X under Pressure from Regulators over Data Harvesting

for Grok AI. *The Guardian.* https://www.theguardian.com/technology/article/2024/jul/26/elon-musks-x-under-pressure-from-regulators-over-data-harvesting-for-grok-ai

Nyob (2024a). *Noyb Urges 11 DPAs to Immediately Stop Meta's Abuse of Personal Data for AI.* https://noyb.eu/en/noyb-urges-11-dpas-immediately-stop-metas-abuse-personal-data-ai

Nyob (2024b). *Twitter's AI Plans Hit with 9 More GDPR Complaints.* https://noyb.eu/en/twitters-ai-plans-hit-9-more-gdpr-complaints

Nyob (2024c). *Case C081-11. Complaint against Meta Platforms Ireland Limited.* https://noyb.eu/files/meta_ai/complaint_ie.pdf

Nyob (2024d). *Case C087-06. Complaint against Twitter International Unlimited Company.* https://noyb.eu/sites/default/files/2024-08/IE_Twitter_AI_bk.pdf

Nyob (2025). *Nyob Sends Meta "Cease and Desist" Letter over AI Training. European Class Action as Potential Next Step.* https://noyb.eu/en/noyb-sends-meta-cease-and-desist-letter-over-ai-training-european-class-action-potential-next-step

Passarella, E. (2024). Elon Musk Asked People to Upload Their Health Data. X Users Obliged. *The New York Times.* https://www.nytimes.com/2024/11/18/well/x-grok-health-privacy.html

Solove, D. J. (2024). Artificial Intelligence and Privacy. *SSRN Electronic Journal.* https://doi.org/10.2139/ssrn.4713111

Twitter (2020). *Privacy Policy Effective January 1st, 2020.* https://x.com/pt/privacy/previous/version_15

X (2024). *Privacy Policy Effective November 15, 2024.* https://x.com/en/privacy