noyb – European Center for Digital Rights
Goldschlagstraße 172/4/3/2
1140 Vienna
AUSTRIA

To:
Data Protection Commission
21 Fitzwilliam Square South
Dublin 2
IRELAND

Vienna, 12 August 2024

noyb case number:	**C087-06**
Complainant:	[redacted] (hereinafter: "the complainant")
Represented under Article 80(1) GDPR by:	***noyb* – European Center for Digital Rights** Goldschlagstraße 172/4/3/2, 1140 Vienna, Austria
Respondent:	**Twitter International Unlimited Company** One Cumberland Place, Fenian Street Dublin 2, D02 AX07, Ireland (hereinafter: "Twitter" or "the controller")
Regarding:	The use of personal data for undefined forms of "machine learning or artificial intelligence models" and the consequent violation of Articles 5(1) and (2), 6(1), 6(4), 9(1), 12(1) and (2), 13(1) and (2), 17(1)(c), 18(1)(d), 19, 21(1) and 25 GDPR

COMPLAINT

INTRODUCTION

We file the following complaint, being aware of the recent action of the Irish Data Protection Commission (hereinafter: DPC) (see Irish High Court case number H.MCA.2024.0000411), which was necessary because the DPC lacks the power to directly order a suspension of processing under Article 58 GDPR. The first public hearing of this application on 08.08.2024 revealed the following background:

- Since September 2023, Twitter and the DPC were engaged in consultations pursuant to Article 36 GDPR, as Twitter itself concluded that the processing of user data for training AI systems would constitute a "*high risk".*
- Twitter started the processing on 07.05.2024, without any public information or notice to data subjects. Based on the exchanges before the Irish High Court, it seems that the DPC was also not informed of the commencement of the processing, despite a pending procedure under Article 36 GDPR.
- The DPC demanded what the parties called "*enhanced mitigation procedures*" before the Irish High Court, which were implemented on 16.07.2024. We assume that these "*enhanced mitigation procedures*" are the opt-out button that Twitter implemented (see 3.2.1 below). There is no other such "mitigation" measure that we are aware of.
- The exchange before the Irish High Court revealed that these "*enhanced mitigation procedures*" failed due to undefined "*technical issues*". This means Twitter did not manage to implement even the mitigation features they agreed upon with the DPC. It is unclear what these "*technical issues*" were.
- Based on the fact that this procedure was brought as an urgent application during the Irish court summer recess, we assume that Twitter has only recently informed the DPC that the processing was actually already ongoing.
- On 08.08.2024, the DPC and Twitter agreed to an undertaking that any further processing of EU personal data (beyond the storage of personal data) for AI training purposed is paused. It is unclear if personal data that was already ingested into the systems will be covered by this undertaking and how the differentiation between EU and non-EU data will be effectively implemented (see 3.4.3. below).
- We are aware that the DPC has issued an urgency procedure under Article 66 GDPR, or is committed to do so soon.

In relation to this complaint, we want to highlight that the pending procedure before the DPC seems to only cover unlawful actions in procedure under Article 36 GDPR.

Given that the DPC has "*negotiated*" the implementation of mitigation measures under Article 6(1)(f) GDPR, it seems to generally accept that the processing may fall under Article 6(1)(f) GDPR, which we fundamentally reject. We also note that there is no indication that all other elements brought up in this complaint are covered by the current actions of the DPC.

Furthermore, we note that the DPC has not taken long-term action but only agreed to an "undertaking" with Twitter, which means that none of the reliefs sought in this complaint are currently implemented.

We therefore believe that the following complaint is not consumed by existing litigation and procedures of the Irish DPC.

1. OVERVIEW

Since 07.05.2024, Twitter International Unlimited Company (hereinafter "Twitter" or "the controller") introduced a new default on its platform "X" to irreversibly ingest the entire data sets of more than 60 million EU/EEA data subjects[1] for undefined "machine learning or artificial intelligence models", without specifying the purposes of such systems. We see the urgent need to file this complaint.

Twitter appears to violate at least Articles 5(1) and (2), 6(1) 6(4), 9(1), 12(1) and (2), 13 (1) and (2), 17(1)(c), 18(1)(d), 19, 21(1) and 25 GDPR. At its core, this complaint relies on the following elements:

- *First*, Twitter has **no legitimate interest** under Article 6(1)(f) GDPR that would override the interest of the complainant (or any data subject) and no other legal basis to process such vast amounts of personal data for undefined purposes.
- *Second*, Twitter unlawfully assumed permission to process personal data for **undefined, broad technical means** ("*machine learning or artificial intelligence models*") **without specifying the purpose** of the processing under Article 5(1)(b) GDPR.
- *Third*, Twitter has taken steps to **deter data subjects from exercising their right to choose** by pretending that data subjects only enjoy a right to object ("*opt-out*") instead of relying on consent ("*opt-in*") and by deterring users from objecting under Article 21 GDPR.
- Fourth, Twitter **fails to provide the necessary** "*concise, transparent, intelligible and easily accessible*" **information**, "*using clear and plain language*".
- *Fifth*, Twitter is **highly unlikely to properly differentiate** (i.) between data subjects where it can rely on a legal basis to process personal data and other data subjects where such a legal basis does not exist and (ii.) between personal data that falls under Article 9 GDPR and other data that does not.
- *Sixth*, the **processing of personal data is highly likely to be irreversible** and thus Twitter is unable to comply with the right to be forgotten once personal data of the complainant is ingested into (unspecified) "*machine learning or artificial intelligence models*".

As a consequence, and given that the processing of the complainant's personal data has **already started and cannot be reversed**, we apply (see section 5 below) that you take (among others) the following urgent action:

- *First*, **immediately issue an urgency decision under Article 66 GDPR** to stop the processing of the personal data of the complainant and over 60 million EU/EEA X users without consent.
- *Second,* **fully investigate the matter** under Article 58(1) GDPR.
- *Third*, **prohibit the use of personal data for undefined "*machine learning or artificial intelligence models*"** without the opt-in consent form the complainant – and indeed other data subjects.

We note that the DPC has itself mentioned that it is "*surprised*"[2] about the steps taken by Twitter and concurs with the urgency of this case.

[1] https://transparency.x.com/en/reports/amars-in-the-eu (accessed on July 29th 2024).
[2] https://www.irishexaminer.com/news/arid-41444617.html (accessed on July 29th 2024).

2. REPRESENTATION

noyb – European Center for Digital Rights is a not-for-profit organisation active in the field of the protection of data subjects' rights and freedoms with its registered office in Goldschlagstraße 172/4/2, 1140 Vienna, Austria, registry number ZVR: 1354838270 (hereinafter: *"noyb")* (**Annex 1**). *noyb* is representing the complainant under Article 80(1) GDPR (**Annex 2**).

3. FACTS OF THE CASE

The following is a brief summary of facts at the time of the filing of this case. These facts may be supplemented by additional information that may arise during the next weeks and in the course of the investigation:

3.1. New Privacy Policy

On 29.09.2023, Twitter updated its privacy policy, claiming that it has the right to use any data related to a user's engagement on its platform, X, to train its AI models and that users grant Twitter a worldwide royalty-free license for such content and personal data. The new privacy policy contains only one mention of artificial intelligence, stating:

> *"We may use the information we collect and publicly available information to help train our machine learning or artificial intelligence models for the purposes outlined in this policy."*[3]

In a separate document called "*Additional information about data processing*"[4] Twitter explains that it relies on Article 6(1)(f) GDPR and publishes the following "analysis" of their overriding legitimate interests that allegedly outweigh the Fundamental Right to Data Protection under Article 8 of the Charter:

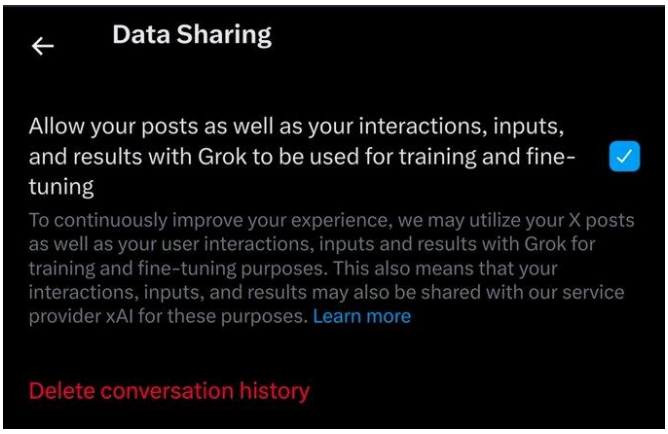
> *"Legitimate interests analysis summary – processing public post data to train machine learning and artificial intelligence models, including generative models*
>
> *X may use information that individuals provide and data that it receives (as described in X's Privacy Policy) to train machine learning and artificial intelligence models, including generative models. This includes public X posts and associated metadata of X users. This helps X offer better services, including summaries of search results and content. Without this training and processing, people would not have access to a large range of information, opinions, viewpoints and accurate summaries and X would have a more difficult time providing relevant, accurate and appropriate responses. To safeguard the rights of those who use our services, users can easily "protect" (limit to a followers-only audience) their posts, or delete their posts at any time, thereby removing their posts and related metadata from being used. X also provides information and user controls to enable X users to opt out of their public post data being used to train an underlying generative model."*

[3] https://x.com/en/privacy (accessed on July 29th 2024).
[4] https://help.x.com/en/rules-and-policies/data-processing-legal-bases (accessed on July 29th 2024).

3.2. New setting on the X website

3.2.1. New interface

Without any announcement or information, sometime in July 2024, Twitter activated a new default setting in the X web user interface that ingests all user posts, interactions, inputs and results to train AI systems. Data subjects can, according to the setting, "*allow*" that "*posts as well as your user interactions, inputs, and results with Grok to be used for training and fine-tuning*."



Screenshot of X's default 'allowance' for Grok AI

It should be noted that default data sharing setting does not just ingest X users' personal data for the purpose of training and developing Grok. These extensive personal data may also be shared with xAI, a separate Elon Musk-led company "*working on building artificial intelligence*" that includes but is not limited to Grok. Indeed, xAI has developed AI with other entities as well (including Open AI GPT-3.5 and 4).[5]

3.2.2. Opt-out only on the 7th (!) step

Twitter has done everything to ensure that data subjects will not change the default setting. X users are only able to opt out by following these steps:

1. Log into X
2. Click on '*More*' in the menu
3. Click on '*Settings and privacy*'
4. Click on '*Privacy and Safety*'
5. Scroll down to '*Data sharing and personalization*'
6. Click on '*Grok*'
7. Untick the box '*Allow your posts as well as your interactions, inputs, and results with Grok to be used for training and fine-tuning*'

It should be noted that initially, the opt-out option was only available in the browser version of X – not in the mobile app.[6] As a result, users could not find this setting if they searched for their privacy features in the app. They also would need to manually log in via a browser, requiring users to find their password (which they would otherwise not need after setup when just opening the app).

[5] https://x.ai/about (accessed on July 29th 2024).
[6] See, e.g.: https://sleonproductions.com/x-activates-a-default-setting-that-gives-it-permission-to-train-grok-ai-on-users-posts-the-setting-can-be-turned-off-on-the-web-but-not-in-the-mobile-app/ (accessed on July 29th 2024).

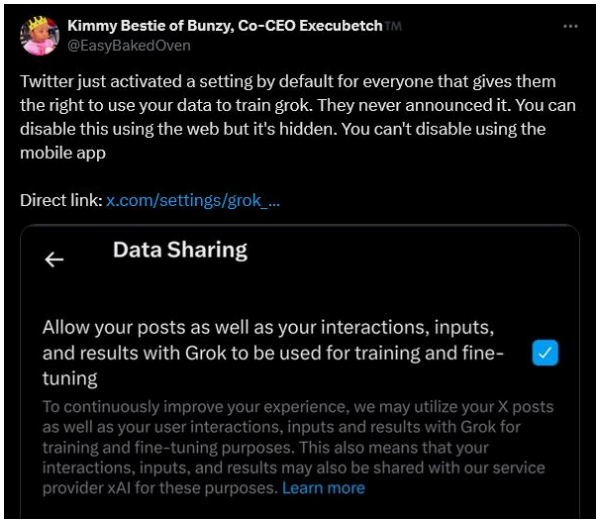
3.2.3. *Twitter "allows" itself to use all personal data*

Data subjects did not "*allow*" such processing for the training of "*machine learning or artificial intelligence models*". Instead, Twitter "allowed" itself to process all relevant personal data, by creating this new setting, pre-ticking it and automatically activating it.

X users were not notified of the new default setting when it was implemented or of any opportunity to opt out of it. The new default setting appears to have taken immediate effect.

3.2.4. *Active information by "@EasyBakedOven" – not by Twitter*

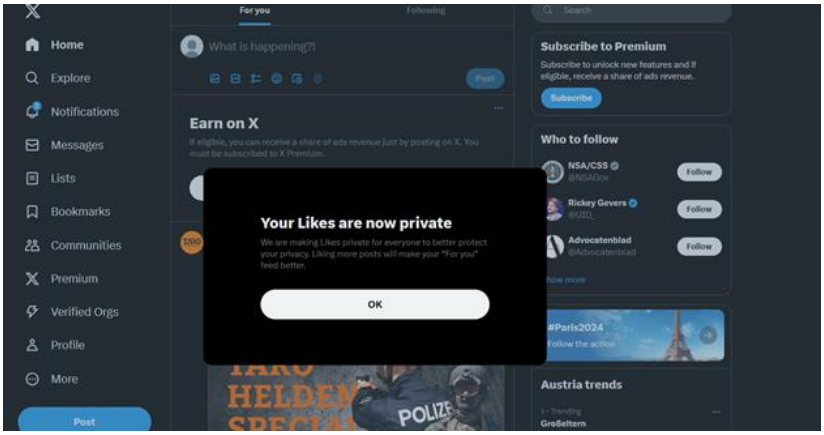
Twitter did not proactively inform users about the fact that all their personal data is being to train AI models. It seems most X users only found out about the new default setting via a "viral" post by an X user named '@EasyBakedOven' on 26.07.2024 – more than two months after Twitter had actually commenced the processing.



Screenshot of viral post by X User '@EasyBakedOven'

X users did not receive any email or pop-up about this new default setting or about X's Privacy Policy update on 29.09.2023 that mention the training of their AI.

This is unusual. X users do typically receive notifications of other privacy updates when logging into their account, such as the notification that liked posts are now private and therefore only visible to the user itself (see screenshot below).



Screenshot of notification on privacy updates

3.3. Scope of processing

Twitter's intended processing of personal data is exceptionally broad. It is also highly questionable whether Twitter is able to properly separate personal data that (i.) falls under Article 6(1)(f) GDPR, (ii.) falls under the application of the GDPR and (iii.) falls under successful objection under Article 21 GDPR.

The exact processing is a matter for further investigation by the authorities under Article 58(1) GDPR and the information below is naturally a preliminary summary:

3.3.1. Personal data that undergoes processing is not defined

In its privacy policy, Twitter does not limit the amount or the type of personal data that may be used to train AI systems. The policy does not specify what information that Twitter collects will be used to train its *"artificial intelligence models*", stating broadly that it may use any information it collects. This indicates that private as well as public posts, messages and interactions may be susceptible to such processing.

On a separate webpage page called "*About Grok, Your Humorous AI Search Assistant on X*" linked in the opt-out setting,[7] there is some information on how to opt out of data processing specifically referring to Grok. However, the webpage only specifies information for opting out of Grok. It does not explain how users can exercise their opt-out rights against the processing of their personal data for training of *"artificial intelligence models"* in general, as are referenced in the privacy policy (see 3.1) and in the default data sharing "allowance" setting (see 3.2.1).

➔ *In other words, according to the privacy policy,* ***any data on Twitter platforms*** *and* ***any publicly available data off-Twitter platforms may be used*** *to train "artificial intelligence models".*

➔ *There is no clear indication if* ***opting out of having data used for Grok*** *also means that* ***any other AI model*** *will not be trained based on that personal data.*

3.3.2. No time limit, allowing use of very old personal data

We note that Twitter has not proposed any limitation on how old the training data could be. Twitter seems to try to use its many "dormant" accounts as a source for personal data, when the user may not even be aware of or reacting to (hidden and un-notified) updates to Twitter's settings. This allows Twitter to generate revenue even from data subjects that have not substantially used the service in years ("data recycling"). Such data should usually have been subject to deletion routines under Article 5(1)(e) GDPR, which Twitter has never implemented.

It should also be noted that Twitter's privacy policies in force in 25.05.2018 did not clearly state that deletion of Twitter's account covered the deletion of all personal data related to that account.[8] Consequently, training data may include personal data that Twitter still retains from deleted accounts.

[7] https://help.x.com/en/using-x/about-grok (accessed on July 29th 2024).
[8] https://twitter.com/privacy/previous/version_14 (accessed on July 29th 2024).

3.3.3. No limitation for "specific purposes" as required by Article 5(1)(b)

From Twitter's privacy policy, it is unclear what the specific purpose of processing is for training the controller's "*machine learning or artificial intelligence models*". The privacy policy states:

> *"We may use the information we collect and publicly available information to help train our machine learning or artificial intelligence models for the purposes outlined in this policy."*[9] (emphasis added)

Under the heading *"How We Use Information*", Twitter lists "*five main ways we use information*", listing extremely broad purposes that apply, in general, to all processing:

1. Operate, improve and personalise services
2. Foster safety and security
3. Measure, analyse and make services better
4. Communicate with users about services
5. Research

Although the privacy policy's only mention of "*artificial intelligence models*" is contained under the sub-heading "*Operate, improve and personalize our services*" (suggesting that operation, improvement and personalisation are the purposes for training AI) the reference broadly notes that use of personal data for AI training could be used for *any* of the extremely broad purposes mentioned in the policy.

This could include basically anything. For example, such broad potential 'purposes' could justify any of the following examples of "*artificial intelligence models*":

- An AI system to detect bots, illegal behaviour and the like (*security*)
- An AI system that allows users to interact and answer questions (*improve services*)
- An AI system to help improve uploaded pictures by users (*improve services*)
- An AI system allowing you to search users within the platform using a photograph (*improve services*)
- An AI system to help find more relevant information in the newsfeed (*personalise services*)
- An AI system to allow advertisers to exploit users' weaknesses (*operate services*)
- An AI system to allow political parties to influence elections (*personalise services*)
- An AI system to allow the finding of potential future criminals using a platform (*security*)

➔ *Overall, the privacy policy's definition of purposes is circular and wholly unclear to data subjects seeking to understand how and why their personal data is processed.*

3.3.4. No anonymisation or pseudonymisation of personal data

We note that Twitter does not even claim to foresee that personal data is minimised or limited in any way, shape or form.

Most notably, the GDPR usually foresees processes like anonymisation or (at least) pseudonymisation as approaches to implement requirements under Article 5(1)(c) GDPR or to comply with the duty to have "*data protection by design and by default*".

None of Twitter's webpages contain any hint, let alone clear legal undertaking, in that direction.

[9] https://x.com/en/privacy (accessed on July 29th 2024).

3.3.5. Summary: No limitation on the processing operations

In summary, Twitter's description of the processing operations foresees none of the typical limitations for the processing of personal data. It seems that Twitter is trying to use the current hype around AI technology and the lack of understanding about it to "slip through" processing operations that would otherwise never be tolerated.

➔ *Twitter foresees the use of any personal data (on Twitter or from a third party), for any purpose, with no time limit and potentially with anyone as the recipient of information from these systems.*

3.4. Foreseeable technical problems in Twitter's implementation

It is clear that the proposed approach by Twitter to have a proper and clear legal basis for any individual piece of information is not achievable in the way Twitter is currently conducting the processing.

3.4.1. Lack of separation between data subjects that agree and/or object

The functioning of a social network, where data is often shared or mixed, would usually mean that any objection would (technically) not apply to data that is not directly linked to an account. This is particularly clear when a user posts personal data about another data subject (who may even not have an X account). The same technical limitation obviously applies to the use of personal data of various users of the service, such as when a user that objected is in a picture that was uploaded by a user that did not object.

Therefore, we struggle to understand how Twitter can separate the personal data of users who opted out from the personal data of other users.

3.4.2. Lack of separation between personal data under Article 6 and 9

Even when it comes to the personal data of a specific data subject, it is unclear whether Twitter can differentiate between personal data falling under Article 6 GDPR and so-called "sensitive" data, which is protected by Article 9 GDPR.

Since other social media platforms such as Meta have long maintained that it is technically impossible to differentiate between data falling under Article 9 GDPR and other personal data,[10] it is extremely unlikely that Twitter can properly distinguish between them when user data is used to train an AI model. The same problem also applies to personal data covered by Article 10 GDPR.

As is explained in more detail below, Article 9 GDPR does not foresee the use of special categories of personal data for "*legitimate interests*", but such personal data would nevertheless be used to train Twitter's AI systems under the same legal basis too.

[10] Meta is currently facing litigation before the CJEU in C-446/21 Schrems, where Meta has submitted that it "does not separate" between special categories of data in accordance with Article 9 GDPR and other categories of data and would therefore be unable to comply with Article 9 GDPR.

3.4.3. Lack of separation between EU and non-EU personal data

It is also unclear how Twitter would separate personal data falling under the territorial scope of Article 3 GDPR from personal data of others. In addition to the fact that joint data (e.g. a US X user mentions an EU X user in a post or a combined picture) can be hard to separate, it also seems from the privacy policy that the joint processing operation has two separate controllers (Twitter in Ireland and "X Corp."), which would suggest a joint controllership of both legal entities.

3.5. Personal data cannot be "forgotten" from an AI system

As is already apparent from other artificial intelligence systems like Large Language Models that are based on artificial neural networks (see, e.g., the *noyb* complaint on OpenAI),[11] personal data that is once entered into an AI system (according to the controllers) cannot be "unlearned", "forgotten", deleted or rectified.

It therefore seems likely that an "objection" now, after the default setting to allow all posts, interactions, inputs and results to be used for training of Grok has already been activated on all X accounts, will not have the effect that personal data is not processed within the LLM anymore – contrary to the obligations under Article 17 GDPR ("*right to be forgotten*"). This irreversible approach by controllers is not just a violation of the GDPR, but an additional factor that gravely undermines the rights and freedoms of data subjects.

Twitter does not acknowledge this irreversibility anywhere. Instead, on webpages once[12] and twice[13] removed from the privacy policy, it baselessly assures data subjects that they can "*easily*" prevent their data from being used by setting posts to private, deleting posts, or opting out of the default setting to share data for training.

[11] See e.g. https://noyb.eu/sites/default/files/2024-04/OpenAI%20Complaint_EN_redacted.pdf
[12] https://help.x.com/en/rules-and-policies/data-processing-legal-bases (accessed on July 29th 2024).
[13] https://help.x.com/en/using-x/about-grok (accessed on July 29th 2024).

4. VIOLATIONS OF THE GDPR

4.1. Violations of Article 5 GDPR

The approach by Twitter violates Article 5 GDPR. Given the need to take many of these factors into account when assessing the legality of processing under Article 6(1)(f) GDPR, these violations also reflect back on the lack of a "legitimate interest" (detailed later in 4.2 to 4.5.11):

4.1.1. Fairness and transparency under Article 5(1)(a)

The use of data subjects' personal data to train AI was intransparent in every respect. As detailed in 4.5.8 and 4.5.9, data subjects using X (and formerly Twitter) were unaware that their data would at some point be used to train boundless "*machine learning or artificial intelligence models*". Data subjects could not have expected that their data would be used for such processing.

Data subjects were not notified when the privacy policy was updated to include information about AI processing. What's more, data subjects were not notified when Twitter introduced the new default setting ingesting all personal data on X to train Grok and other xAI "*machine learning or artificial intelligence models*." As a result, data subjects could not opt out before the processing occurred. Twitter also ensures that the opt-out is not prominently displayed to users. First, as noted in section 3.2.2, the pre-ticked box "allowing" personal data to be used to train AI was initially not accessible from the X app. Instead, users could only access it by logging into X on their browsers. In addition, once you are logged in, the opt-out takes six additional steps to access at all.

These deliberate choices to prevent data subject awareness of processing and to minimise opt-outs rates are clearly not "*fair*" and infringe Article 5(1)(a)'s principle of fairness. The lack of proper information under Article 12 and 13 GDPR (see below) also leads to a violation of the transparency requirement in Article 5(1)(a) GDPR.

4.1.2. Purpose limitation under Article 5(1)(b) and 6(4)

Article 5(1)(b) GDPR clearly states that personal data need to be collected for "specified" purposes. As already highlighted under 3.3.3 above, Twitter does not name any "specific purpose" for the processing of personal data via "*machine learning and artificial intelligence models*" but instead cites extremely broad purposes of operating, improving and personalising their services.

- If according to the former Working Party 29, the purpose of "*improving users experience*", "*marketing purposes*", "*IT-security purposes*" or "*future research*" are all purposes that (without more detail) are too vague or general and do "*not meet the criteria of being 'specific'*", how can Twitter's purposes "*to offer better services*" or "*to operate, improve and personalize our services*" be considered "specified" purposes"?[14]
- If according to the EDPB in its Binding Decision 5/2022, an average user cannot fully grasp what is meant by processing for service improvement where a company's contract lacks clarity, how can X users fully grasp what is meant by processing "*to offer better services*" or "*to operate, improve and personalize our services*" with no further detail?[15]

[14] Article 29 Working Party, Opinion 03/2013 on purpose limitation, 2 April 2013, 00569/13/EN, WP203, p. 16
[15] EDPB Binding Decision 5/2022 on the dispute submitted by the Irish SA regarding WhatsApp Ireland Limited (Art. 65 GDPR), paras 111 and 114.

- If according to the EDPB in its Guidelines on processing of personal data through video devices "*[v]ideo surveillance based on the mere purpose of 'safety' or 'for your safety' is not sufficiently specific*", how can Twitter's purposes "*to offer better services*" or "*to operate, improve and personalize our services*" be "sufficiently specific"?[16]
- If according to the EDPS in its decision on the investigation into the European Commission's use of Microsoft 365, the purposes "*to provide an online service*", including the "*ongoing improvement*", and in particular of "*making improvements to user productivity*", to "*quality*" and to "*efficacy*" cannot be considered specified purposes, how can Twitter's purposes "*to offer better services*" or "*to operate, improve and personalize our services*" be considered "specified purposes"?[17]

It is clear that Twitter's stated purposes "*to offer better services*" or "*to operate, improve and personalize our services*" cannot therefore not be seen as sufficiently specified for the processing operation. Further, given that the privacy policy's sole mention of processing of data for training "*artificial intelligence models*" notes that it may be based on any purpose mentioned in the policy, Twitter's AI could be used for wholly unrelated other purposes (see examples above under 3.3.3).

Under the criteria listed in Article 6(4) GDPR, it is clear that the processing of personal data shared by Twitter's users for such broad and unspecified purposes is not compatible with its initial purpose, which is the provision a social network:

- There is no link between this initial purpose and the purpose of the intended further processing. Twitter's envisioned use of personal data for the training of AI models is not due to any link with the initial purpose, but rather arises from the fact that such training needs large amounts of data and Twitter happens to possess large quantities of data that it wants to use.
- The context in which the personal data was collected contradicts the use for the intended further processing. Information was initially shared on Twitter's platform in order to participate in the social network provided by Twitter and share information with certain people. The complainant and certainly also other Twitter users did not anticipate that this information would be used to train AI models for all kind of undetermined future applications.
- The nature of the personal data, in particular the fact that special categories of personal data are processed, also contradicts the compatibility with the processing for training purposes of AI-models.
- The complainant can only speculate on the existence of any appropriate safeguards, but there is no documentation of such safeguards. It will be up to Twitter to demonstrate whether such safeguards are in place. But even the existence of safeguards does not change the fact that overall, the further processing is incompatible with the initial processing.

Since a compatibility test in accordance with Article 6(4) GDPR shows an incompatibility between the initial purpose and the further processing for the training of unspecified future "*artificial intelligence technology*", Twitter could not base the further processing on a legitimate interest (even if there was a legitimate interest which is challenged in this complaint).

Overall, Twitter clearly violates the purpose limitation principle in Article 5(1)(b) GDPR.

---

[16] EDPB, Guidelines 3/2019 on processing of personal data through video devices, 29 January 2020, para 15, p. 9.
[17] EDPS decision on the investigation into the European Commission's use of Microsoft 365, Case 2021-0518, 8 March 2024, para 97, p. 34.

4.1.3. Data minimisation under Article 5(1)(c)

As highlighted under 3.3.1 to 3.3.3, Twitter's privacy policy does not limit the processing of personal data in any way (scope, sources, types of data or time limits). It theoretically permits the use of any "*information we collect*" to be ingested in the AI systems. There is also no limitation via anonymisation, pseudonymisation or other privacy enhancing technologies.

Thereby, Twitter also violates the data minimisation principle in Article 5(1)(c) GDPR.

4.1.4. Accuracy under Article 5(1)(d)

We further note that AI systems still have a very low accuracy rate.[18] While AI generated pictures of people with four fingers may be tolerable, inaccurate information on an individual can lead to serious harms. It is likely that any results that relate to a data subject will regularly produce false results, which will likely violate Article 5(1)(d) GDPR.

4.1.5. Storage limitation under Article 5(1)(e)

As far as the information provided by Twitter goes, it plans to process personal data ingested into its "*machine learning or artificial intelligence models*" indefinitely. No storage limitation period is specified in the privacy policy or elsewhere. This would likely constitute an additional breach of Article 5(1)(e) GDPR.

4.2. The lack of a legal basis under Article 6(1) GDPR

The use of any personal data to train "*machine learning or artificial intelligence models*" is clearly "processing" of personal data under Article 4(2) GDPR, which requires a "legal basis" pursuant to Article 6(1) GDPR, as processing of personal data is by default illegal under the GDPR.

Twitter seems to rely on alleged overriding "*legitimate interests*" under Article 6(1)(f) GDPR to justify the use of personal data (including X posts and user interactions) of over 60 million EU/EEA data subjects.[19]

We are surprised that Twitter is arguing that it has a "legitimate interest" in using all personal data of over 60 million EU/EEA users[20] when the CJEU has recently, explicitly and clearly held in C-252/21 *Bundeskartellamt* that a controller does not even have a "legitimate interest" to use personal data for advertisement.

It seems clear that the bar set by the CJEU would not allow for the irreversible ingestion of their personal data into undefined "*machine learning or artificial intelligence models*" without any purpose limitation and with an undisclosed number of recipients that will be able to access personal data ingested into such a system.

➔ *Given that the CJEU has clearly taken the view that the use of data for personalized advertisement is not a "legitimate interest", it is painfully obvious that the processing of personal data via new means for any purpose (in all likelihood including "personalized advertisement") cannot be legal under Article 6(1)(f) GDPR.*

[18] https://noyb.eu/en/chatgpt-provides-false-information-about-people-and-openai-cant-correct-it
[19] https://transparency.x.com/en/reports/amars-in-the-eu (accessed on July 29th 2024).
[20] https://transparency.x.com/en/reports/amars-in-the-eu (accessed on July 29th 2024).

For the avoidance of doubt, we nevertheless want to highlight each element of the typical 3-step test under Article 6(1)(f) GDPR that Twitter fails:

4.3. Lack of a "legitimate interest" under Article 6(1)(f) GDPR (Step 1)

Twitter must claim and prove to have a "legitimate interest", according to the established 3-step test for Article 6(1)(f) GDPR.[21] In the current case, the analysis is already failing in the first step, as Twitter neither claims – let alone proves – such a legitimate interest:

4.3.1. Twitter relies on "technical means" – not a "legitimate interest"

In a webpage separate from its privacy policy (that is only available in English and not in any other language) Twitter states that it has a legitimate interest in:

> *"processing public post data to train machine learning and artificial intelligence models, including generative models."*[22]

Usually any "legitimate interest" analysis starts with the interest or the aim of the processing activity – in other words, the "purpose" of the processing operation. As noted in 3.3.3 above, Twitter does not name any specific purpose for the processing of personal data via "*machine learning and artificial intelligence models*" but instead cites any "*purposes outlined in this policy*."

These vague justifications for the processing of data, as detailed extensively in 4.1.2 above, cannot constitute a specific purpose under Article 5(1)(b) GDPR. Such unspecified purposes are just as much a legitimate interest as any other means to process personal data (like "*store all data in a database*", "*run a social network*", "*find correlations in your data*" or "*to do Big Data analysis*").

What Twitter is describing is not a legitimate interest, but merely a means (see e.g. Article 4(7) GDPR "*purposes and means*") to achieve various broad and unspecified purposes.

4.3.2. "Legitimate interests" recognised by the GDPR are usually defensive

The examples in Recitals 47 to 49 of the GDPR are predominantly defensive legitimate interests (like network security, information security or preventing fraud). In such cases, the legislator has indicated an openness to recognise the processing of personal data as a "legitimate interest", given that the controller is merely acting in a defensive way.

Instead, Twitter seems to want to offensively use the personal data of over 60 million EU/EEA data subjects[23] to extract profits from (often long abandoned) social media profiles. The GDPR and its recitals do not provide or hint that such processing of personal data could be seen as a legitimate interest.

4.3.3. Making money itself is not a "legitimate interest"

Despite claims to the opposite by controllers, the mere interest in making money is itself not a "legitimate interest", as can be seen from the countless decisions on the sale of personal data, the use for personalized advertisement and the like.[24]

---

[21] CJEU 4 May 2017, C-13/16 (*Rigas*), para. 28.
[22] https://help.x.com/en/rules-and-policies/data-processing-legal-bases (accessed on July 29th 2024).
[23] https://transparency.x.com/en/reports/amars-in-the-eu (accessed on July 29th 2024).
[24] See e.g. https://autoriteitpersoonsgegevens.nl/documenten/ap-normuitleg-grondslag-gerechtvaardigd-belang (accessed on July 29th 2024).

4.3.4. Mere data extraction is itself not a "legitimate interest"

Equally, it is not a legitimate interest to simply buy and collect personal data from third parties ("*data brokerage*") and use internal data for totally unrelated new business ideas.

If the mere extraction of personal data from various systems to support any type of new processing for any undefined purpose were a "legitimate interest", this would literally mean that any controller could use any personal data from any source for any new purpose. This narrative entertained by Twitter is therefore totally outside of the common understanding under the GDPR.

4.3.5. Violation of Articles 5, 12, 13, 17(1)(c), 18, 19, 21(1) and 25 GDPR

As demonstrated above and below (see 4.1 and 4.6 to 4.9) the proposed AI system of Twitter and the way it was introduced clearly violates at least Articles 5(1), 5(2) 12, 13, 17(1)(c), 19, 21(1) and 25 GDPR. The violation of other provisions of the GDPR is another major factor for why any balancing of interests under Article 6(1)(f) GDPR must fail.

An artificial intelligence system that is based on the violation of eight (!) Articles of the GDPR in one go cannot ever be seen as "legitimate".

4.3.6. Inclusion of "sensitive data" under Article 9 GDPR

As stated in 3.4.2 above, it is unclear whether Twitter can differentiate between sensitive data under Article 9 GDPR and other data. We therefore note that Twitter likely lacks the option to rely on a "legitimate interest" as its processing likely includes special categories of data that do not fall under Article 6(1)(f) GDPR and where relying on a "legitimate interest" is simply not available.

4.3.7. Lack of separation between data subjects' personal data

As discussed in section 3.4.1 above, Twitter is likely not in a position to separate personal data of (i.) data subjects that objected and (ii.) personal data relating to data subjects that did not object (and that potentially are not even X users).

This leads to the inevitable conclusion that X users who object could still have some of their data processed when it is uploaded or published by other users. It is thus reasonable to assume that the right to object under Article 21(1) GDPR cannot be fully complied with.

Reliance on legitimate interest as a legal basis always requires compliance with the law, including that the data subject has the right to object. As this is not always possible, or at least not for all data, Twitter cannot use Article 6(1)(f) GDPR as a legal basis for this processing activity.

4.3.8. Summary on the existence of a "legitimate interest"

- ➔ *Twitter does not pursue any legitimate interest recognizable under Article 6(1)(f) GDPR.*
- ➔ *The mere use of a broad category of various technologies constitutes so-called "means" not a legitimate interest in itself.*
- ➔ *Compared to the legitimate interests named in the GDPR or accepted in case-law, the mere extraction of personal data to use for commercial gain is not a "legitimate interest".*
- ➔ *Finally, Twitter tries to process an enormous pool of personal data, which (at least partly) inevitably contains personal data that cannot be processed based on a "legitimate interest".*

4.4. All data for any purpose is not "strictly necessary" processing (Step 2)

Very much overlapping with the principle of data minimisation in Article 5(1)(c) GDPR and the duty to engage in data protection by design and by default in Article 25 GDPR (see below), the second element of the CJEU's legitimate interest test requires that personal data be "*strictly necessary*".

In C-252/21 *Bundeskartellamt* the CJEU held at paragraph 108 that:

> *"…that condition requires the referring court to ascertain that the legitimate data processing interests pursued cannot reasonably be achieved just as effectively by other means less restrictive of the fundamental rights and freedoms of data subjects, in particular the rights to respect for private life and to the protection of personal data guaranteed by Articles 7 and 8 of the Charter…"*

The question is not if the processing would be better, easier or more convenient for the controller, but if it is "strictly necessary" to reach an aim or purpose. It is clear that the "strictly necessary" test must fail for Twitter:

- It should be stressed that assessing the necessity of a certain processing operation is very difficult when the specific purposes are not even disclosed. As stated above, "*machine learning and artificial intelligence models*" are not a purpose but rather a broad group of means of processing. Processing can never be "necessary" to entertain technological "means".
- That being said, whatever the purposes may be, it is highly unlikely that they strictly require the use of *any "information we collect and publicly available information"* of *all* EU/EEA users, without any anonymisation or pseudonymisation measures in place and with no time limit.
- This can also be demonstrated by the fact that many controllers have already developed "*machine learning and artificial intelligence models*" without the use of such vast data sources.
- In addition, it must be noted that the fact that only some types of "*machine learning and artificial intelligence models*" require a large amount of data to be trained does not authorise Twitter to process any data potentially available to them. For example, "Reactive Machines" fall under the definition of "artificial intelligence" and are not based on past experiences to take decisions. It can therefore not logically be "strictly necessary" to use all personal data for any "artificial intelligence technology".
- Finally, Twitter would have the option to limit the processing to persons that actually want to use the Grok AI functions. It is not clear to what extent the use of anyone else's personal data is necessary to provide AI services to a small group of actual users.

➔ *Overall, it seems obvious that Twitter attempts to process personal data far beyond anything that is "strictly necessary" for the (undisclosed) potential purposes.*
➔ *This can also be demonstrated by the many existing AI systems that were trained and run on much smaller dataset.*

4.5. Twitter also cannot overcome the balancing test (Step 3)

Even if Twitter were found to pursue a "legitimate interest" and the processing of (all) personal data it holds on data subjects were considered "strictly necessary", the third level of Article 6(1)(f) – the overall "balancing" test – would also clearly fail for Twitter:

4.5.1. Interpretation in light of Articles 7, 8 and 52(1) of the Charter

Obviously, Article 6(1)(f) GDPR must be interpreted in the light of the Charter, especially as Article 6(1)(f) GDPR has a similar function to the proportionality test in Article 52(1) of the Charter.

- If under C-293/12 *Digital Rights Ireland* (and many following judgements by the CJEU) the "mere" storage of communication metadata for the rather important purpose of national security is not "proportionate", how can the use of (almost) all personal data of over 60 million active EU/EEA users[25] of a social network be "proportionate" to train an AI model with unclear future use?
- If in C-311/18 *Schrems II* the "mere" scanning of traffic data and the access to stored data for national security purposes violates Article 7 and 8 of the Charter, how can the use of all of this data be "proportionate" when training an AI model?
- If in joined cases C-203/15 and C-698/15 *Tele2* the "mere" retention of traffic data and location data for the purpose of fighting crime violates Articles 7 and 8 of the Charter, how can the use of all this data be "proportionate" when training an AI model?

Already in comparison with CJEU case law on Article 7 and 8 of the Charter, it seems apparent that the use of much vaster amounts of personal data, for much more trivial purposes (like generating a "*humorous AI search*" or improving a chat bot) cannot be proportionate under Article 7 and 8 of the Charter and consequently, also cannot be proportionate under Article 6(1)(f) GDPR.

4.5.2. Unlawful initial collection of personal data

Any balancing of interests must already fail, because Twitter had largely no legal basis for the initial collection of large amounts of personal data that it has apparently used for training.

Before the coming into force of the GDPR on 25.05.2018, Twitter relied on consent under Article 7(a) of Directive 95/46. However, this consent was bundled, based on the mere use of the website (no "opt-in") and was clearly far from compliant with Article 4(11) GDPR or Article 7(a) of Directive 95/46/EC.[26] Twitter can therefore not rely on consent obtained from data subjects up until 25.5.2018 for the processing of personal data.

Thereafter, Twitter either did not specify a legal basis or relied on legitimate interest, as it does now, to collect large amounts of the personal data it collects, including information data subjects share with Twitter, additional information Twitter receives about data subjects and inferences Twitter makes about data subjects.[27] In the CJEU judgement C-252/21 *Bundeskartellamt*, paragraph 117 noted that the user of an online social network

> *"cannot reasonably expect that the operator of the social network will process that user's personal data, without his or her consent, for the purposes of personalised advertising."*

Processing of information for such purposes, absent the data subject's consent, is only justified where it meets narrow necessity requirements pursuant to Article 6(1)(b) or (f).[28] Thus, it is

---

[25] https://transparency.x.com/en/reports/amars-in-the-eu (accessed on July 29th 2024).
[26] For example, Twitter Privacy Policy, effective on 14 May 2007 "*By using our Site you are consenting to our processing of your information as set forth in this Privacy Policy now and as amended by us*." (https://x.com/en/privacy/previous/version_1) (accessed on July 29th 2024).. Twitter Privacy Policy, effective on 21 October 2013 "*When using any of our Services you consent to the collection, transfer, manipulation, storage, disclosure and other uses of your information as described in this Privacy Policy*." (https://x.com/en/privacy/previous/version_8) (accessed on July 29th 2024).. Twitter Privacy Policy, effective on 18 June 2017 "*When using any of our Services you consent to the collection, transfer, storage, disclosure, and use of your information as described in this Privacy Policy. This includes any information you choose to provide that is deemed sensitive under applicable law*." (https://x.com/en/privacy/previous/version_13) (accessed on July 29th 2024).
[27] https://help.x.com/en/rules-and-policies/data-processing-legal-bases (accessed on July 29th 2024).
[28] C-252/21 *Bundeskartellamt*, para. 92.

likely that Twitter also did not have a proper legal basis to collect large amounts of the personal data that it has obtained from 25 May 2018 to the present day.

4.5.3. Exceptionally large and unlimited amount of personal data

Furthermore, the personal data that Twitter is processing goes far beyond any "data pool" that is used for similar purposes:

- According to the privacy policy, the processing could concern any personal data collected by Twitter since the complainant signed up to the service – spanning a long time and including personal data that is deleted, archived data and personal data of other users.
- Such information can contain sensitive information revealing political leaning, financial background, sexual orientation, health problems, criminal offences, events that people attended or children's data.
- The processing also concerns online tracking data that Twitter collects on third pages, personal data uploaded by others (individuals and businesses) and the like.
- In 2021, X reported that it processes 400 billion events in real time and generates a Petabyte of data every day.[29]

Compared to typical examples of an overriding "legitimate interest" (e.g. the mere storage of CCTV pictures for a limited space and time or the keeping of an IP address for security reasons), Twitter engages in processing of totally unheard-of dimensions for undefined future purposes.

4.5.4. Non-public personal data

As discussed in 3.3.1 above, it is unclear whether Twitter will use data from private accounts, private messages between X users or private liked posts that are not visible to the general public.

Twitter's privacy policy states broadly that "*we may use the information we collect and publicly available information to help train our machine learning or artificial intelligence models*". This characterisation is expansive and does not exclude any type of personal data. It indicates that any and all use of X, private or public, is potential fodder for AI training.

In C-252/21 *Bundeskartellamt*, the CJEU held at paragraphs 84 and 85 that even information known to the network, is not "fair game" and is generally protected by the GDPR. Data subject intent, the CJEU emphasises, is key:

> *"[...] Article 9(2)(e) of the GDPR must be interpreted as meaning that, where the user of an online social network visits websites or apps to which one or more of the categories set out in Article 9(1) of the GDPR relate, the user does not manifestly make public, within the meaning of the first of those provisions, the data relating to those visits collected by the operator of that online social network via cookies or similar storage technologies.*
>
> *Where he or she enters information into such websites or apps or where he or she clicks or taps on buttons integrated into those sites and apps, such as the 'Like' or 'Share' buttons or buttons enabling the user to identify himself or herself on those sites or apps using login credentials linked to his or her social network user account, his or her telephone number or email address, that user manifestly makes public, within the meaning of Article 9(2)(e), the data thus entered or resulting from the clicking or tapping on those buttons only in the circumstance where he or she has explicitly made the choice*

[29] https://blog.x.com/engineering/en_us/topics/infrastructure/2021/processing-billions-of-events-in-real-time-at-twitter- (accessed on July 29th 2024).

> *beforehand, as the case may be on the basis of individual settings selected with full knowledge of the facts, to make the data relating to him or her publicly accessible to an unlimited number of persons."*

Similar statements can be found in C-362/14 *Schrems I*, C-311/18 *Schrems II* or C-468/10 *ASNEF*, where the CJEU has consistently held that communication data and content data is especially protected. It is obvious that Twitter (operating a "social network") is predominantly using "communication data" and/or "content data" for the relevant processing activities.

4.5.5. High-risk technology with regular problems

In their current state, AI systems are still an unproven and speculative technology. This increases the risks for data subjects in an enormous way. Given that Twitter also does not explain what the AI system will be used for, any product may be used against the interest of a data subject or may produce errors that lead to real-life consequences for the data subject.

This is not just theoretical; these are news headlines of the past year(s). To name just some (of many) examples:

- Microsoft had to turn off an AI chatbot after it "*turned into a Nazi*".[30]
- Google rolled back its AI Search function given countless errors.[31]
- Facebook had to shut down AI bots after they spoke to each other in their own language that was incomprehensible to humans.[32]
- OpenAI had its systems used for phishing and scams.[33]
- California has banned "self-driving" cars, following recurring issues.[34]

The lack of accurate results (see Article 5(1)(d) GDPR) and the overall unclear power and use of such systems makes the complainant fearful of having their own personal data ingested into such a system that may later also be used against the complainant.

The processing of personal data contrary to the interests of the data subject is another major factor that leads to a negative outcome in any balancing test.

4.5.6. No right to object or erase once personal data is used ("No way back")

As noted above at 3.5 above, artificial intelligence models have an 'unlearning' problem. It is widely considered "*virtually impossible to make an AI Model 'forget' the things it learns from private user data*" after it has been trained on such information.[35]

The consequence for "*machine learning or artificial intelligence models*" like Twitter's Grok or other such technologies developed by xAI is that objections to processing can only impact the use of personal data *going forward* — not the use of personal data which has already been ingested by the AI models. Contrary to Articles 17(1)(c), 19 and 21(1) GDPR, this means that while no new

---

[30] https://www.cbsnews.com/news/microsoft-shuts-down-ai-chatbot-after-it-turned-into-racist-nazi/ (accessed on July 29th 2024).
[31] https://www.nytimes.com/2024/06/01/technology/google-ai-overviews-rollback.html (accessed on July 29th 2024).
[32] https://www.firstpost.com/tech/news-analysis/facebook-researchers-shut-down-ai-bots-that-started-speaking-in-a-language-unintelligible-to-humans-3876197.html (accessed on July 29th 2024).
[33] https://tech.co/news/chatgpt-ai-scams-watch-out-avoid#phishing (accessed on July 29th 2024).
[34] https://slate.com/business/2023/10/cruise-suspended-california-robotaxis-self-driving-cars-san-francisco.html (accessed on July 29th 2024).
[35] See: https://fortune.com/europe/2023/08/30/researchers-impossible-remove-private-user-data-delete-trained-ai-models/ (accessed on July 29th 2024); see also: https://www.theregister.com/2019/07/15/ai_delete_data/ (accessed on July 29th 2024).

personal data may be ingested into an AI system, Twitter has no way to delete personal data that its "*artificial intelligence model*" was already trained on. This is the clear opposite of a "*right to be forgotten*", which by definition also requires deletion of previously obtained personal data.

The fact that the use of personal data seems to be (technically) irreversible violates the right to object to any future processing under Article 21 GDPR.

In the Joined Cases C-26/22 and C-64/22 *SCHUFA*, the CJEU has already decided that any processing of (public) personal data must end as soon as the published data is deleted (in this case, within 6 months). Twitter's training approach does not permit the removal of such personal data once it is ingested into the system.

The fact that the processing is allegedly irreversible is another serious factor that would usually tip any balancing test towards a negative outcome.

4.5.7. X has been designated as a VLOP by the European Commission

On 25.04.2023, the European Commission designated X, which boasts over 60 million active users in the EU,[36] a Very Large Online Platform (VLOP) under the Digital Services Act.[37] This is another marker of the immense data collection and processing power that X has over EU/EEA users. It highlights the perilous impact this unlawful processing has on the fundamental right to data protection of millions of EU data subjects whose X data has been unwittingly used to train AI systems.

4.5.8. Typical case of unlimited "secondary processing"

Sometimes the use of personal data for a closely related purpose (e.g. the option to apply an AI filter to an uploaded picture) may be in line with the expectations of a data subject and purposes of the processing.

However, the use of all personal data (no matter the purpose for which it was shared or generated) for an undisclosed future purpose contemplated by Twitter via any form of current or future "*machine learning or artificial intelligence models*" is a typical case of unrelated "secondary processing", which the GDPR explicitly tries to prevent.

As a social media platform designed for sharing user's information within its ecosystem, Twitter is intuitively understood to collect and process personal data primarily for providing the user with the service. This is particularly true for users who set up their accounts in 2007, when Twitter entered the European market. According to its first privacy policy, the primary purpose of data processing was explained as follows:

> "*Our primary goals in collecting personally identifiable information are to provide you with the product and services made available through the Site, including, but not limited, to the Service, to communicate with you, and to manage your registered user account, if you have one.*"[38]

Although Twitter's privacy policies have changed over time, Twitter has consistently underlined that processing activities aimed at providing their services to the user. Under this framework, it

---

[36] https://transparency.x.com/en/reports/amars-in-the-eu (accessed on July 29th 2024).
[37] See press release: https://ec.europa.eu/commission/presscorner/detail/en/ip_23_2413 (accessed on July 29th 2024).
[38] Dated on 14 May 2007.

is apparent that the training of AI systems does not fall within the scope of initial purpose of data processing.

4.5.9. Expectation of data subjects

In using X (formerly Twitter,) data subjects entered agreed to use a service allowing them to share posts, look at cat pictures or chat with friends. Data subjects (who may have signed up years ago) had no expectation that personal data entered into a social network would be used in 2024 to train AI systems with an undefined future purpose.[39]

As the CJEU had held in C-252/21 *Bundeskartellamt* at paragraph 117:

> *"In this regard, it is important to note that, despite the fact that the services of an online social network such as Facebook are free of charge, the user of that network cannot reasonably expect that the operator of the social network will process that user's personal data, without his or her consent, for the purposes of personalised advertising. In those circumstances, it must be held that the interests and fundamental rights of such a user override the interest of that operator in such personalised advertising by which it finances its activity, with the result that the processing by that operator for such purposes cannot fall within the scope of point (f) of the first subparagraph of Article 6(1) of the GDPR."*

Between 2007 and May 2018, Twitter has clearly stated in its privacy policy[40] that the main purpose of Twitter services was to help users "*share information with the world*". The expectation of Twitter users was that processing activities would extend to their intentional sharing of data with others in a social setting — not that their information would be used to train an AI.

Further, as noted in 4.5.8 above, for years, Twitter's privacy policy linked the purpose of services' improvement with existing services, and manifestly indicating a bounded scope of potential improvement.[41] Such a purpose of processing could not reasonably cover the creation and training of an unforeseeable and altogether *new* service or technology like an advanced AI system.

Prior to August 2023 when the privacy policy was updated to include a reference to "*artificial intelligence models*", data subjects had no "reasonable expectation" that their personal data might be processed for training of AI system. The vast majority of X users in the EU — *at least* 60.9 million active users— were on the platform prior to August 2023[42] and thus would have been wholly unaware of such a potential usage.

---

[39] cf. Recital 47 GDPR: "[...] *At any rate the existence of a legitimate interest would need careful assessment including whether a data subject can reasonably expect at the time and in the context of the collection of the personal data that processing for that purpose may take place. The interests and fundamental rights of the data subject could in particular override the interest of the data controller where personal data are processed in circumstances where data subjects do not reasonably expect further processing*. [...]"

[40] Privacy policies applicable to data processing activities before the GDPR came into force.

[41] For example: Twitter Privacy Policy, effective on18 June 2017: "*We may use and store information about your location to provide features of our Services, such as allowing you to Tweet with your location, and to improve and customize the Services, for example, with more relevant content like local trends, stories, ads, and suggestions for people to follow*." - Twitter Privacy Policy, effective on 16 October 2010 "*We do this to help improve our Services, including advertising, and to be able to share aggregate click statistics such as how many times a particular link was clicked on*"

[42] Between July 2023 and August 2024, the number of active X account holders increased by 0.9 million. While we cannot confirm precisely how many users created accounts during this period or reactivated their usage only in the last year, this data indicates that the number of *new* users since August 2023 is extremely low compared to the over 60 million active users prior to that date. It should also be noted that this number only includes *active* users and does not account for a significant number of users who have not been active in a number of years but still have accounts, see: https://transparency.x.com/en/reports/amars-in-the-eu (accessed on July 29th 2024).

4.5.10. Industry standards

While industry standards under the GDPR are often a "low bar" given that many controllers do not comply with the law, most currently known systems (that can already be highly problematic in relation to the GDPR) are trained with dedicated data that was obtained by the controller (e.g. scans of streets for self-driving cars), publicly available information (e.g. web scraping) or are otherwise limited in scope.

The most comparable industry practice is seen with Meta, which announced in May 2024 that within one month it would begin training its artificial intelligence technology using EU data subjects' personal data on Meta platforms. In response to numerous GDPR complaints concerning Meta's plans to use EU/EEA data subjects' posts and engagement on its platforms to train Meta AI, Meta announced that it would not proceed with such training in the EU.[43] Given Meta's decision, we are not aware of any social networks or any other consumer-facing controller to use all available personal data for AI systems.

What's more, Twitter at no point announced that it would use EU data subjects' personal data to train Grok or other xAI "*machine learning or artificial intelligence models*". It did not provide EU data subjects an opportunity to opt out prior to using their data to train such AIs. Instead, Twitter began training without notice to data subjects, providing only a hidden (see 3.2.2 above) and after-the-fact opt-out that does not permit data subjects to retroactively erase personal data used from the AI (see 4.5.6 above).

4.5.11. Twitter fails the overall balancing test

Given the initial unlawful collection of personal data, the exceptionally large and unlimited amount of personal data (apparently including non-public data), the highly risky nature of the technology involved, the impossibility to object or erase once one's data is has already been used, the disproportionate market power that Twitter exercises over its users, the existence of a further processing clearly unrelated to the original one, a scope of processing well beyond the expectations of the data subject and even a lack of compliance with (minimal) industry standards, Twitter fails the balancing test and consequently cannot rely on legitimate interest under Article 6(1)(f) GDPR.

4.6. Violation of Article 12 GDPR

Twitter does not provide "*concise, transparent, intelligible and easily accessible*" information according to Article 12 GDPR, nor does it inform the complainant in "*clear and plain language*". On the contrary, Twitter conceals relevant information and rights, as highlighted in section 4.1.1 of this complaint.

Furthermore, as discussed in 3.2.2 above, Twitter seeks to deter data subjects from exercising their rights by adopting a complex procedure instead of a "one-click" objection. The user has to go through 7 different steps in order to simply submit an objection. In doing so, Twitter acts in violation of Article 12(2), which requires the controllers to "*facilitate the exercise of data subject* rights".

[43] https://www.theguardian.com/technology/article/2024/jul/18/meta-release-advanced-ai-multimodal-llama-model-eu-facebook-owner (accessed on July 29th 2024).

4.7. Violation of Article 13 GDPR

As is already apparent under 3.1 above, Twitter's new privacy policy violates Article 13 GDPR by failing to include several elements of this Article, as follows:

- Twitter fails to inform the complainant of the exact purpose of processing, instead referring broadly to a number of vague purposes that could be used to justify virtually any type of "*artificial intelligence model*". However, the disclosure of the specific purposes is obligatory under Article 13(1)(c) GDPR.
- Twitter should have informed data subjects about the claimed legitimate interest it pursued in the processing, according to Article 13(1)(d) GDPR. Instead, Twitter only informs data subjects about the technical means ("*machine learning and artificial intelligence models*").
- Twitter's privacy policy does not provide any information on the duration of the processing nor on the criteria used to determine it, as mentioned in section 4.1.5 above of the complaint, therefore violating Article 13(2)(a) GDPR.

Therefore, Twitter acts in violation of multiple elements of Article 13 GDPR.

4.8. Violation of Articles 17(1)(c), 19 and 21(1) GDPR

As shown at 3.5 above, any objection or other finding that personal data is processed without a legal basis would not lead to the end of processing within an artificial intelligence system when data has already been ingested. This is contrary to the "right to be forgotten" and limits the rights of data subjects under Articles 17 and 19 GDPR as well as under Article 21(1) GDPR to a mere "*right to not have even more data processed*".

4.9. Violation of Article 25 GDPR

From the documentation that was provided by Twitter, it appears evident that Twitter has not entertained any technical and organisational measures to:

- limit the processing of personal data or the impact on the fundamental rights of data subjects (such as an opt-in system or clear controls for data subjects),
- implement an approach of data minimisation in practice,
- limit the processing only to strictly "necessary" personal data,
- limit the processing to anonymised or pseudonymised personal data,

or indeed any other publicly available and enforceable measure. By failing to do so, Twitter has also violated its duties under Article 25 GDPR ("data protection by design and default") when simply declaring the personal data of all its users worldwide to be the "new oil" for any future AI machine.

5. APPLICATIONS

Based on the above facts and law, and indeed any other facts or legal arguments that may arise during the procedure, we make the following applications:

5.1. Duty to act

The CJEU has repeatedly held that supervisory authorities have a positive duty to act if they are made aware of a GDPR violation. In C-311/18 *Schrems II* the CJEU held at paragraph 111:

> *"In order to handle complaints lodged, Article 58(1) of the GDPR confers extensive investigative powers on each supervisory authority. If a supervisory authority takes the view, following an investigation, that a data subject whose personal data have been transferred to a third country is not afforded an adequate level of protection in that country, it is required, under EU law, to take appropriate action in order to remedy any findings of inadequacy, irrespective of the reason for, or nature of, that inadequacy. To that effect, Article 58(2) of that regulation lists the various corrective powers which the supervisory authority may adopt."*

In the Joint Cases C-26/22 and C-64/22 *SCHUFA* the CJEU has further highlighted at paragraph 57:

> *"In order to handle complaints lodged, Article 58(1) of the GDPR confers extensive investigative powers on each supervisory authority. Where, following its investigation, such an authority finds an infringement of the provisions of that regulation, it is required to react appropriately in order to remedy the shortcoming found. To that end, Article 58(2) of that regulation lists the various corrective measures that the supervisory authority may adopt."*

In C-768/21 *Land Hessen*, the AG has further issued an opinion saying at paragraph 82:

> *"[...] that the supervisory authority has an obligation to act when it finds a personal data breach in the course of investigating a complaint. In particular, it is required to define the most appropriate corrective measure(s) to remedy the infringement and ensure that the data subject's rights are respected.* [...]*"*

An equal result can be derived from the general duty of public authorities to uphold fundamental rights - like the right to data protection in Article 8 of the Charter. There is consequently no question that the DPC has a duty to act in this case.

5.2. Investigation under Article 58(1) GDPR

Given that some of the details of Twitter's processing are unclear, we hereby apply for a full investigation using all powers under Article 58(1) GDPR, which should at least include the following steps:

- Clarification of the concrete "artificial intelligence technology" that will be used.
- Clarification of the personal data that will be ingested into such systems.
- Clarification on how Twitter intends to separate EU/EEA personal data, data falling under Article 9 GDPR and data for which users have exercised choice (opt-in or opt-out) from data of data subjects that have taken the opposite decision.
- Clarification on the options to exercise the "right to be forgotten" under Article 17 GDPR, but also other GDPR rights (like the right to access or rectification) once personal data is ingested into such systems.
- Demanding any "legitimate interest" assessment that Twitter may have conducted under Article 6(1)(f) GDPR.

- Demanding the record of processing activities under Article 30 GDPR.
- Demanding the documentation of any Data Protection Impact Assessment under Article 35 GDPR that Twitter should have produced on these systems.

5.3. Preliminary stop of the processing activities under Article 58(2) GDPR

Given the exceptional circumstances of this case (see below), we apply to have a preliminary stop of any processing activities enforced under Article 58(2) GDPR in parallel with the "Urgency Procedure" under Article 66 GDPR.

5.3.1. The conditions required by Article 66(1) GDPR are met

As outlined under 3.2 above, Twitter has already started using the complainant's personal data for some types of AI technology. This means that personal data of the data subject and of more than 60 million affected people[44] are being processed to train Twitter's AI technology. This processing, which constitutes an "exceptional circumstance", is unlawful as stated in the previous section.

As detailed throughout this complaint, Twitter's further processing has and can continue to seriously impair the data subjects' rights and freedoms. It is thus urgently necessary and appropriate to immediately stop any further use of personal data of over 60 million people in the EU/EEA[45] until the matters raised in this complaint are sufficiently investigated and decided.

5.3.2. No imminent threat to Twitter & limitation to three months

On the other hand, a preliminary halt of processing activities would merely amount to a "delay" of the processing operations - if the supervisory authorities may (opposite to any suggestion in the case law) later take a view that the approach by Twitter was in fact legal.

According to Article 66(1) GDPR, any urgency action is also limited to three months, which would allow Twitter to explain how this approach is legal.

5.3.3. The urgency procedure is available to every DPA, including the LSA

Finally, it should be noted that the urgency procedure under Article 66 GDPR is available to every supervisory authority and also to the lead supervisory authority (LSA).

Therefore, it follows from the above and from each DPA's duty to act (see above 5.1) that the DPC should immediately adopt provisional measures.

5.4. Corrective powers under Article 58(2) GDPR

Even before any investigation may have come to a final conclusion, we urge the authority to take imminent, preliminary steps to ensure that Twitter does not pursue the processing operations any further, including but not limited to:

- Immediately issue a warning under Article 58(2)(a) GDPR, highlighting the unlawfulness of the intended processing.

[44] https://transparency.x.com/en/reports/amars-in-the-eu (accessed on July 29th 2024).
[45] https://transparency.x.com/en/reports/amars-in-the-eu (accessed on July 29th 2024).

- Order Twitter to stop processing personal data of affected users for artificial intelligence purposes under Article 58(2)(d) and (f) GDPR.

5.5. Penalty

We assume that Twitter's violations of Articles 5(1) and (2), 6(1), 6(4), 9(1), 12(1) and (2), 13(1) and (2), 17(1)(c), 18(1)(d), 19, 21(1) and 25 GDPR overall amount to a clear intentional breach of the law - especially in the light of the long list of previous CJEU, EDPB and SA decisions. We also reiterate that given that this irreversible processing has already taken place, the harms to the rights and freedoms of data subjects are incomprehensibly significant.

We note that Article 83(1) GDPR requires that Supervisory Authorities issue fines that are "*effective, proportionate and dissuasive*".