A Decision-Making Process to Implement the 'Right to be Forgotten' in Machine Learning

Katie Hawkins*[0000-0001-7927-5856] , Nora Alhuwaish*[0000-0002-5509-3625], Sana Belguith[0000-0003-0069-8552], Asma Vranaki[0000-0001-8944-6532], Andrew Charlesworth[0000-0002-9571-7383]

University of Bristol, Bristol, UK
{katie.hawkins, hw19625, sana.belguith, asma.vranaki, a.j.charlesworth}@bristol.ac.uk

Abstract. The unprecedented scale at which personal data is used to train machine learning (ML) models is a motivation to examine the ways in which it can be erased when implementing the GDPR's 'right to be forgotten'. The existing literature investigating this right focus on a purely technical or legal approach, lacking the collaboration required for this interdisciplinary space. Recent works has identified there is no one solution to erasure in ML and this must therefore be decided on a case-by-case basis. However, there is an absence of guidance for controllers to follow when personal data must be erased in ML. In this paper we develop a novel, decision-making flow that encompasses the necessary considerations for a controller. Addressing, in particular, the interdisciplinary considerations relevant to the EU GDPR and data protection scholarship, as well as concepts from computer science and its application in industry. This results in several optimal solutions for the controller and data subject, differing with levels of erasure. To validate the proposed decision-making flow a real case study is discussed throughout the paper. The paper highlights the need for a clearer framework when personal data must be erased in ML; empowering the regulator, controller and data subject.

Keywords: Right to be Forgotten, GDPR, Erasure, Machine Learning, Machine Unlearning, Decision-Making.

1 Introduction

In 2017, a company known as Clearview AI created a database that now has more than 30 billion facial images scraped from online accessible sources, including social media networks and videos extracted from online platforms [1]. Facial images are considered as biometric data,[1] and are particularly sensitive, due to the link to a person's physical identity and its unique way to identify someone. The vast majority of people whose images were scraped and processed were unaware of Clearview AI's methodology [2].

[1] GDPR Article 4(14) defines biometric data as "personal data resulting from specific technical processing relating to the physical, physiological or behavioural characteristics of a natural person, which allow or confirm the unique identification of that natural person."

Several complaints were made by data subjects, defined as natural persons that had their images processed by the company [3, 4]. After numerous investigations in various countries in the EU, such as France [2], it was found that Clearview AI was in breach of the General Data Protection Regulation (GDPR) [3]. This regulation aims to ensure the protection of natural persons' fundamental rights and freedoms, and sets rules for the free movement of personal data [5]. The GDPR grants data subjects several rights, including an important and novel right; the right to erasure, otherwise known as the 'right to be forgotten' (RTBF) articulated in Art.17 GDPR. The RTBF is defined as the data subject's right to obtain erasure of personal data related to the data subject without undue delay, on six grounds, explained in Section 3.1. In this case, Clearview is defined as the controller, where the company determines the purposes and means of processing personal data [6]. This controller's processing of personal data, breached various elements of the GDPR including processing without a legal basis and failure to take into account data subjects' rights, such as the right to erasure [2]. Clearview AI was given formal notice to cease the collection and use of personal data in the absence of a legal basis and to comply with erasure requests within a period of two months [2].

This is not a unique case, there are several other cases where a controller has been ordered to erase personal data from their facial recognition systems [7]. These systems take advantage of Machine Learning (ML) – training models on a large database of images, that provide the capability to identify faces of known individuals, comparing faces, and detecting similar faces in a database [8]. Following on from these examples, the question regarding erasure arises: how should controllers erase personal data that has already been deployed in ML models?

Answering this question is significantly challenging due to the complexities of ML, which does not retain data in raw format, but rather embeds the data throughout the system, as part of the development of the model [9]. The erasure problem for ML models, as expressed by Dang, is similar to asking a person to forget a single lesson from an entire educational background [9]. Numerous studies have shown that some ML models tended to 'memorise' data on which they had been trained [10, 11]. This memorisation may lead to a 'leak' of those data, and in some cases, result in re-identification of data subjects [11, 12]. This is a privacy attack known as membership inference [13]. The attacks have been applied on many supervised models and generative models [11].

The RTBF has captured the attention of many researchers due to its importance in strengthening data protection rights [14] and the challenges of practically implementing its requirements [15]. Both legal and technical literature have attempted to overcome this challenge, but not without several limitations that have inhibited the effective application in ML [16, 17]. This results in multiple problems; firstly, researchers have argued that certain ML models could be classified as personal data under EU's GDPR [11, 12]. The GDPR adopts a wide definition of personal data, as any information related to an identified or identifiable natural person – going beyond name or phone numbers to include dynamic IP addresses, browser fingerprints or smart meter readings. This wide approach is also adopted by the Court of Justice of the EU (CJEU) [18].

Secondly, this creates problems from the perspective of the data subject. If the ML model is classified as personal data but erasure is too complex, it negates the ability of the data subject to exercise this right. This could breach individuals' fundamental rights

to privacy and data protection in Articles 7 and 8 of the Charter of Fundamental Human Rights [19], as well as the RTBF in Article 17 GDPR [3].

Thirdly, public reports from controllers such as Google have shown an increase in requests to delist content under European privacy law [20]. Erasure requests are likely to continue to increase as the creation of new privacy attacks are able to successfully identify data subjects [11]. Therefore, there is a need for developing erasure techniques to efficiently deal with large volumes of requests as well as ensuring personal data is erased from the ML model currently in operation.

A final problem is the flexible interpretation of the RTBF from a legal perspective, noted by the European Union Agency for Cybersecurity (ENISA) [21]; the deliberate generality and extensiveness of the RTBF results in "a range of interpretations appropriate for many different situations". This may allow controllers to have more than one suitable approach for erasure, but it is not clear what the approaches are or how those approaches can be determined from a technical perspective.

1.1 Paper Contributions

This paper focuses on bridging the gap between the legal and technical literature for the RTBF in ML models. In particular, the paper is the first to integrate the multidisciplinary problem space of the EU GDPR and data protection scholarship, with concepts from computer science and industry application. In doing so, the authors create a novel decision-making flow that provides a practical outcome for the controller to implement the RTBF in ML models. The decision flow identifies the relevant decision, outlines the necessary legal and technical considerations and assesses alternative resolutions. This supports the need for a clearer framework when personal data must be erased in ML; empowering the regulator, controller and data subject.

1.2 Paper Structure

The rest of this paper is organised as follows: Section 2 provides an overview of the basic concepts related to ML, followed by a literature review. Section 3 demonstrates the decision-making flow alongside a discussion of GDPR's requirements for the RTBF, followed by the erasure techniques. The paper concludes with a summary of the discussion and a proposal for future works.

2 Background

2.1 Machine Learning

To gain a grasp of the challenges and this paper's recommendations, a brief introduction to the field of ML is necessary. ML is a set of techniques that allow computers to learn by creating or using algorithms based on data [22]. Most of the literature refers to ML models, this represents the output of a ML algorithm that is run on some data. ML varies in complexity due to the variety of options available, including the data processing

methods, algorithm types and objectives. For simplicity here, the development of a trained ML model is split into two phases: (1) Training and Validation Phase, and (2) Deployment and Monitoring Phase.

Within the first phase, an objective for the model is defined, for example, the ability to identify and classify an image as a dog or cat. It then proceeds with selecting an algorithm, as well as gathering and preparing data. The prepared data will be used to train, test and validate the model. For the purpose of enforcing the RTBF, the assumption is that personal data has been collected and processed within this phase. In the second phase, the model has been created and now deployed for use on new unseen data, hence 'learning'. The deployed model continues to be maintained and its performance monitored.

2.2 Literature Review

The literature review aims to demonstrate the inspiration for the authors' proposed decision-making flow. The literature can be categorized into three categories. First, legal literature that considers the RTBF. Second, literature that considers the application of the RTBF in ML, including the proposed technical solutions. Finally, literature that considers both the legal requirements and technical solutions.

After the GDPR's explicit protection of the RTBF, numerous regulatory guidelines and academic papers have been published to explain the right [23–27]. However, understanding how the RTBF is to be adequately implemented in ML practices remains inconclusive. There are three types of legal literature on the RTBF that are considered significant for this paper. The first type is the literature that discusses the RTBF's scope, grounds, exemptions and the need to balance the RTBF with other rights and interests [28]. Such literature provides the basis for the authors to understand and identify the legal requirements for implementing the RTBF, which frames the first part of the proposed decision-making flow, explained below. However, the majority of this literature lacks consideration of the technical development and application of ML which could significantly impact practical interpretation of the RTBF. The second type of legal literature focuses solely on the implementation of the RTBF in the search engine field, as well as analysing critical cases that mainly addressed its implementation in relation to Google, notably the CJEU decision on *Google Spain* [29] or *Google vs CNIL and GC and Others* [30], since this is the most common practice of the right [31–35]. The CJEU's flexible and subjective interpretation of the right in its jurisprudence leads the authors to construct an adaptive holistic approach to implementing the RTBF. However, the *Google Spain* case was pre-GDPR and much current jurisprudence in the CJEU or Member States' Courts and EU DPA guidance addresses particular scenarios like delisting requests received by search engines [24, 27]. So there are limits to how much extrapolation can be employed given the differing contexts. The third type of legal literature focuses on the barriers and challenges to applying the RTBF, and either criticises the vagueness of the legal requirements for implementing the right, or expresses the technical difficulty and impracticality of applying the right [16, 36, 37]. These challenges inspired the authors to construct a decision-making flow that could aid in overcoming these challenges.

In the context of the technical literature within the RTBF, the majority focus on a range of techniques to determine *how* to erase training data from models. The objective is to determine how a controller can remove training data from the established knowledge of a deployed ML model in phase 2. The field of research is known as machine unlearning, proposed by Cao and Yang [38]. The current state-of-the-art attempts to produce machine unlearning solutions that overcome challenges relevant to the practical deployment, for example, reducing the computational efficiency, cost and skills required [39]. Limitations usually arise in the applicability of machine unlearning solutions, as proposals lack a broad scope for ML models. It is crucial that these limitations are presented to a controller, as this will determine the appropriate erasure technique. Therefore, the technical considerations of each erasure technique, including the applicability, is discussed in Section 3.2. A further limitation is the absence of legal analysis within the proposed techniques.

In the decision-making literature, there has been no attempt to create a decision-based process relevant to the RTBF in ML. Therefore, the final part of the literature review focuses on the literature that involves both a legal and technical discussion. One of the first papers to highlight the interdisciplinary gap was published in 2018, arguing that the current privacy regulation is not fit to handle the challenges of AI [17]. It provides a technical and legal discussion of the problem space and calls for further research to investigate the balance between the RTBF and a ML model's need to remember information used to train it. However, it limits the technical solutions to differential privacy and data minimisation. These techniques are preventative measures, as it assumes the model has not yet been trained. It also lacks the capability for exact erasure, and remains prone to privacy attacks [40]. Another study investigates the RTBF and its implementation in ML [9]. They review the definitions of the RTBF in several major legal documents, and its application in practice. It highlights similar questions relating to the level of erasure and the techniques required, but limits its discussion to a brief analysis. The paper argues that differential privacy can be considered as the framework to define the RTBF, whilst machine unlearning is a usable technique to practice the RTBF. However, the paper lacks the required analysis and understanding of the RTBF requirements and other GDPR provisions related to the RTBF, such as the RTBF's grounds and exemptions. In another paper, the types of techniques for erasure are expanded from differential privacy, and include influence functions and machine unlearning [41]. The majority of research within this field concludes with the need for more interdisciplinary researchers to identify other technologies that can be used, as well as discuss the wider problem space for both the legal and technical fields [16, 17, 41, 42]. Other interdisciplinary papers on the topic focus on classifying models as personal data [12]. For example, Veale, Binns and Edwards' paper explains model inversion and membership inference attacks, and how the GDPR is likely to classify models as personal data. It then describes selected consequences for data subjects' rights to have access to new information, erasure and objection. Although the paper limits the legal considerations to the applicability of personal data in ML, it helps to shape the understanding of personal data and privacy harm in ML models. This interpretation is incorporated into the considerations when deciding on the level of erasure within the authors' decision-making process.

3 Decision-Making Flow

Figure 1 presents the first decision-making flow for implementing the RTBF in ML. Produced by the authors, it aims to illustrate (at a high-level) the steps and decisions once a RTBF is requested. The following section walks through the flow in greater detail. Section 3.1 discusses the initial legal steps in the flow, including the level of erasure, grounds and exemptions. Section 3.2 then considers the lower level of the flow, where erasure techniques must be applied.

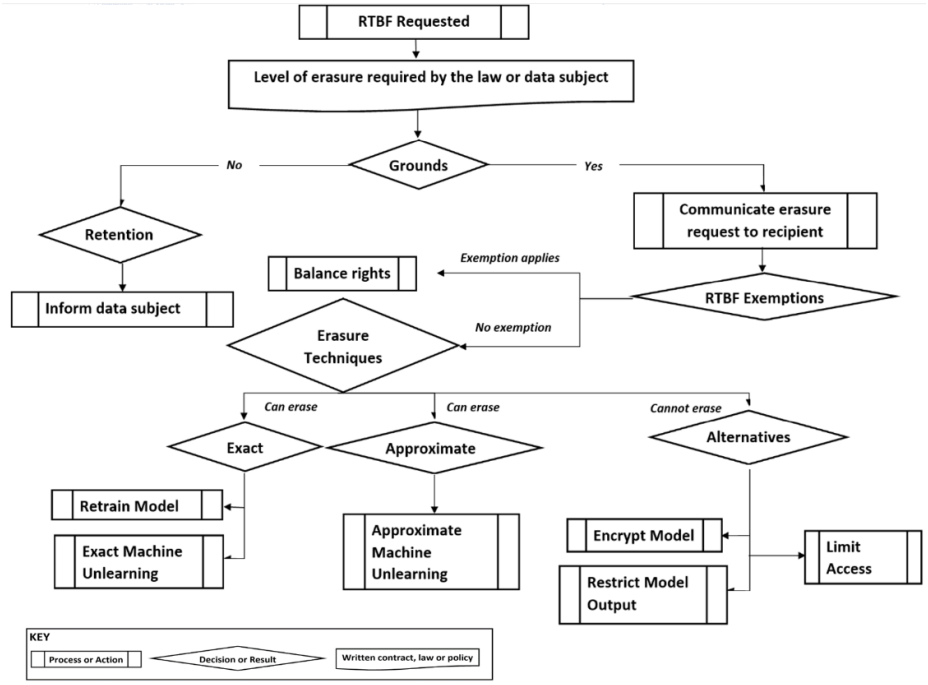

Fig. 1. Decision-Making Flow

3.1 Legal Requirements for Implementing the RTBF

There are multiple layers of legal requirements that should be considered in order to implement the RTBF. This section analyses the legal requirements for the RTBF and argues that each legal requirement should be considered and assessed in a decision-making flow in order to implement it in ML.

Level of Erasure required by the law or data subject. As can be seen in Figure 1, the legal requirements that the controller should consider in the RTBF decision-making flow start with identifying the required level of erasure. The erasure level is determined by two factors: the context of the data subject's request and the law. The GDPR does not specify how the controller should receive and understand a data subject's request,

which leaves it to the controller to inform the data subject of the differences between erasure levels and manage such requests accordingly [43]. Controllers should be absolutely clear with data subjects about what is meant by erasure and what actually happens to their personal data once the controller has erased it [44]. The controller can require data subjects to state the exact data points they want to be removed (e.g., remove a name or address) [44]. In that way**,** the controller should be able to identify the level of erasure requested by the data subject [43]. Regardless of the technical possibility of erasure levels in ML systems, the data subject can either request limited erasure where the personal data is removed from a specific level of the system, such as the user interface, or request complete erasure that removes personal data from the whole system. If the data subject requests complete erasure, to what extent does the GDPR require the controller to erase the data in complicated systems such as ML?

The GDPR does not clarify the extent of erasure required. Both the CJEU and national courts have interpreted the RTBF as limited erasure by, for example, restricting access [45] or removing the link between personal data and search results associated with the data subject's name [29, 30]. The European Data Protection Board (EDPB) [46], the body responsible for ensuring the consistent application of the GDPR throughout the EU, has published guidelines on interpreting the RTBF [24]. Yet, the guidelines focus only on delisting requests submitted by data subjects in search engines and assert that the delisting request does not result in the personal data being completely erased, as the requested personal data is not erased from either their source (the website) or the search engine's index [24]. Personal data can remain publicly available and accessible, but no longer be linked to the results of searching on the data subject's name [24]. The guidelines do, however, emphasise that search engine providers are not exempt from the duty to fully erase in exceptional cases. Unfortunately, the guidelines do not provide further information about these exceptional cases, they only provide an example. Therefore, it appears, erasure can be interpreted as limited (without the personal data being completely erased, as is the case with a delisting request) or complete erasure (in some exceptional cases). Interpreting erasure in ML based on the analogy of delisting in search engines is challenging. Unlike search engines, the desired and required impact of erasure in ML is not to remove the data from the public. Another difference is that ML processing does not contradict the right to freedom of expression or use of personal data for journalistic purposes, which may require making personal data available at a certain level. Rather it may contradict other interests, such as the controller's legitimate interest or legal obligations. In addition, the function of search engines is dissimilar to ML, which is often both more complex and difficult to understand and explain.

Local Data Protection Authorities (DPA), independent public authorities that supervise, investigate and have corrective powers to the application of the GDPR, provide expert advice on GDPR issues, and handle complaints lodged against violations of the GDPR and the relevant national laws [47], have published guidelines. For example, the French regulator's AI compliance guidelines assume that personal data can be present in all life cycles of AI including training data, deployment data and data in the model; therefore, data subjects' rights will apply across all these cycles [48]. It is important to note that the ML phases may differ from AI phases explained by the guidelines. This is

because ML is just a subset of AI and that AI has much broader application. Thus, this paper adopts the ML phases outlined in Section 2.1.

Even if the RTBF applies to all phases and throughout the life cycles of the ML system, it is still questionable whether the law requires an exact or approximate degree of erasure from the deployed model, as it can be technically difficult to guarantee 100 percent erasure. The GDPR neither requires complete erasure nor prevents it. However, the legislation appears to strengthen the RTBF, encouraging controllers to devise various techniques that help meet the objective of the GDPR in protecting fundamental rights, for two main reasons. First, the difficulty and impracticality of complete erasure must have been envisaged by the multidisciplinary experts who participated in drafting the GDPR. This is evident in the proposal for the GDPR, which allowed restriction instead of erasure when technically difficult to erase personal data, and limited this exemption to systems that were designed before the application of the GDPR [49]. This paragraph was omitted in the final version of the GDPR, which may indicate that technical difficulties to erase are no longer considered a reasonable excuse for restricting personal data instead of erasure [50]. In addition, erasure may differ from one type of data to another. For example, in the case of images by "blurring the picture with no retroactive ability to recover the personal data that the picture previously contained, the personal data are considered erased in accordance with GDPR" [51]. Additionally, local DPAs have the discretion to assess the need for erasure on a case-by-case basis as a result of the normative evaluation, which places the controller in a flexible position where it may be required to refuse the erasure, implement complete, limited, or approximate erasure based on data type, level of harm, the grounds and exemptions of the RTBF, as discussed below.

Grounds. The second legal requirement that should be considered by the controller in the RTBF decision-making flow is establishing a ground for the erasure request. The RTBF is not an absolute right, and the erasure request should be established on one of six grounds specified in Article 17(1) GDPR. These different grounds partially overlap [52]. Three of the grounds assume the lack of a legal basis for processing: when the personal data are no longer necessary for the purpose of collection and processing (subparagraph (a)), data subject withdraws consent, and there is no other ground for processing (subparagraph (b)), or the controller processes personal data unlawfully (subparagraph (d)). The latter ground, which can be seen as a general clause, is clarified by Recital 65 in which it is stated that the RTBF can be invoked by a data subject where the processing does not comply with the GDPR. The fourth ground is based on Article 21, the invocation of the right to object when processing is necessary for public interest or in the exercise of official authority vested in the controller in point (e) and on legitimate interest in point (f) of Article 6(1) where there are no other overriding interests, or when the objection is based on direct marketing (subparagraph (c)). The fifth ground is when processing personal data of a child in relation to the offer of information society services based on the child's consent. In this situation, the data has to be erased upon simple request. Recital 65 clarifies that the data subject should be able to exercise the RTBF notwithstanding the fact that they no longer are a child. It is emphasised that the RTBF is relevant in particular where the data subject has given his/her consent as a

child and is not completely aware of the risks involved by the processing, and later wants to erase that personal data (subparagraph (e)). Finally, (subparagraph (f)) requires the controller to erase personal data when this is mandated under EU or Member State law.

As illustrated in Figure 1, the erasure request requires a decision that it is either accepted as establishing a ground or be refused. The controller will need to internally assess whether or not the purpose of processing was, and is still, necessary, and identify the legal base for the personal data collection. If the initial legal base is consent, then this cannot be changed to another legal base if the data subject withdraws his/her consent. The controller should have a clear understanding of all the GDPR requirements in order to determine whether the personal data is lawfully processed. The controller also needs to take the data subject's age into consideration. In the Clearview AI example, the most applicable reason for the application of the RTBF is unlawful processing. This is because Clearview AI violated several GDPR requirements; processed personal data without a legal basis, and not adequately informing data subjects about, or facilitating the exercise of, their rights to access and erasure [53].

If it has been decided that the personal data in question provides grounds for erasure, the controller must take reasonable steps to inform other controllers of the data subject's request, such as when personal data has been made public (for example by broadcasting) [54]. These steps must include technical measures which take into account available technology and implementation cost. In addition, the controller should, to the extent possible, notify anyone to whom the personal data has previously been disclosed according to Article 19 GDPR. The controller must inform the data subject about those recipients if the data subject requests it.

Alternatively, if the erasure request fails to establish that any of the aforementioned grounds are applicable, the controller can retain the personal data and must inform the data subject of the decision without unduly delay, which was specified as a period of a month, in accordance with Article 12 of the GDPR.

Exemptions. Even when one of the above grounds applies, the controller must make sure that no exemption applies to the personal data before erasing the personal data. In the GDPR, the controller is obliged to refuse the erasure request if the processing is necessary for one or more of the following. First, exercising the right of freedom of expression and information (subparagraph (a)). Second, complying with a legal obligation under the EU or national laws that the controller is subject to, for the performance of a task carried out in the public interest or in the exercise of official authority (subparagraph (b)). Third, cases in which the processing of special categories of personal data is necessary for public health in certain grounds provided for in Article 9 GDPR (subparagraph (c)). The same holds true when processing is necessary for archiving purposes in the public interest, scientific or historical research purposes or for statistical purposes in accordance with Article 89(1) GDPR (subparagraph (d)). This exemption can be invoked insofar as the RTBF is likely to render impossible or seriously impair achieving the objectives of the processing. The final exemption can be invoked when processing is necessary for the establishment, exercise or defense of legal claims (subparagraph (e)).

In the Clearview AI example, the first exemption is unlikely to be relevant. Although the RTBF in ML may contradict other fundamental rights or public interests, it rarely contradicts the right to freedom of expression and information, as it is the case in search engines. It does not seem, from the company's function, that it has legal obligations to retain personal data in, for example, taxation or labour requirements. Clearview AI processing might be necessary for the performance of a task carried out in the public interest or in the exercise of official authority, as the company offers its services to law enforcement authorities in order to help identify perpetrators or victims of crime, so the second exemption may apply. Clearview AI is using people's images, which are considered a special category of the personal data in the GDPR, but it seems that the company does not need the personal data for reasons of public interest in public health, which renders that third exemption as inapplicable. The fourth exemption applies to Clearview AI when the processing is necessary for scientific research or statistical purposes. Similarly, when processing is necessary for exercise or defense of legal claims.

The way the GDPR articulates the exemptions creates three main considerations when applying the RTBF in ML. First, the GDPR shifts the responsibility of balancing the right to privacy and data protection and other rights from the government to the controller [31]. In other words, controllers are required to determine whether the data subject's rights outweigh, for example, other rights, interests or the legal requirements. The EU and national courts' interpretation of balancing has been inconsistent and based on a case-by-case approach [55]. The courts sometimes order erasure [56], or require restricted access [45], or refuse the erasure request and allow the controller to retain the personal data [57]. The criteria for conducting such balancing are subjective, and available guidelines are mainly relevant to search engines and, as such, are inadequate to be implemented in ML [24].

Second, a problem may arise when controllers, instead of undertaking a long and subjective assessment of the exemptions and balancing of the RTBF against other public interests, automatically erase personal data upon a RTBF request. This would lead to 'over-erasure' out of an abundance of caution to avoid financial penalties [58]. However, the choice to favour the RTBF may be hypervigilant, and the controller is obliged to consider the exemptions, especially when the RTBF contradicts other fundamental rights or legal requirements. Third, before erasing personal data based on the RTBF, the controller should be aware of all legal requirements from EU laws or national regulations that apply within the jurisdiction relevant to the controller.

There are many variables at play and each RTBF request will have to be evaluated individually. After assessing the applicability of the five exemptions, controllers have two tracks: either to invoke an exemption or not. In the first track, the controller should balance the RTBF with the interests in the applied exemptions. The result of the balance can be refusing the RTBF request and retaining personal data, erasing personal data or coming up with an alternative solution that balances the RTBF and other interests. Erasure and other alternatives are discussed in the next section. The second track is the way when there is no exemption to be considered and the controller must erase the personal data.

3.2 Machine Learning Techniques for Erasure

At this point within the decision-making flow, the controller is required to erase the data subject's personal data. Thus, the discussion moves to the ways in which data can be erased from the ML model.

Before considering the techniques for erasure, it is crucial to understand that personal data can be contained at numerous points within ML phases. As mentioned in Section 2, within the training and validation phase, personal data can be collected and used within the training dataset. This is not the case in all models, where personal data is not used or personal data is removed during pre-processing. For the purposes of the RTBF, this paper assumes some form of personal data is held within the training dataset and possibly in ML models after it has been deployed [59].

Exact Erasure. Exact erasure techniques focus on the challenging case where personal data is involved in the ML model itself. The goal is to identify the exact same deployed model as if the user data was never part of the training dataset. For exact erasure, it is assumed that simply erasing the data subject's data from the training dataset (independent of the trained model) is insufficient. It is the strictest requirement for removal as the requested data must be removed from the deployed model (phase 2) and the training dataset (phase 1). This complete and thorough erasure is most often what is desired by data subjects [39], but can be the most challenging to implement.

Although the controller may not be able to determine whether or not personal data can be inferred from the model, exact erasure is necessary when the Data Protection Authority (DPA) or the data subject present evidence that their personal data can be inferred from the deployed model [48]. This links back to the notion of privacy harms, where data subjects may use research on inference attacks as evidence to illustrate the unanticipated use of personal data. Thus, the data subject will likely have to apply or prove such attacks to the ML model in question to sufficiently support the plausibility of re-identification of their personal data. Some may argue that neither the DPAs nor the data subjects have the technical capabilities to provide such evidence, however, recent publications could aid data subjects (and potentially adversaries) to determine whether their data was used to train a ML model [60]. Whilst the publication is only applicable to text-generation models, increased RTBF requests could in turn generate an industry for gathering such evidence for the data subject. Other technical research has investigated the quality of privacy protection and the detection of privacy violations (e.g. RTBF), where the authors suggest these as ways to verify possible misuse of the data in ML [61].

Thus, this scenario is relevant in cases where other removal methods are insufficient, and the data subject has proven evidence of personal data inference. This is likely to be a growing future scenario that could be adopted as more exact erasure techniques are identified and endorsed by industry.

Techniques. The first and naïve approach to erasing a data subject's data from many basic ML models is to completely retrain the deployed model on the remaining training dataset. This approach requires the controller to retrain the model from scratch after

every erasure request. This retraining carries significant energy, time, labour and costs [12]. The controller must consider the likelihood of a request and whether the computational time and effort to retrain is the best solution. For most controllers, this would not be a viable option. This impractical approach motivates academics to research efficient ways for models to "unlearn" the requested data from their existing deployed model.

The second approach is to apply exact unlearning methods. This proposed solution allows the system to 'unlearn' a piece of data without having to retrain the entire model and the associated relationships between data [62]. Thus, it is more efficient and practical from the controller's perspective than the previous option, especially where complete erasure is necessary for more than one RTBF request. This area of research is still active and there is no endorsed technique, nor a technique applicable to every ML model, and researchers have stated it is likely impossible to have a technique that would be able to fit such criteria [17]. Currently, there are only two methods relevant to exact unlearning. The first method provided a general definition for an unlearning algorithm in a general case, i.e. without being specific to a particular training dataset [63]. The paper then proposes an unlearning algorithm that is more efficient than retraining from scratch. The limitation is the applicability, as it can only be applied to controllers required to erase data in clustering problems, specifically k-means clustering. Possibly the best progress for model-applicability was published in 2021, where the authors proposed a framework that is applicable to any unlearning algorithm but designed to achieve the largest improvements for deep neural networks [64]. It partitions the deployed model into smaller sub-models, removing the need to retrain the entire model and instead retrain a sub-model that contains the personal data requested for erasure.

Before applying exact unlearning techniques, controllers may need to consider the capability of the controller. For example, the unlearning solution published in 2021 [64] will still require technical expertise to ensure the framework can be successfully tailored to the controller's deployed model. It is also dependent on the type of erasure request, this proposal only considers the type where certain items are removed from the training data, it does not include class removal, for example, where image removal is required. As the goal for exact erasure is to identify the exact same deployment model as if the user data was never part of the training dataset, controllers may question how the removal of personal data might affect the performance of the deployed model. It is hard to predict how erased points will change the model [65], but research in unlearning has looked at how many data subject requests can be performed before degrading the accuracy of the deployed model [66]. The current state-of-the-art claims it could handle a higher batch of unlearning requests than the estimated unlearning requests anticipated by Google [64]. Also, the controller may have to consider the architecture of their ML system, as the ease of erasure will depend on the way in which the model is deployed, including whether the model is outsourced. In this case, some controllers may not be able to erase data from or even understand the outsourced ML model.

It is important to note that exact erasure is not always the best solution for the data subject. For example, exact erasure does not imply complete privacy of the individual [63], as it could be possible to re-identify the data subject that had their data erased (in the rare case where someone has access to both the old and new model).

Ultimately, there is an extremely long way to go to reach any endorsement or use of exact unlearning in practice, and therefore remains a great opportunity to produce more exact unlearning approaches that can be realistically adopted.

Approximate Erasure. At the same point within the decision-making diagram as exact erasure, approximate erasure is another approach to erase data from the deployed model. The difference is that it relaxes the strict requirements of exact erasure to make the problem computationally manageable [62]. This results in a range of techniques that do not guarantee all the data subjects' personal data has been removed, rather, providing a statistical bound as to its ability to remove the requested data from the existing deployed model. The goal of approximate unlearning is to approximate the model parameters one would obtain by exact unlearning [67].

Approximate erasure may be deemed necessary when the controller is unable to retrain the model from scratch and the exact unlearning techniques are not applicable or computationally expensive. Due to the limited techniques for exact unlearning and current lack of endorsement in industry, this is a likely scenario.

Techniques. The techniques considered for approximate erasure include any technique that replaces/removes the majority of data that identifies the data subject. For approximate unlearning, most of the approaches either perform less computationally expensive actions on the parameters or modify the architecture [39]. As the requirements are more relaxed, there are more solutions applicable to a range of ML models, generating roughly 40 proposed solutions. As with exact unlearning, there is no one endorsed method, and the techniques differ on the erasure request and ML model used by the controller. The following approximate unlearning techniques have been chosen due to their wider applicability to ML models and efficiency of unlearning.

In one approach, the authors propose a differentially private unlearning mechanism for streaming data removal requests [68]. It is the first paper that examines the provision of deletion guarantees with the motivation that users may wish to erase their data exactly because of what deployed models reveal about them. This motivation for removal is likely to increase as new membership attacks re-identify data subjects from the deployed model. This is a suitable legal justification, requiring revocation under the "unlawful processing" grounds explained above. Another proposed solution for approximate erasure uses a certified removal technique applied to linear models [69]. This also takes advantage of differential privacy's objective, stating that the model after removal is indistinguishable from a model that never trained on the removed data. This unlearning solution is mostly applicable to deep neural networks in the image domain. Another relevant to classification models looks at approximate unlearning for the specific setting of class-wide erasure requests, for example facial recognition [42].

Similar to the considerations in exact erasure, it is difficult to determine how the techniques can be applied, and there is an assumption in the literature that the controller has the relevant expertise to be able to apply such methods. As the techniques exist currently, there is no evidence that these have been applied to industry. The majority of the paper's proposals do use example data sets as part of its evaluation, but the effectiveness for industry is not yet clear.

Some academics argue that approximation techniques are more beneficial for the privacy of the data subject: the techniques for approximate unlearning, such as differential privacy, do not focus on data erasure but attempt to make data private or non-identifiable. As approximate unlearning techniques do not erase all the requested personal data, the data subject may want some form of verification that the unlearning technique has been applied. The controller is obliged according to Article 12 GDPR to inform the data subject of the controller's decision on the request. Although there are no direct requirements for the verification of erased data points, it is worth mentioning that one of the main principles of the GDPR is accountability, which means that the controller is the one responsible for complying with the law and being able to demonstrate compliance. Controllers may demonstrate compliance by stating whether the erasure request is implemented or not, without providing details of how the system functions. This undetailed notice to the data subject may be a result of wishing to protect the trade secret and intellectual property in the system. Either way, whether the erasure is proved or not, the controller remains responsible for any breach or non-compliance. That means if the controller was not considering a technical approach that applies the required erasure, the controller may face the risk of being held non-compliant. Therefore, the technical literature does focus on providing theoretical guarantees alongside their proposed unlearning techniques [39]. It would be advantageous for controllers to have the ability to efficiently and confidently confirm to data subjects their requested removal was successful. There are various ways to verify erasure, including measuring erasure via applying privacy attacks [70], information leakage [71] or apply cryptography with verifiable proof [72].

Alternatives. As research in machine unlearning is still evolving, it is likely that controllers may struggle to adopt the exact and approximate unlearning measures discussed above. Therefore, the decision-making flow considers that some controllers cannot erase the data and alternative approaches must be applied. This is based on the assumption that the chosen technique is adequate to meet the requirements of the data subject and the law. This is because the alternatives do not erase the subject's personal data in the ML model, but they do attempt to decrease the privacy harms associated with ML. Alternatives may be necessary where data cannot be erased, for example, if the erasure of data will destroy the model or result in unfounded complexities, especially where third parties are involved.

Techniques. One of the possible alternatives is restricting the model output to prevent privacy harms against the deployed model. This could be relevant where the data subject has provided sufficient evidence of re-identification of the deployed model. For example, mitigation strategies have been proposed to prevent membership inference attacks against ML models if the data cannot be erased [13]. This includes limiting the model's predictions to top k classes and decreasing the precision of the prediction. This approach reduces the work required for the controller but in turn reduces the accuracy of the model's output.

Encrypting the ML model is another technique used to protect the confidentiality of the model output, without affecting the performance. Homomorphic encryption has

been proposed on the gradients [73]. The scheme can prevent information leakage to the honest-but-curious cloud server, focusing on controllers using collaborative deep learning. A simpler technique is to erase or restrict access to the training dataset only. Restricting access could include storing encrypted copies of the training data. This was a practical interpretation by ENISA under the RTBF [21].

It is important to note that the RTBF and the right to restriction are different rights in the GDPR. Article 18 grants data subjects the right to restrict the processing of their personal data in specific circumstances different from the RTBF grounds [74]. As clarified by Recital 67 GDPR, restriction of processing can include temporarily moving the selected personal data to another processing system; making the selected dataset unavailable to users, or temporarily removing published data from a web page [75]. However, restriction can be imposed as a result of balancing different rights or interests, as stated in the CJEU judgement in the *Camera di Commercio* case, to balance the data subject's right to erasure and legal obligations [45]. Similar to restriction, in backup systems, it can be impractical to erase data because of the technical difficulty or security requirement [36]. In those cases, the UK DPA, the Information Commissioner (ICO), for example, directs controllers to explain this clearly to the data subject and not to use the backup data for any other purposes; to put the requested data 'beyond use', publicly and privately [26].

For some controllers, the alternatives could be the only viable option to implement, especially as the machine unlearning techniques are still in their infancy. The techniques are less costly in terms of time and computation. However, this may also raise the concern that controllers will use these alternatives as an 'easier' way of complying with the data subject's erasure request. Thus, these techniques must only be used where the controller has provided sufficient evidence that it is unable to use erasure techniques.

4 Conclusion

This paper is the first to investigate a wider scope of the problem space for the RTBF in ML, and in turn hopes to aid both regulators and controllers for future cases. It illustrates that implementing the RTBF is not a matter of mere erasure of personal data or refusal of the erasure request, rather it is about establishing a decision-making process that begins by identifying the required level of erasure, establishing a ground for erasure, balancing the right requested against other rights and interests, and then deciding the most appropriate technique in a case-by-case approach.

In the case of Clearview AI, the decision-making flow demonstrates that erasure is required, but the level of erasure was not mentioned. Firstly, if Clearview AI was to only erase the training database of the facial images, independently to the deployed model, the level of erasure would not be sufficient; the deployed model would still be able to detect the erased images as the biometric template that has been trained on has not changed. Although erasing the whole model is not what the GDPR nor its guidance directly require, it is likely a scenario that may be requested by the DPA or courts. Thus, the decision-flow would find that exact erasure is necessary. Instead, Clearview AI did

not address the formal notice and the French DPA imposed a penalty of 20 million euros [2]. Similar fines and erasure orders were imposed by the competent authority in the UK, Italy, Greece, Australia and the US [76–79]. For future erasure requests, and to avoid future fines, Clearview AI (and other controllers) must allow the use of efficient unlearning techniques, weighing up the considerations (both legally and technically) to decide the best approach.

In summary, this work sheds light on critical decision-making challenges that warrant further investigation. Specifically, greater interdisciplinary research exploring how personal data can be inferred from the output of the ML model will underscore the need for more stringent erasure techniques. In addition, the authors recommend testing the proposed solution on both hypothetical and real-world scenarios to solidify its validity and feasibility. This could also involve considering other complex ML applications such as federated learning, repurposing, transfer learning and one-shot learning. Overall, addressing these challenges will contribute to the development of efficient and easy to apply erasure techniques that prioritise the privacy of the data subject.

Acknowledgements. Nora Alhuwaish is a PhD student sponsored by Kind Saud University. Katie Hawkins is a PhD student sponsored by the EPSRC Centre for Doctoral Training in Cyber Security.

References

1. Clearview AI | Facial Recognition, https://www.clearview.ai, last accessed 2023/01/31.
2. Facial recognition: 20 million euros penalty against CLEARVIEW AI | CNIL, https://www.cnil.fr/en/facial-recognition-20-million-euros-penalty-against-clearview-ai, last accessed 2023/01/19.
3. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). Hereinafter [GDPR]. Official Journal of the European Union L119, pp. 1–88, May 2016, http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:L:2016:119:TOC, last accessed 2022/09/11.
4. GDPR, Article 4(1).
5. GDPR, Article 1.
6. GDPR, Article 4(7).
7. Mann, M., Smith, M.: Automated facial recognition technology: Recent developments and approaches to oversight. The University of New South Wales Law Journal. 40, 121–145 (2017). https://doi.org/10.3316/ielapa.771179858194317.
8. Ugail, H.: Chapter 6 - Deep face recognition using full and partial face images. In: Davies, E.R. and Turk, M.A. (eds.) Advanced Methods and Deep Learning in Computer Vision. pp. 221–241. Academic Press (2022). https://doi.org/10.1016/B978-0-12-822109-9.00015-1.

9. Dang, Q.-V.: Right to Be Forgotten in the Age of Machine Learning. In: Antipova, T. (ed.) Advances in Digital Science. pp. 403–411. Springer International Publishing, Cham (2021). https://doi.org/10.1007/978-3-030-71782-7_35.
10. Hartley, J., Tsaftaris, S.A.: Measuring Unintended Memorisation of Unique Private Features in Neural Networks, http://arxiv.org/abs/2202.08099, (2022). https://doi.org/10.48550/arXiv.2202.08099.
11. Rigaki, M., Garcia, S.: A Survey of Privacy Attacks in Machine Learning, http://arxiv.org/abs/2007.07646, (2021). https://doi.org/10.48550/arXiv.2007.07646.
12. Veale, M., Binns, R., Edwards, L.: Algorithms that remember: model inversion attacks and data protection law. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences. 376, 20180083 (2018).
13. Shokri, R., Stronati, M., Song, C., Shmatikov, V.: Membership Inference Attacks against Machine Learning Models. arXiv (2017). https://doi.org/10.48550/arXiv.1610.05820.
14. Edwards, L., Veale, M.: Slave to the Algorithm? Why a "Right to an Explanation" Is Probably Not the Remedy You Are Looking For. Social Science Research Network, Rochester, NY (2017). https://doi.org/10.2139/ssrn.2972855.
15. Szeghalmi, V.: Difficulties Regarding the Right to Be Forgotten in the Case Law of the Strasbourg Court. Athens Journal of Law. 4, 255–270 (2018). https://doi.org/10.30958/ajl.4-3-4.
16. Fabbrini, F., Celeste, E.: The Right to Be Forgotten in the Digital Age: The Challenges of Data Protection Beyond Borders. German Law Journal. 21, 55–65 (2020). https://doi.org/10.1017/glj.2020.14.
17. Villaronga, E.F., Kieseberg, P., Li, T.: Humans forget, machines remember: Artificial intelligence and the Right to Be Forgotten. Computer Law & Security Review. 34, 304–313 (2018). https://doi.org/10.1016/j.clsr.2017.08.007.
18. Case C-434/16 Peter Nowak v Data Protection Commissioner [2017] THE COURT (Second Chamber) ECLI:EU:C:2017:994.
19. Charter of Fundamental Rights of the European Union., https://eur-lex.europa.eu/EN/legal-content/summary/charter-of-fundamental-rights-of-the-european-union.html, last accessed 2023/02/06.
20. Requests to delist content under European privacy law – Google Transparency Report, https://transparencyreport.google.com/eu-privacy/overview?hl=en_GB, last accessed 2023/01/27.
21. The right to be forgotten - between expectations and practice, https://www.enisa.europa.eu/publications/the-right-to-be-forgotten, last accessed 2023/01/19.
22. Alzubi, J., Nayyar, A., Kumar, A.: Machine Learning from Theory to Algorithms: An Overview. J. Phys.: Conf. Ser. 1142, 012012 (2018). https://doi.org/10.1088/1742-6596/1142/1/012012.
23. Ausloos, J.: The Right to Erasure in EU Data Protection Law. OUP Oxford, Oxford ; New York (2020).
24. Guidelines 5/2019 on the criteria of the Right to be Forgotten in the search engines cases under the GDPR (part 1) | European Data Protection Board, V.2 Adopted on

7 July 2020, https://edpb.europa.eu/our-work-tools/our-documents/guidelines/guidelines-52019-criteria-right-be-forgotten-search-engines_en, (2022).
25. Do we always have to delete personal data if a person asks?, https://ec.europa.eu/info/law/law-topic/data-protection/reform/rules-business-and-organisations/dealing-citizens/do-we-always-have-delete-personal-data-if-person-asks_en, last accessed 2022/12/07.
26. Right to erasure, https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/individual-rights/right-to-erasure/, last accessed 2022/11/12.
27. The right to de-listing in questions | CNIL, https://www.cnil.fr/en/right-de-listing-questions, last accessed 2022/11/28.
28. Kuner, C., Bygrave, L.A., Docksey, C., Kuner, C., Bygrave, L.A., Docksey, C., Drechsler, L. eds: The EU General Data Protection Regulation (GDPR): A Commentary. Presented at the February 13 (2020). https://doi.org/10.1093/oso/9780198826491.002.0001.
29. Google Spain SL and Google Inc v Agencia Española de Protección de Datos (AEPD) and Mario Costeja González. (2014).
30. Case C-136/17, GC and Others v Commission nationale de l'informatique et des libertés (CNIL), judgment of 24 September 2019 (Grand Chamber) (ECLI:EU:C:2019:773).
31. Tzanou, M.: The Unexpected Consequences of the EU Right to Be Forgotten: Internet Search Engines As Fundamental Rights Adjudicators, https://papers.ssrn.com/abstract=3277348, (2018).
32. Verschaeve, S.: Going Dark or Living Forever: the Right to be Forgotten, Search Engines and Press Archives, https://papers.ssrn.com/abstract=3669865, (2020). https://doi.org/10.2139/ssrn.3669865.
33. Klinefelter, A., Wrigley, S.: Google LLC v. CNIL: The Location-Based Limits of the EU Right to Erasure and Lessons for U.S. Privacy Law, https://papers.ssrn.com/abstract=3844968, (2021).
34. Globocnik, J.: The Right to Be Forgotten is Taking Shape: CJEU Judgments in GC and Others (C-136/17) and Google v CNIL (C-507/17). GRUR International. 69, 380–388 (2020). https://doi.org/10.1093/grurint/ikaa002.
35. Razmetaeva, Y.: The Right to Be Forgotten in the European Perspective. TalTech Journal of European Studies. 10, 58–76 (2020). https://doi.org/10.1515/bjes-2020-0004.
36. Politou, E., Alepis, E., Virvou, M., Patsakis, C.: The "Right to Be Forgotten" in the GDPR: Implementation Challenges and Potential Solutions. In: Politou, E., Alepis, E., Virvou, M., and Patsakis, C. (eds.) Privacy and Data Protection Challenges in the Distributed Era. pp. 41–68. Springer International Publishing, Cham (2022). https://doi.org/10.1007/978-3-030-85443-0_4.
37. Yoo, C.S.: The Overlooked Systemic Impact of the Right to Be Forgotten: Lessons from Adverse Selection, Moral Hazard, and Ban the Box, https://papers.ssrn.com/abstract=4124596, (2022).

38. Cao, Y., Yang, J.: Towards Making Systems Forget with Machine Unlearning. In: 2015 IEEE Symposium on Security and Privacy. pp. 463–480 (2015). https://doi.org/10.1109/SP.2015.35.
39. Nguyen, T.T., Huynh, T.T., Nguyen, P.L., Liew, A.W.-C., Yin, H., Nguyen, Q.V.H.: A Survey of Machine Unlearning, http://arxiv.org/abs/2209.02299, (2022). https://doi.org/10.48550/arXiv.2209.02299.
40. Protivash, P., Durrell, J., Ding, Z., Zhang, D., Kifer, D.: Reconstruction Attacks on Aggressive Relaxations of Differential Privacy, http://arxiv.org/abs/2209.03905, (2022).
41. Shintre, S., Roundy, K.A., Dhaliwal, J.: Making Machine Learning Forget. In: Naldi, M., Italiano, G.F., Rannenberg, K., Medina, M., and Bourka, A. (eds.) Privacy Technologies and Policy. pp. 72–83. Springer International Publishing, Cham (2019). https://doi.org/10.1007/978-3-030-21752-5_6.
42. Baumhauer, T., Schöttle, P., Zeppelzauer, M.: Machine unlearning: linear filtration for logit-based classifiers. Mach Learn. (2022). https://doi.org/10.1007/s10994-022-06178-9.
43. Gutmann, A., Warner, M.: Fight to Be Forgotten: Exploring the Efficacy of Data Erasure in Popular Operating Systems. In: Naldi, M., Italiano, G.F., Rannenberg, K., Medina, M., and Bourka, A. (eds.) Privacy Technologies and Policy. pp. 45–58. Springer International Publishing, Cham (2019). https://doi.org/10.1007/978-3-030-21752-5_4.
44. Deleting Personal Data, https://ico.org.uk/media/for-organisations/documents/1475/deleting_personal_data.pd.
45. Case C-398/15: Request for a preliminary ruling from the Corte suprema di cassazione (Italy) lodged on 23 July 2015 — Camera di Commercio, Industria, Artigianato e Agricoltura di Lecce v Salvatore Manni. (2015).
46. GDPR, Article 70(1)(a) to (y).
47. GDPR, Article 51 to 59.
48. AI: ensuring GDPR compliance | CNIL, https://www.cnil.fr/en/ai-ensuring-gdpr-compliance, last accessed 2022/12/05.
49. Council of the European Union, Proposal for a Regulation of the European Parliament and of the Council on the Protection of Individuals with Regard to the Processing of Personal Data and on the Free Movement of Such Data (General Data Protection Regulation), revised and consolidated draft, Interinstitutional File: 2012/0011 (COD).
50. P7_TA(2014)0212 Protection of individuals with regard to the processing of personal data. (2014).
51. Guidelines 3/2019 on processing of personal data through video devices | European Data Protection Board, https://edpb.europa.eu/our-work-tools/our-documents/guidelines/guidelines-32019-processing-personal-data-through-video_en, last accessed 2023/01/14.
52. Kranenborg, H.: Article 17 Right to erasure (‘right to be forgotten’). In: The EU General Data Protection Regulation (GDPR). Oxford University Press (2020). https://doi.org/10.1093/oso/9780198826491.003.0049.

53. Facial recognition: the CNIL orders CLEARVIEW AI to stop reusing photographs available on the Internet | CNIL, https://www.cnil.fr/en/facial-recognition-cnil-orders-clearview-ai-stop-reusing-photographs-available-internet, last accessed 2022/12/05.
54. GDPR, Article 17(2).
55. Frantziou, E.: Further Developments in the Right to be Forgotten: The European Court of Justice's Judgment in Case C-131/12, Google Spain, SL, Google Inc v Agencia Espanola de Proteccion de Datos. Human Rights Law Review. 14, 761–777 (2014). https://doi.org/10.1093/hrlr/ngu033.
56. Google LLC. v. Audiencia Nacional (Spanish), https://globalfreedomofexpression.columbia.edu/cases/google-llc-v-audiencia-nacional/, last accessed 2023/02/02.
57. NT1 v Google LLC, NT2 v Google LLC EWHC 799 (QB) (UK). (2018).
58. Kelly, M., Satola, D.: The right to be forgotten. University of Illinois law review. 2017, 1–64 (2017).
59. How do we ensure individual rights in our AI systems?, https://ico.org.uk/for-organisations/guide-to-data-protection/key-dp-themes/guidance-on-ai-and-data-protection/how-do-we-ensure-individual-rights-in-our-ai-systems/, last accessed 2023/01/27.
60. Song, C., Shmatikov, V.: Auditing Data Provenance in Text-Generation Models, http://arxiv.org/abs/1811.00513, (2019). https://doi.org/10.48550/arXiv.1811.00513.
61. Pyrgelis, A., Troncoso, C., De Cristofaro, E.: Knock Knock, Who's There? Membership Inference on Aggregate Location Data, http://arxiv.org/abs/1708.06145, (2017). https://doi.org/10.48550/arXiv.1708.06145.
62. Izzo, Z., Smart, M., Chaudhuri, K., Zou, J.: Approximate Data Deletion from Machine Learning Models: Algorithms and Evaluations. (2020).
63. Ginart, A., Guan, M., Valiant, G., Zou, J.Y.: Making AI Forget You: Data Deletion in Machine Learning. 14 (2019).
64. Bourtoule, L., Chandrasekaran, V., Choquette-Choo, C.A., Jia, H., Travers, A., Zhang, B., Lie, D., Papernot, N.: Machine Unlearning. arXiv:1912.03817 [cs]. (2020).
65. Sarker, I.H.: Machine Learning: Algorithms, Real-World Applications and Research Directions. SN COMPUT. SCI. 2, 160 (2021). https://doi.org/10.1007/s42979-021-00592-x.
66. Ullah, E., Mai, T., Rao, A., Rossi, R., Arora, R.: Machine Unlearning via Algorithmic Stability, http://arxiv.org/abs/2102.13179, (2021). https://doi.org/10.48550/arXiv.2102.13179.
67. Thudi, A., Jia, H., Shumailov, I., Papernot, N.: On the Necessity of Auditable Algorithmic Definitions for Machine Unlearning, http://arxiv.org/abs/2110.11891, (2022). https://doi.org/10.48550/arXiv.2110.11891.
68. Gupta, V., Jung, C., Neel, S., Roth, A., Sharifi-Malvajerdi, S., Waites, C.: Adaptive Machine Unlearning, http://arxiv.org/abs/2106.04378, (2021). https://doi.org/10.48550/arXiv.2106.04378.

69. Guo, C., Goldstein, T., Hannun, A., van der Maaten, L.: Certified Data Removal from Machine Learning Models. arXiv:1911.03030 [cs, stat]. (2020).
70. Jagielski, M., Thakkar, O., Tramèr, F., Ippolito, D., Lee, K., Carlini, N., Wallace, E., Song, S., Thakurta, A., Papernot, N., Zhang, C.: Measuring Forgetting of Memorized Training Examples, http://arxiv.org/abs/2207.00099, (2022). https://doi.org/10.48550/arXiv.2207.00099.
71. Goel, S., Prabhu, A., Kumaraguru, P.: Evaluating Inexact Unlearning Requires Revisiting Forgetting, http://arxiv.org/abs/2201.06640, (2022). https://doi.org/10.48550/arXiv.2201.06640.
72. Eisenhofer, T., Riepel, D., Chandrasekaran, V., Ghosh, E., Ohrimenko, O., Papernot, N.: Verifiable and Provably Secure Machine Unlearning, http://arxiv.org/abs/2210.09126, (2022). https://doi.org/10.48550/arXiv.2210.09126.
73. Liu, B., Ding, M., Shaham, S., Rahayu, W., Farokhi, F., Lin, Z.: When Machine Learning Meets Privacy: A Survey and Outlook, http://arxiv.org/abs/2011.11819, (2020). https://doi.org/10.48550/arXiv.2011.11819.
74. The GDPR, Article 18 - Right to restriction of processing. (2019).
75. Vollmer, N.: Recital 67 EU General Data Protection Regulation (EU-GDPR), https://www.privacy-regulation.eu/en/recital-67-GDPR.htm, last accessed 2023/02/02.
76. ICO fines facial recognition database company Clearview AI Inc more than £7.5m and orders UK data to be deleted, https://ico.org.uk/about-the-ico/media-centre/news-and-blogs/2022/05/ico-fines-facial-recognition-database-company-clearview-ai-inc/, last accessed 2023/01/31.
77. Greek DPA imposes 20M euro fine on Clearview AI for unlawful processing of personal data, https://iapp.org/news/a/greek-dpa-imposes-20m-euro-fine-on-clearview-ai-for-unlawful-processing-of-personal-data/, last accessed 2023/01/31.
78. Facial recognition: Italian SA fines Clearview AI EUR 20 million | European Data Protection Board, https://edpb.europa.eu/news/national-news/2022/facial-recognition-italian-sa-fines-clearview-ai-eur-20-million_en, last accessed 2023/01/31.
79. Clearview AI breached Australians' privacy, https://www.oaic.gov.au/updates/news-and-media/clearview-ai-breached-australians-privacy, last accessed 2023/01/31.