HUMANS FORGET, MACHINES REMEMBER: ARTIFICIAL INTELLIGENCE AND THE RIGHT TO BE FORGOTTEN

Eduard Fosch Villaronga, Peter Kieseberg, Tiffany Li [*]

"Nothing fixes a thing so intensely in memory as the wish to forget it" – Montaigne

INTRODUCTION

Popular society is often still impressed at the pace of new artificial intelligence (AI) advancements: In 1996, IBM's Deep Blue beat a reigning world champion in a game of chess.[1] Twenty years later, Google's AlphaGo beat a Grandmaster at Go, a game long considered to be a challenge too complex and difficult for AI.[2] artificial intelligence success at mastering Go is only one small example of the great strides AI technologies have made in the past few decades, but it is a sign of the exponentially increasing power and importance of AI in human society. Artificial Intelligence is rapidly developing, and it is necessary for lawmakers and regulators to keep up with the pace of this new and increasingly important technology.

Unfortunately, our current laws[3] are not fit to handle the complexities and challenges of artificial intelligence. One area in which current law is insufficient is privacy regulation.

While it may be easy to dismiss legal questions of AI and privacy as mere iterations of Easterbrook's "law of the horse,"[4] artificial intelligence fundamentally changes our current understanding of privacy because much of what scholars conceive to be privacy today rests on an understanding of how humans process information – especially, how humans remember and forget. This deficiency in understanding is especially apparent when considering the privacy law concept of the "Right to be

---

[*] Authors are in surname alphabetical order. Eduard Fosch Villaronga is Postdoctoral Researcher in Law, Governence and Technology at the University of Twente (Enschede, Netherlands); Peter Kieseberg is Senior Researcher at SBA Research (Vienna, Austria); Tiffany Li is Affiliate Scholar with Princeton University's Center for Information Technology Policy (Princeton, New Jersey, USA) and Fellow with the Internet Law and Policy Foundry (Washington, D.C., USA). The authors thank the participants of the Internet Law Works-in-Progress Conference 2017 for their helpful feedback.

[1] https://www.theatlantic.com/technology/archive/2016/02/when-computers-started-beating-chess-champions/462216/

[2] https://www.wired.com/2016/03/googles-ai-wins-fifth-final-game-go-genius-lee-sedol/

[3] This Article focuses on the laws of the E.U., due to the comparatively advanced nature of privacy laws in the E.U. as compared to that of other jurisdictions.

[4] *Easterbrook, Frank H. (1996). "Cyberspace and the Law of the Horse" (PDF). University of Chicago Legal Forum.*

Forgotten."

The Right to be Forgotten has risen to prominence alongside the rising importance of privacy law in general, particularly as understood in regulations like the European Regulation 679/2016 on Data Protection, (the "General Data Protection Regulation" or "GDPR")[5]. The Right to be Forgotten is essentially the concept that individuals have the right to request that their data (collected by others) be deleted. This concept of "data deletion" has come to the forefront of many juridical discussions of the Right to be Forgotten.

While "data deletion" may seem to be a straightforward topic from the point of view of many regulators, this seemingly simple issue poses many practical problems in actual machine learning environments. In fact, "data deletion" requirements can be considered to actually border on the edge of impossibility.

The problem with the Right to be Forgotten and its inapplicability to AI may be due to our inaccurate understanding of privacy in relation to AI. People often view privacy as, metaphorically, hiding their information from others. This is especially apparent when examining the principle of the Right to be Forgotten, under which individuals can request that information made public be deleted (and thus, made private). In the case of public information that is made private, the metaphor of a human mind forgetting a piece of information applies well. When individuals make previously-public information private, they metaphorically request that others forget that information. However, this metaphor is unique to human minds only and does not necessarily translate to the AI/machine learning era.

To understand the Right to be Forgotten in context of artificial intelligence, it is necessary to first delve into an overview of the concepts of human and AI memory and forgetting. Our current law appears to treat human and machine memory alike – supporting a fictitious understanding of memory and forgetting that does not comport with reality. (Some authors have already highlighted the concerns on the *perfect remembering*.[6])

This Article will examine the problem of AI memory and the Right to be Forgotten, using this example as a model for understanding the failures of current privacy law to reflect the realities of AI technology.

First, this Article analyzes the legal background behind the Right to be Forgotten, in order to understand its potential applicability to AI, including a discussion on the antagonism between the values of privacy and transparency under current E.U. privacy law. Next, the Authors explore whether the Right to be Forgotten is practicable or beneficial in an AI/machine learning context, in order to understand whether and how the law should address the Right to Be Forgotten in a post-AI world. The Authors

[5] 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)

[6] Mayer-Schönberger, V. (2011). *Delete: The virtue of forgetting in the digital age*. Princeton University Press.

discuss the technical problems faced when adhering to strict interpretation of data deletion requirements under the Right to be Forgotten, ultimately concluding that it may be impossible to fulfill the legal aims of the Right to be Forgotten in artificial intelligence environments. Finally, this Article addresses the core issue at the heart of the AI and Right to be Forgotten problem: the unfortunate dearth of interdisciplinary scholarship supporting privacy law and regulation. While this Article approaches that larger systemic deficiency through a contrasting legal and technical analysis of the Right to be Forgotten, the Authors' ultimate goal is to encourage greater interdisciplinary research in all facets of privacy law as applied to new technologies, particularly including artificial intelligence.

I. LEGAL ANALYSIS OF THE RIGHT TO BE FORGOTTEN

A. A Brief Legal History of the Right to be Forgotten

The legal history of the Right to be Forgotten can be said to have begun in 2010. That year, a Spanish citizen (together with the Spanish National Data Protection Agency) sued both a Spanish newspaper and Google, Inc. The Spanish citizen argued that Google was infringing on his right to privacy, due to the fact that Google's search results included information relating to a past auction of the man's repossessed home. The plaintiff requested that his information be removed from both the newspaper and from Google's search engine results.

Representatives for Google explained that even if the company could censor certain search results, as it had done in, for example, Google China, the censored information would still remain in the original websites from which the Google results were created. Google effectively argued that they were data processors and not data controllers (two distinct classes with much different privacy obligations under E.U. privacy law).

Ultimately, the Division of Administrative Law of the Spanish National Court agreed to submit to the European Court of Justice (ECJ) a question of interpretation regarding certain provisions of the Data Protection Directive from 1995 on the protection of personal data. The questions were: 1) whether the Data Protection Directive applied to search engines; 2) whether the EU Law applied to Google Spain if the server was in the United States; 3) and whether a data subject could request to have his/her data removed from accessibility via search engines.

In 2014, the ECJ ruled in favor of the Spanish citizen (C-131/12).[7] The court stated that, according to the Art. 4.1 a) of the Data Protection Directive 95/46 EC[8], the

[7] Court of Justice of the European Union (2014) C-131/12 Google Spain SL, Google Inc v. Agencia Española de Protección de Datos (AEPD), Mario Costeja González. Available at: curia.europa.eu/juris/document/document.jsf?text=&docid=152065&doclang=EN

[8] Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of

European Data Protection Directive applies to search engine operators if one or more of the following three conditions are met: 1) if they have a branch/subsidiary in a Member State[9] which promotes the selling of advertising space offered by the search engine to the inhabitants of that Member State; 2) if the parent company designates a subsidiary company in a Member State and it is responsible of two filling systems concerning data from the data subjects of such Member State; or 3) if the branch/subsidiary forwards to the non-EU parent company located outside the EU any requests and requirements from the data subjects or from authorities in charge of surveilling the data protection right even if these forwards are engaged in voluntarily.

As long as at least one of these conditions is met – in the aforementioned case, it was the first condition that the Court deemed Google to have met – the Court deemed this sufficient to qualify the search engine company as a data controller. As data controllers, the national laws that pursue the objectives of the directive 95/46/EC would fully applies to the search engine companies. For the Google Spain case, this meant that the Court affirmed the right of data subjects to ask search engine companies to remove links that contained personal information about the data subjects. The Court stated that: 1) the removal of data could be required under certain conditions, e.g., when the information is inaccurate, inadequate, irrelevant, or excessive for the purposes of the data processing; and 2) that the right was not absolute and needed to be balanced with other compelling rights such as the freedom of expression.[10]

B. Current Regulatory Definitions for the Right to be Forgotten

Although referring to the applicability of the national law that transposed the Directive 95/46/EC to search engines, the C-131/12 ruling constituted the basis of a new understanding of the territorial scope of the European data protection rules, the Right to be Forgotten, and the applicability of the EU data protection rules to a search engine[11]. Recognizing the existence of the Right to be Forgotten, the C-131/12 identified a general principle which until now was scarcely mentioned in the data protection directive 95/46/EC.

Indeed, the Right to be Forgotten is not hundred percent new. The data protection directive 95/46/EC already contained the "right of access" on its Article 12, which somehow already contemplated the possibility to enforce the erasure of incomplete, inaccurate or illegal data from the data controller. Currently the new GDPR has included the right to erasure on its Article 17,, which refers to the right of the data subject "to obtain from the controller the erasure of personal data concerning him or her without undue delay". According to the European Commission, art. 7 GDPR

---

such data

[9] Member Sates are nations that are party to the E.U. Data Protection Directive.

[10] *See* europa.eu/rapid/press-release_MEMO-15-6385_en.htm

[11] Factsheet on the Right to be Forgotten ruling (C-131/12). Cfr.: http://ec.europa.eu/justice/data - protection/files/factsheets/factsheet_data_protection_en.pdf

strengthens the principle and improves legal certainty.[12]

This right is an obligation for the controller who shall erase the personal data without undue delay when a) the personal data is no longer necessary in relation to the purposes for which it was collected or otherwise processed; b) the data subject withdraws the consent on which the processing is based on a given consent for a specific purpose, or on a given explicit consent for special categories of data and where there is no other legal ground for the processing; c) the data subject objects to the processing – pursuant his/her right to object – and there are no overriding legitimate grounds for the processing, or the data subject objects to the processing for direct marketing purposes; d) the personal data has been unlawfully processed; e) the personal data has to be erased for compliance with a legal obligation in Union or Member State law to which the controller is subject; f) the personal data has been collected in relation to the offer of information society services referred to in Article 8(1) GDPR, i.e. of a child.

In the light of public data erasure obligation, and taking into account the state of technology and the cost of the implementation, the controller has the obligation to take reasonable steps – including technical measures – to inform the processors of the data subject request and to delete any link or copy or replication of such personal data.

The GPDR states exceptions for such deletion including the exercise of the freedom of expression and information; when the data in question is processed due to a legal obligation under the EU or Member State law; for reasons of public interest in the area of public health, public interest, scientific, historical research or statistical purposes which would be rendered practically impossible to achieve their objectives without the processing of this data; or when such processing involves the establishment or defense of legal claims.

Formally recognizing this right in the GDPR signifies democratizing something private companies were exploiting at users' expenses.[13] It also results in a more general obligation towards the protection of privacy that was only covered partially by some existing sector-specific rights – e.g. bankruptcy law already offered debtors a fresh start through the forgiveness of debts,[14] and criminal law was already in favor of the expunction of criminal sentences too (which are the equivalent of "never convicted.")[15] This means that the principle of equality is ensured irrespectively of the context;[16] which implies the abolition of certain discriminations, for instance the

---

[12] Ibidem.

[13] *See* the website www.reputation.com

[14] Reifner, U. et al. (2003) Consumer Over indebtedness and Consumer Law in the European Union. Final Report. Presented at the Commission of the European Communities, Health and Consumer Protection Directorate-General. Available at: www.ecri.eu/new/system/files/26+consumer_overindebtedness_consumer_law_eu.pdf

[15] Jacobs, J.B. and Laurrari, E. (2015) Expungement of Criminal Records in Europe. Collateral Consequence Resources Center. Collateral Consequences of Criminal Conviction and Restoration of Rights: News, Commentary, and Tools.

[16] Equal Rights Trust (2008) Declaration of Principles on Equality. Available at:

"offender-type discrimination", i.e. in some countries sexual offenders were not eligible for the expungement of their sentences.[17]

C. Legal Controversies Regarding the Right to Be Forgotten

There are already some notable concerns with the Right to be Forgotten – both in theory and in implementation:

The main problem concerning the Right to be Forgotten (RTBF) lies in the clash between the good intentions of the regulators – written from an abstract point of view – and the actual complexity of real-life technical environments. The vagueness of the Article's definition, however, rubs the impossibility of its application: the Article seems to push towards the simple deletion of the personal data or the folder containing the personal data from the data controller's system, as if data on a computer was like a physical file that can simple be destroyed. Interestingly, the word "deletion" does not appear in the GDPR; and the word "remove" only appears twice but does not refer to the RTBF. The word used for the GDPR to refer to *deletion* is "erasure" and it is not explained throughout the text.

In the light of the exceptions – for reasons of public interest in the area of public health, public interest, scientific, historical research or statistical purposes – some authors believe that the problem lies on determining what information may have value in the future.[18] Ambrose argues that the immediate value and the remote value of the information play a major role in shaping the difficulties associated with the enforcement of this right, e.g. he claims that it can be dangerous in scenarios involving people running for political officers, for instance.

Any misunderstanding concerning the RTBF might not matter in the light of a GDPR infringement. Penalties for non-compliance with the GDPR reach up to 4% of the undertaking annual revenue or include fines up to EUR 20 million. Moreover, those in the organization in charge of personal data protection can be criminally liable.

Regulatory fines aside, one larger question that is somewhat outside the scope of this Article is whether the Right to be Forgotten matters – that is, whether there is a political, sociological, or moral need to protect the Right to be Forgotten. It seems so, according to the GDPR which is based on the fundamental right to data protection of the European Charter of Fundamental Rights. However, this assumption is not without defensible challenges, particularly from the free speech community. The United States, which could be considered by some to be an international free speech country, does not legally recognize the Right to be Forgotten. U.S. civil liberties advocates and technology corporations have also fought against similar rulings.

The legal discussion seems not to be of help either. In December 2015, the

---

www.equalrightstrust.org/content/declaration-principles-equality Accessed 26 July 2017.

[17] Ibidem.

[18] Ambrose, M. L. (2013). It's About Time: Privacy, Information Life Cycles, and the Right to Be Forgotten. *Stan. Tech. L. Rev.*, *16*, 369.

European Commission announced that while an individual that has given his/her consent to processing for a specific purpose has the right to get his/her data removed from the system when s/he does not want it processed anymore, still, "this does not mean that on each request of an individual all his personal data are to be deleted at once and forever"[19]. The European Commission argues that retention of the data may be allowed for contract performance or for legal compliance reasons, and that data can be kept as long as it is necessary for that purpose.[20] It is not surprising, therefore, that without any other clarification legal scholars and engineers are confused by the extent of such right.

Some legal scholars see the main problem of the RTBF with the freedom of expression, of media and other compelling rights. Rosen believes that unless the right is defined more clearly, this right will make the gap between the understanding of privacy and freedom of speech between Europe and United States even wider, beyond the possibility that it will lead to a less open Internet.[21] To that, the European Commission argues that, in theory, the Right to be Forgotten is about protecting the privacy of the individuals not about erasing past events or restricting freedom of press.[22]

II. TECHNICAL BACKGROUND FOR DATA DELETION IN MACHINE LEARNING

A. Understanding Human and AI Memory

To understand the Right to be Forgotten in context of artificial intelligence, it is necessary to first delve into an overview of the concepts of human and AI memory and forgetting. Our current law appears to treat human and machine memory alike – supporting a fictitious understanding of memory and forgetting that does not comport with reality. (Some authors have already highlighted the concerns on the *perfect remembering*.[23])

Cognitive psychologists believe there are two primary systems of memory in the human mind: a short-term memory and a long-term memory[24]. However, there is not yet consensus on what the major differences between the two are. What gets stored in long-term memory may depend on multiple factors, including the meaningfulness of the memory.[25] It's not quite clear what those factors (including "meaningfulness") are. In fact, there is not even a solid agreed-upon estimate of how much raw data a

[19] *See* europa.eu/rapid/press-release_MEMO-15-6385_en.htm

[20] Ibidem.

[21] Rosen, J. (2011). Free speech, privacy, and the web that never forgets. *J. on Telecomm. & High Tech. L.*, *9*, 345.

[22] *See* europa.eu/rapid/press-release_MEMO-15-6385_en.htm

[23] Mayer-Schönberger, V. (2011). *Delete: The virtue of forgetting in the digital age*. Princeton University Press.

[24] https://www.scientificamerican.com/article/why-do-we-forget-things/

[25] https://www.scientificamerican.com/article/why-memory-so-good-bad/

human mind can store.[26] Thus, it is sufficient it to say that our current understanding of the human mind and memory is nascent at best, flagrantly incorrect at worst.

In contrast, scholars do know how "minds"[27] in the world of artificial intelligence work, if for no other reason than that human beings are the ones who create the logical processes behind artificial intelligence. To be sure, individual artificial intelligence systems may design their own processes without direct human instruction (and will likely do so increasingly more in the future). Indeed, much has been said about the "black box" nature of self-instructing AI and the difficulty of understanding advanced artificial intelligence decision-making[28]. However, at the core, computer scientists still know, generally, what the foundations of artificial intelligence decision-making are, and they know this arguably better than neuroscientists understand the foundations of human decision-making. At minimum, there is strong scientific understanding of how AI treats data input, storage, and deletion.

In brief, while scholars may not fully understand the decision-making process of a specific AI, it is possible to understand, generally, how AI "minds" work – at least in context of data input, storage, and deletion. Understanding the difference between human and AI "memory" provides a greater understanding of the deficiencies of current privacy law, particularly related to the Right to be Forgotten.

B. A Technical Analysis of AI Data Deletion ("Forgetting")

As discussed, the Right to be Forgotten requires the deletion of previously public data. Essentially, the Right to be Forgotten applies the human memory metaphor of "forgetting" information. When individuals request that their personal information be deleted, this is equivalent to metaphorically requesting that others forget that information. However, this metaphor is unique to human minds only and does not translate to the AI/machine learning era.

Specifically, the Right to be Forgotten requirements of data deletion do not easily translate because AI does not "forget" data in the way that humans do. Data deletion in artificial intelligence contexts is much more complex.

The first aspect of deletion in machine learning focuses on the question whether deletion is actually possible in modern data-driven environments. We argue that data removal is actually extremely complex in current systems. We will illustrate the fundamental problems of the technical implementation of the Right to be Forgotten with the example of a modern relational database management system (DBMS), from

[26] http://www.bbc.com/future/story/20150401-whats-the-most-we-can-remember

[27] The Authors understand that there is some controversy regarding the semantics of AI "mind" and "memory." This Article does not debate the fitness of these terms but uses them as convenient metaphors for understanding legal concepts based on human mind and memory.

[28] https://www.nature.com/news/can-we-open-the-black-box-of-ai-1.20731

now on simply referred here to as “database”.

Databases are programs designed for the efficient provisioning of data, where the term “efficient” is typically referred to as the speed with which data can be searched for. While the problem of efficient searching seems trivial when thinking about small amounts of data, it is one of the fundamental classes of algorithms in computer science and one of the oldest applications besides solving mathematical problems. Relational databases typically work by indexing data, i.e. the data records are stored on the disk inside files, but the layout of this file is structured in the form of a (mathematical) B-Tree (more precisely, a B^+-Tree is typically used). Trees are data structures that are very search-efficient and allow fast retrieval of information. Furthermore, in addition to the tree structure that defines the physical location of the information on the disc, additional search indexes can be constructed to allow to speed up specific search queries. Of course this navigation through the search trees is not conducted by the user directly, but by using an interface, e.g. the SQL querying language for explicitly defining the data records that should be retrieved from the databases. Modern databases are able to support searches in data sets containing billions of records using rather affordable hardware.

There exist some requirements for real-life databases that have direct effects on the problem of data removal. They are normally called ACID, the acronym for atomicity, consistency, isolation and durability. These terms are outlined thusly:

- ***Atomicity***: Atomicity means that a set of operations is done either as a whole or not at all. Example: Insert of a data record needs to be done for the whole record or not at all, just adding half of the record is inacceptable (the same is asked for whole sets of records being added/removed). This is especially interesting with respect to database crashes during operation, here mechanisms must be in place that undo the incomplete transactions and get back to the last state. This is called a “rollback”.
- ***Consistency***: After an operation is finished that database must be back in a consistent state, meaning that all relations are unambiguous and the database is normalized.
- ***Isolation***: In case of parallel transactions, the database must ensure that they do not interfere with each other. This is typically done via “locking”, i.e. the data to be changed is marked as locked and cannot be touched by other operations until the first operation is finished.
- ***Durability***: Data must be stored permanently in the database, especially considering system errors or server crashes. Especially crashes must not result in data loss or inconsistencies. Mechanisms like transaction logs are typically used in order to support this requirement.

Databases following them being called “ACID-compliant”. All major databases currently in use are ACID-compliant. On another side, users of databases typically expect additional features from the database in order to provide a usable environment:

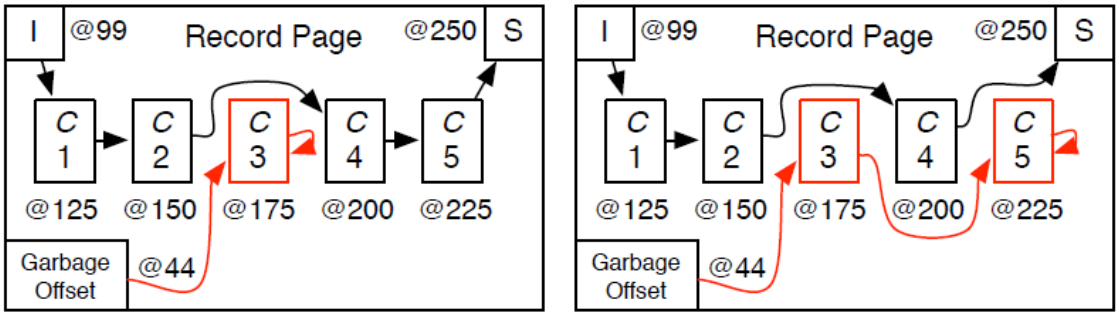
- ***Efficient operation***: Especially retrieval must be done as fast as possible, typically resulting in the fact that operations that are not required to be done at once are

postponed to later times. This is also an aspect regarding data deletion that is discussed later in this Article.

- ***Intended Rollbacks***: In addition to rollbacks done in order to mitigate errors introduced by crashes, the database needs to have enough history stored on previous states in order to be able to roll back in time for a certain amount of transactions (typically, there exist transactions that cannot be rolled back though).
- ***Audit & Control***: Many regulations, but also internal security policies, require mechanisms that provide transparency, i.e. mechanisms that make it possible to control when, which data was changed and by whom, at what time, and through which action.
- ***Replication and Backups***: Protection against negative effects from disasters is a fundamental requirement in most modern IT-systems. Replication techniques guarantee having several instances of the same database, with the same data content dynamically updated, spread across a (possibly large) geographical area.

From the requirements outlined above it is clearly visible that every data record added to the database might not only reside at one specific point in the file system, but might be stored at various locations inside internal database mechanisms, as well as across different replicated databases, in log-files and backups. When the Right to be Forgotten asks for permanent deletion of the data, these requirements must be taken into account. When asking for deletion in a strict sense, these spaces must be identified and overwritten with random information. In several internal mechanisms like the database transaction log, the latter is especially impossible without seriously endangering the consistency of the database, or even simply breaking it altogether.

Leaving aside the issue of internal mechanism, since it can always be discussed whether adhering to the GDPR includes making the deleted information safe against recovery in high-profile forensic investigations targeting volatile internal mechanisms, the principle method of deletion must be analyzed. In actually all relevant databases, when a record is deleted via the (SQL) interface, it is not overwritten with data or the space filled up with zeroes, it is only marked as deleted and removed from the search indexes. This hearkens back to the issues of performance, actually deleting and overwriting space would be a tremendous additional effort with serious impact on the actual performance (it would mean that a delete-request would be more expensive than several data insertions).



(a) One deleted record (b) Two deleted record

Figure 1 – Deletion in MySQL[29]

We demonstrate this issue with respect to our example, the MySQL-Database (see Figure 1):

- Figure 1 (a) shows the state of our database before deletion. While most parts of the picture are given for accuracy, the important parts for the deletion are the five data records themselves, sitting in the spaces C1 to C5, as well as the start (I) and the end (S) of this part of the database. Furthermore, the figure shows one deleted record at C3 which is linked by the so-called "Garbage Offset", a collection of deleted and now free space. In order to reduce complexity, this figure does not show the whole search tree, but only a small part of it, a so-called "page".
- When the database is searching for data, it locates the page inside the search tree, where the needed information must reside. Inside the page, it starts at node I and follows the path of arrows, until the required data is found. If the search ends at node S without any result, the data was not found.
- The task in our example is the removal of the data stored in C5.
- The database searches for the data in C5 and navigates through the tree until C5 is found.
- The space is now "marked for deletion":
- The arrow pointing to C5 is bent in order to show to the node after C5 (in this case node S), the arrow pointing from C5 is bent in order to refer back to C5.
- C5 is then added to the garbage offset by bending the arrow from C3 to show to C5.
- Effectively, C5 is moved from the list of active records to the list of deleted records indicated by the garbage offset. The data is still stored in the database, but when the database requires space for storing a new record, the list started by

[29] Fruhwirt, Peter, Peter Kieseberg, and Edgar Weippl. "USING INTERNAL MySQL/InnoDB B-TREE INDEX NAVIGATION FOR DATA HIDING." In IFIP International Conference on Digital Forensics, pp. 179-194. Springer International Publishing, 2015.

the garbage offset can be searched for suitable space to overwrite, instead of allocating new space on the disk.

As illustrated, the data is thus not really deleted, it is solely removed from the search index. In reality, it can take a long time until the deleted space is reused again (which effectively destroys the old data inside it), as databases often rather append new data than searching for existing free space due to performance issues

With this background, it is thus necessary to discuss what the term "deletion" of the Right to be Forgotten is actually referring to the simple removal from the search index, overwriting in the file system, deletion from log-files and backups, or even removal from all internal mechanisms. Depending on the actual requirements, deletion might become infeasible in real-life environments operating under economic principles.

III. Does the Right to Be Forgotten Make Sense for AI?

The last section discussed the problems of actual forgetting (a.k.a. deletion) in modern data driven environments, especially when they are based on databases, which is more or less normal with a given size. This section assumes that a legal way of obeying the Right to be Forgotten has been established in order to simplify discussion on the issues with respect to machine learning.

One major question lies in analyzing the effects of data removal on the quality of the results. This is especially interesting considering algorithms that use a so-called "knowledge-base" for calibration, i.e. the algorithm takes the knowledge-base with pre-calculated results as reference data and extracts the common artifacts. It then uses this "learned" rules on new data, which has to be very close to the training data in terms of data structure and statistical properties. Furthermore, the resulting categorizations are again fed into the knowledge base in order to get even better training data for the next run, thus iteratively extending the knowledge base. In a recent work[30], the effects of deletion on a set of prominent categorization algorithms were studied. Summarized, the deletion of single data points did not have any large-scale effects. Still, it must be noted that in these experiments the data points scheduled for removal had been selected at random, which could be different in real-life cases where it could be the case that people that want their data to be removed share some commonalties that are then missing from the data set at all.

Another major research questions targets methods for removing the need for deletion altogether by changing the underlying data in a way to make it less sensitive enough not to need deletion anymore. Currently, several approaches exist, but none of

[30] Bernd Malle, Peter Kieseberg, Edgar Weippl, and Andreas Holzinger, "The Right to Be Forgotten: Towards Machine Learning on Perturbed Knowledge Bases", Workshop on Privacy Aware Machine Learning (PAML), August 2016

them is fit enough to be used in real-life applications:

The most practical, using trusted environments, has the problem that according to the GDPR the data analyst needs to have informed consent of the data owners for each analysis, which is highly impractical. Furthermore, this approach runs into problems when considering shared environments or the use of Cloud-computing for better performance.

Another idea lies in the utilization of functional encryption. Functional encryption algorithms are secure cryptographical functions that establish an isomorphism, i.e. it is possible to perform mathematical operations on the encrypted data without being able to decrypt it. Let $F(x)$ be the encryption function, and $F^{-1}(y)$ the decryption function, then for functional encryption it holds true that $x+y = F^{-1}(F(x) + F(y))$. While solving many issues regarding data privacy in theory, in practice all algorithms known today are simply far too inefficient to be used even on data sets of moderate sizes, not to mention the area of big data.

Pseudonymization works by exchanging sensitive attributes for placeholders before the calculation. While this is practical, according to the GDPR pseudonymized data needs to be treated just as the original sensitive data, thus nothing is won from a legal perspective.

Anonymization works by transforming the original data set into a derivative form that blurs the sensitive information enough to make the user unidentifiable. Typical methods like k-anonymity work by generalizing the sensitive attributes to a point, where they are no longer sufficient for the identification of persons. This is also not a technologically strong solution. In related experiments in the same work as the one analyzing the effects of anonymization, the authors studied the effects of anonymization on machine learning algorithms, more precisely on popular classifiers. They used the very popular notion of k-anonymity. The main result was that anonymization using k-anonymity resulted in quite significant distortions in the results, even when considering a very low security/privacy margin ($k=3$ or 4), and made the results practically worthless when using higher privacy margins. The analysis was performed on the same data set as the analysis on deletion, making these results comparable.

In conclusion, the implementation of the Right to be Forgotten has a serious impact on machine learning environments. In order to circumvent the problem of deletion in complex data centered systems, more research in the area of privacy aware machine learning (PAML), i.e. in the development of algorithms resilient against the effects of anonymization, has to be conducted. Future revisions of the GDPR or future explanations from competent authorities – e.g. European Data Protection Supervisor – should consider the technical side of information systems to ensure an adequate balance between the wording of the Law and its applicability. Law-makers have to be aware of this fact and of the impact on innovative products and services this can have especially when considering the economic race with other countries spurring far less restrictive data protection laws like the US and China.

IV. POTENTIAL SOLUTIONS AND FURTHER QUESTIONS

Now, assuming that the Right to be Forgotten is a valuable right that should be protected, and given that this Article proves the impracticability of applying the Right to be Forgotten in machine learning environments, one can now turn to whether legal or policy solutions could exist to protect this legal right (or at least the underlying spirit of it).

While a comprehensive index of all potential solutions is beyond the scope of this Article, we believe that there are indeed some potentially effective and practical solutions that would protect the Right to be Forgotten, or a version of it, in machine learning/artificial intelligence environments.

A. Data Minimization

On the most basic level, if companies or governments do not collect certain personal information, there will be no information to forget. To put it in another way, not existing is the simplest way to be forgotten. (As the saying goes, if a tree falls in a forest, but no one is around.…)

Collecting less data, or "data minimization," has already been strongly advocated for by many privacy advocacy groups. Yet, companies continue to collect ever greater amounts of information. Simply telling data controllers to collect less data is likely not a viable solution. Perhaps the competent authorities could encourage such data minimization by issuing guidance documents to help inform data controllers and ease the processes involved. As an example, the Dutch data regulator issued a guidance document on how to copy information from identification documents. The Dutch data regulator advised that data controllers need to cover the photo and the unique personal Dutch number (called BSN).[31] This guidance was effective in minimizing the collection of those pieces of data. Now, many companies use a mold to cover that relevant information when scanning the identity document.[32]

B. Innovative Technical Solutions

As we examined in our analysis of data deletion techniques, there are different methods to "delete" data in AI environments. Current law does not make clear which forms of deletion would be sufficient for legal requirements under the Right to be Forgotten. We have also addressed the impossibility or impracticability of some

[31] T. Jocnckheer (2012) Copying ID Documents – Dutch Data Regulator Issues Guidance. Cfr.: http://www.privacyandcybersecuritylaw.com/copying-id-documents-dutch-data-regulator-issues-guidance

[32] T-Mobile Privacy Statement (in Dutch) Cfr.: http://www.t-mobile.nl/global/media/pdf/privacy-statement.pdf

concepts of data deletion or AI "forgetting." The law also does not address whether this impossibility or impracticability would allow for alternative methods to suffice.

More interdisciplinary research is needed to understand the full spectrum of options for implementing the Right to be Forgotten or at least for following the spirit of the law.

Many innovative solutions exist that do not necessarily deal strictly with deletion of data. For example, data controllers could make sure to collect data in more disparate ways, storing different types of data in siloes so that they cannot be recombined to identify a person. Techniques to guard against re-identification could be just as effective as directly deleting data, if effectiveness is understood as complying with the letter or spirit of the law.

Differential privacy is another new privacy-supporting technology that could help protect personal information in ways that could reflect the intentions of the regulators and Courts in creating and enforcing the Right to be Forgotten.

C. Integrated Technological and Legal/Policy Solutions

In his essay "Recoding Privacy Law: Reflections on the Future Relationship Among Law, Technology, and Privacy," Urs Gasser argues for the potential benefits of "how law and technology can advance the state of the practice through a *mutually productive relationship*."[33] Gasser writes:

> "Taken together, the development of privacy tools that aim to *integrate* legal and technical approaches could help pave the way for a more strategic and systematic way to conceptualize and orchestrate the contemporary interplay between law and technology in the field of information privacy. … Such integrated approaches recognize the rich roles that law can play alongside the technical space and hint at how more robust and effective privacy protections can emerge by melding different instruments and methods — both at the conceptual and implementation levels."[34]

Integrated privacy tools may be ideal solutions to the problem of privacy laws that do not reflect the actualities of current technologies, particularly artificial intelligence.

While the technology behind these integrated privacy tools is still nascent, there have been few legislative developments in the data protection realm that address the so-called privacy enhancing technologies (PETs). Increasingly, more initiatives in the technical world are noting the need to develop legal instruments that can validate the advancements on the technical side that aim at protecting privacy. Regarding databases and other complex systems, there arises the need for integrated solutions that support the physical deletion of information, while retaining e.g. ACID-compliance, as these features are indispensable for the correct functioning of large data driven

[33] Urs Gasser. 2016. "Recoding Privacy Law: Reflections on the Future Relationship Among Law, Technology, and Privacy." Harvard Law Review

[34] Id.

environments.

There are also some innovative solutions that arise when technologists and legal/regulatory scholars work together directly to address these clashes. As one example, in 2012, Professor Jean Yang at Carnegie Mellon University, created a programming language called Jeeves that aims at enforcing data protection policies.[35] The program allows the programmer to write policy-agnostic programs, separately implementing policies on sensitive values from other functionality. Harvard's Berkman Klein Center for Internet Studies has also contributed substantial research to the topic of integrated privacy tools through the Harvard University Privacy Tools Project.[36]

What is uncertain is, whether these integrated solutions will be accepted as legally sufficient by regulators, Courts, and the legal community at large. There is still work to be done to prove the efficacy of these innovative new tools. For example, in the case of the Jeeves programming language, the legal community may ask relevant questions including: Does the language fully promote the awaited aspects in the next general data protection regulation? Could the use of this programming language be enough to protect privacy? Could this language solve the problems arisen by the use of machine learning techniques?

D. Law and Policy Solutions

Another route to bridging the gap between law and technology is to address the deficiencies in the law. This Article has addressed deficiencies in the both the regulations and case law surrounding the Right to be Forgotten, especially as those deficiencies relate to and/or are a result of the inapplicability of current understandings of "forgetting" to the actualities of data deletion in AI environments.

It may appear that the strongest solution would be to simply change the law to reflect new technologies like AI. However, one does not "simply" change law, especially large omnibus regulations like the EU GDPR. Changing, reforming, or updating regulations like the GDPR is an onerous process that takes years (if not decades). Regulations like the GDPR will likely always fall behind new advances in technology, if only because the pace of regulatory change is much slower than that of technological change.

Even in the absence of regulations that perfectly address all new technologies, there may still be legal or policy solutions. Regulators could provide guidance to

[35] One can read on her website: "It is increasingly important for applications to protect the privacy and security of data. Unfortunately, it is often non-trivial for programmers to enforce privacy policies. We have developed Jeeves to make it easier for programmers to enforce information flow policies: policies that describe who can see what information flows through a program. Jeeves allows the programmer to write policy-agnostic programs, separately implementing policies on sensitive values from other functionality". *See* projects.csail.mit.edu/jeeves/. *See* the academic paper Yang, J. et al. (2012) A Language for Automatically Enforcing Privacy Policies, POPL.

[36] See, generally, the work available on their website: https://privacytools.seas.harvard.edu/.

interpreting or complying with the regulations, and such guidance (or "standards") may be more able to accurately reflect the pace of new technology. A technical standard that could set the ground for some common terminologies and procedures could be the solution to the de-codification problem engineers have when they attempt to comply with legal regulations.

In context of the GDPR, one can look to standards like the ISO/IC 29134:2017 as examples of helpful regulatory guidance. The ISO/IEC 29134:2017 "Guidelines for privacy impact assessment" standard aims to provide much more detailed directions for the privacy impact assessment process and the structure of its report than the current Art. 35 GDPR. Standards help provide risk management assistance by limiting liability and helping producers meet market demands[37]. However, standards like these are considered soft law[38] and are not directly enforceable.

Soft legislation provides good alternatives for dealing with many international issues that are new, specific and complex, especially when States cannot foresee the consequences of a legal document. Standards are flexible, seen as a tool of compromise, and sometimes the basis of legal corpuses such as the Machinery Directive 2006/42/EC or the Medical Device Directive 93/42/EEC[39].

At the same time, however, soft legislation is non-binding and only voluntarily adopted. In the case of private standards, one could argue that these standards represent the capitalization or monetization of Law (due to their cost-intensive nature) and that such standards are merely self-interpretations of the industry reality. Leaving full room to standards to specify the content of the Right to be Forgotten, thus, will have to be carefully addressed.

Ultimately, a collaborative approach between the specificity of standards, and the *bindingness* of hard law (capacity for enforcement, consequences for violations, etc.)[40], should be preferred.

E. Questions for Further Research

There is great potential for new research on these issues and many questions that are yet to be answered. Here are only a few of the many questions still be to answered:

- How can we ensure a balance between the Right to be Forgotten and a machine learning model's need to remember information used to train it?

---

[37] Nelson, R. (2015) Robot Safety Standard Update. Presentation available at: www.robotics.org/userassets/riauploads/file/TH_RIA_Roberta_Nelson_Shea.pdf

[38] Shelton, D. (2003). Commitment and compliance: The role of non-binding norms in the international legal system. Oxford University Press

[39] See ec.europa.eu/growth/single-market/european-standards/harmonised-standards/index_en.htm See also Krut, R., and Gleckman, H. (2013). ISO 14001: A missed opportunity for sustainable global industrial development. Routledge

[40] Kenneth W. Abbott and Duncan Snidal (2000). Hard and Soft Law in International Governance. International Organization, 54, pp 421-456

- Can artificial intelligence models be created that "learn" from new data without storing personal information that could be used in a Right to be Forgotten request?

  How should we balance the competing interests of innovative data use and personal data privacy rights?
  Can we create a dynamic regulatory framework that manages to address all user' needs, perhaps on a case-by-case basis?
- On a broader level, how should regulators attempt to enforce laws that are not actually practicable with new technology?
- Do we need a new ontology to create a common understanding between technologists and policymakers?
- Is there a larger set of standards that can be created to address cases like these where law and technology do not agree?

V. CONCLUSIONS

The Right to be Forgotten is only one small aspect of current privacy laws. It is the Authors' intent that analyzing this one aspect of privacy law, in context of an actual technical analysis, will provide an example of the need for greater interdisciplinary work in this field.

The Right to be Forgotten may very well be a well-intentioned regulatory protection, and many would argue that it is an important right to be protected. However, there is a clear disconnect here between law and technical reality. Similar to what privacy researchers have seen in Privacy by Design implementation, it is difficult to implement and enforce legal requirements in data-processing systems.[41] Using two different languages in the legal and in the technical approach to concepts like data deletion leads to a problematic miscommunication that could have unfortunate consequences. It is necessary to bridge that divide in languages and understandings of concepts like "memory" and "forgetting."

As Vint Cerf, Internet pioneer, put it: "You can't go out and remove content from everybody's computer just because you want the world to forget about something. [That's not] a practical proposition at all."[42] This inability for machines to "forget" is especially true for large and complex systems, especially databases, where the ability to go back to an older state of the system, as well as to be able to give detailed information on past system states, is a vital requirement in order to be ACID-compliant and usable in practical applications. As shown, it may be impossible for AI to truly "forget" – at least in the context of the Right to be Forgotten

[41] Bert-Jaap Koops & Ronald Leenes (2014) Privacy regulation cannot be hardcoded. A critical comment on the 'privacy by design' provision in data-protection law, International Review of Law, Computers & Technology, 28:2, 159-17

[42] Warman, M. (2012) Vint Cerf attacks European internet policy. Telegraph. Available at: www.telegraph.co.uk/technology/news/9173449/Vint-Cerf-attacks-European-internet-policy.html

Throughout this Article, the Authors have repeatedly called for greater interaction between lawmakers and technologists in order to actually understand the technological consequences of privacy regulations, especially when concerning new technologies like artificial intelligence. We conclude by strongly advocating for greater interdisciplinary research into the technological aspects of artificial intelligence and implications for legal and regulatory privacy regimes.

This Article analyzed the legal and technological conceptions of the Right to be Forgotten, an analysis that in itself is an example of the strengths of interdisciplinary research in forming practical understandings of law and technology. However, in the spirit of advocating for greater interdisciplinary research, we recognize that greater interdisciplinary research can still be conducted, bringing in more perspectives in addition to the legal and technological. For example, in understanding the Right to be Forgotten, we must return once again to the neuroscientific understanding of memory with which we began this Article.

The Right to be Forgotten, as currently understood by Courts and regulators, relies on conceptions of how human memories function and how humans forget. As noted, these metaphors do not strictly apply to technologies like artificial intelligence. In this Article, we addressed this problem by looking at AI and the Right to be Forgotten from both a legal and technological lens. However, this may not be sufficient. To fully comprehend how to protect the spirit of the Right to be Forgotten, it may be necessary to also study this problem from the perspective of different fields, including neuroscience, cognitive science, anthropology, psychology, and sociology.

Artificial intelligence is rapidly developing, changing our society in ways we may not currently be able to predict. The law must keep pace with technology, and the best solutions to any gaps that develop between law and technology will likely be found through interdisciplinary research. Today, our understanding of AI is limited enough that we must rely on outdated metaphors like remembering and forgetting. Perhaps future scholars will have better metaphors, supported by greater interdisciplinary research. For now, we can only conclude by stating that the AI and Right to be Forgotten problem can be summed thusly: Humans forget, but machines remember.

* * *