INTERNATIONAL JOURNAL FOR LEGAL RESEARCH AND ANALYSIS

Open Access, Refereed Journal Multi Disciplinary
Peer Reviewed

www.ijlra.com

 ISSN: 2582-6433

DISCLAIMER

 ISSN: 2582-6433

EDITORIALTEAM

EDITORS

Dr. Samrat Datta

Dr. Samrat Datta Seedling School of Law and Governance, Jaipur National University, Jaipur.Dr. Samrat Datta is currently associated with Seedling School of Law and Governance, Jaipur National University, Jaipur. Dr. Datta has completed his graduation i.e., B.A.LL.B. from Law College Dehradun, Hemvati Nandan Bahuguna Garhwal University, Srinagar, Uttarakhand. He is an alumnus of KIIT University, Bhubaneswar where he pursued his post-graduation (LL.M.) in Criminal Law and subsequently completed his Ph.D. in Police Law and Information Technology from the Pacific Academy of Higher Education and Research University, Udaipur in 2020. His area of interest and research is Criminal and Police Law. Dr. Datta has a teaching experience of 7 years in various law schools across North India and has held administrative positions like Academic Coordinator, Centre Superintendent for Examinations, Deputy Controller of Examinations, Member of the Proctorial Board

Dr. Namita Jain

Head & Associate Professor

School of Law, JECRC University, Jaipur Ph.D. (Commercial Law) LL.M., UGC -NET Post Graduation Diploma in Taxation law and Practice, Bachelor of Commerce.

Teaching Experience: 12 years, AWARDS AND RECOGNITION of Dr. Namita Jain are - ICF Global Excellence Award 2020 in the category of educationalist by I Can Foundation, India.India Women Empowerment Award in the category of "Emerging Excellence in Academics by Prime Time &Utkrisht Bharat Foundation, New Delhi.(2020). Conferred in FL Book of Top 21 Record Holders in the category of education by Fashion Lifestyle Magazine, New Delhi. (2020).Certificate of Appreciation for organizing and managing the Professional Development Training Program on IPR in Collaboration with Trade Innovations Services, Jaipur on March 14th, 2019

 ISSN: 2582-6433

Mrs.S.Kalpana

Assistant professor of Law

Mrs.S.Kalpana, presently Assistant professor of Law, VelTech Rangarajan Dr.Sagunthala R & D Institute of Science and Technology, Avadi.Formerly Assistant professor of Law,Vels University in the year 2019 to 2020, Worked as Guest Faculty, Chennai Dr.Ambedkar Law College, Pudupakkam. Published one book. Published 8Articles in various reputed Law Journals. Conducted 1Moot court competition and participated in nearly 80 National and International seminars and webinars conducted on various subjects of Law. Did ML in Criminal Law and Criminal Justice Administration.10 paper presentations in various National and International seminars. Attended more than 10 FDP programs. Ph.D. in Law pursuing.

Avinash Kumar

Avinash Kumar has completed his Ph.D. in International Investment Law from the Dept. of Law & Governance, Central University of South Bihar. His research work is on "International Investment Agreement and State's right to regulate Foreign Investment." He qualified UGC-NET and has been selected for the prestigious ICSSR Doctoral Fellowship.He is an alumnus of the Faculty of Law, University of Delhi. Formerly he has been elected as Students Union President of Law Centre-1, University of Delhi.Moreover, he completed his LL.M. from the University of Delhi (2014-16), dissertation on "Cross-border Merger & Acquisition"; LL.B. from the University of Delhi (2011-14), and B.A. (Hons.) from Maharaja Agrasen College, University of Delhi. He has also obtained P.G. Diploma in IPR from the Indian Society of International Law, New Delhi.He has qualified UGC – NET examination and has been awarded ICSSR – Doctoral Fellowship. He has published six-plus articles and presented 9 plus papers in national and international seminars/conferences. He participated in several workshops on research methodology and teaching and learning.

 ISSN: 2582-6433

ABOUT US

INTERNATIONAL JOURNAL FOR LEGAL RESEARCH & ANLAYSIS ISSN

2582-6433 is an Online Journal is Monthly, Peer Review, Academic Journal, Published online, that seeks to provide an interactive platform for the publication of Short Articles, Long Articles, Book Review, Case Comments, Research Papers, Essay in the field of Law & Multidisciplinary issue. Our aim is to upgrade the level of interaction and discourse about contemporary issues of law. We are eager to become a highly cited academic publication, through quality contributions from students, academics, professionals from the industry, the bar and the bench. INTERNATIONAL JOURNAL FOR LEGAL RESEARCH & ANALYSIS ISSN 2582-6433 welcomes contributions from all legal branches, as long as the work is original, unpublished and is in consonance with the submission guidelines.

 ISSN: 2582-6433

REINVENTING THE RIGHT TO BE FORGOTTEN IN THE AGE OF DATA PERMANENCE AND GENERATIVE ARTIFICIAL INTELLIGENCE

AUTHORED BY - PRAMADVARA KUSHWAHA
Assistant Professor, Galgotias University

ABSTRACT

The advent of generative artificial intelligence (AI) presents unprecedented challenges to the legal and normative contours of the Right to Be Forgotten (RTBF). While this right gained legal visibility in the wake of the landmark *Google Spain v. González* decision in 2014, the evolving capabilities of generative AI, particularly its expansive memory and capacity to reconstruct identifiable personal data from minimal inputs are destabilizing conventional understandings of digital forgetting. This article critically examines the trajectory of the RTBF, tracing its transformation from a subsidiary element of privacy law to an autonomous right grounded in individual informational self-determination.

However, the paper argues that traditional mechanisms such as data anonymization, deletion, and geoblocking are increasingly insufficient to curtail the data retention tendencies of generative AI models. The inherent difficulty in identifying data derivatives, coupled with the technical burden of model retraining, renders effective implementation of RTBF both impractical and legally tenuous. Accordingly, this work calls for a recalibrated legal framework that preserves the core values of the RTBF while accommodating the operational realities and constraints of contemporary AI technologies.

While existing instruments such as the General Data Protection Regulation (GDPR) offer foundational guidance, the article emphasizes the necessity for dynamic, responsive regulatory oversight in conjunction with industry cooperation. This study advances an interdisciplinary analysis drawing from legal theory, technological architecture, and policy studies to propose adaptive strategies that uphold individual dignity and autonomy in the face of rapidly advancing AI systems. The novelty of this contribution lies in its holistic and integrative approach to redefining the RTBF within the architecture of generative AI, offering pragmatic pathways for reconciling digital privacy with technological progress.

 ISSN: 2582-6433

Keywords- Artificial Intelligence, RTBF, Machine Learning, GDPR

INTRODUCTION

Artificial Intelligence (AI), often introduced subtly through cloud-based infrastructures, has now become deeply embedded in everyday life, transforming the very fabric of societal functions. Among the most influential technologies within the broader AI ecosystem is machine learning (ML), whose pervasive influence, whether consciously recognized or not, continues to shape social, economic, and cultural domains. Tony Tether, former director of the United States Défense Advanced Research Projects Agency (DARPA)[1], has aptly described machine learning as the anticipated future of the Internet, underscoring its transformative potential.

While ML technologies offer promising solutions to pressing global challenges, such as climate change mitigation and energy efficiency, through innovations like Google DeepMind, they simultaneously introduce complex ethical and human rights concerns. The same intelligence that drives innovation and optimization may, paradoxically, undermine fundamental values of autonomy, privacy, and dignity.[2] This dual capacity of ML to act both as a technological boon and a source of sociotechnical disruption calls for a critical examination of its integration within the digital and legal landscape, particularly in relation to emerging rights frameworks such as the Right to Be Forgotten.

The rapid advancement of generative artificial intelligence has revived critical discourse surrounding the contemporary relevance of the "Right to Be Forgotten" (RTBF) in the digital age. Originally conceived as a mechanism to empower individuals to erase their personal information from the online domain, thereby safeguarding privacy, this right gained international prominence following the landmark 2014 decision in *Google Spain v. González*[3] by the European Union. However, the emergent capabilities of generative AI, characterized by vast data retention and autonomous content regeneration, pose unprecedented challenges to the enforceability and conceptual integrity of RTBF.

This article traces the evolution of RTBF from a derivative of privacy rights to a distinct right

[1] Ronald Leenes & Silvia De Conca, Artificial Intelligence and Privacy - AI Enters the House Through the Cloud
[2] Cindy Gordon, Google Faced With An AI Privacy Challenge: Do I Have The Right To Be Forgotten?
[3] Google Spain SL v. Agencia Española de Protección de Datos, 2014 E.C.R. 317

 ISSN: 2582-6433

of informational self-determination. Yet, it contends that generative AI's near-indelible memory and capacity for synthetic reconstruction of personal data render traditional tools such as data erasure and anonymization increasingly inadequate. In light of these challenges, the article advocates for a pragmatic recalibration of the RTBF, one that retains its normative essence while recognizing its technological constraints.

It calls for a robust and adaptive legal framework bolstered by regulatory vigilance and proactive industry engagement. The proposed approach underscores the necessity for nuanced strategies that not only protect personal dignity and autonomy but also harness the transformative potential of generative AI in service of the public good. Adopting a multidisciplinary lens, the article integrates legal doctrine, ethical considerations, technological implications, and policy dimensions to reassess the RTBF in the era of generative AI.

JURISPRUDENTIAL EVOLUTION OF THE RIGHT TO BE FORGOTTEN

The origins of the right to be forgotten (RTBF) can be traced to foundational European Union directives, notably the Data Protection Directive and the E-Commerce Directive[4], which collectively laid the groundwork for obligating online service providers, particularly search engines to delist or remove links within the EU when warranted. The jurisprudential turning point came with the landmark Google Spain v. González ruling in 2014[5], where the European Court of Justice (ECJ) interpreted these directives to affirm an individual's entitlement to request the erasure of search results that are outdated, irrelevant, or disproportionate. The case centered around Mr. González, who sought the removal of links to a newspaper article detailing a past real estate auction linked to him, asserting that the information no longer served a legitimate public interest.[6] The ECJ, referencing Article 12(b) of Directive 95/46/EC, held that under certain conditions, search engines must delist such links to safeguard personal data.

The formal recognition of RTBF within European data law was initiated in 2012 with its inclusion in the draft General Data Protection Regulation (GDPR), albeit initially restricted to the protection of minors. However, with the GDPR's adoption in 2018, Article 17 expanded

[4] Andreas von Arnauld, Kerstin von der Decken, & Mart Susi, The Right to Be Forgotten, in THE CAMBRIDGE HANDBOOK OF NEW HUMAN RIGHTS: RECOGNITION, NOVELTY, RHETORIC
[5] Supra note 3
[6] ibid

 ISSN: 2582-6433

the right's applicability to all natural persons, thereby embedding it firmly into the legal framework governing data protection.[7]

The RTBF conceptually builds upon the notion that information, though initially lawful, can lose its relevance or legitimacy over time.[8] Legal interpretations of the right must be situated in both temporal and territorial contexts, factoring in both subjective interests of the data subject and objective considerations. The right comprises two interrelated aspects: the right to forget, which aims to sever individuals from stigmatizing past events to preserve dignity, and the right to erasure,[9] which empowers individuals with control over their digital identities, anchored in principles of informational self-determination, privacy, and data sovereignty.

Over time, RTBF has evolved from a data protection mechanism into a substantive human rights doctrine. Positioned at the intersection of privacy, freedom of expression, and data governance, it plays a critical role in balancing competing rights. The Belgian case Olivier G. v. Le Soir[10] significantly contributed to the judicial shaping of RTBF, further highlighting the legal and ethical tensions between personal data removal and freedom of information. Recent decisions of the European Court of Human Rights have attempted to reconcile these conflicts, emphasizing the need for context-specific, proportional approaches that uphold both human dignity and democratic transparency.

DEVELOPMENT OF RTBF IN INDIA

In the Indian legal system, privacy has long been a contentious right. The discussion has always been on the status which should be given the right to privacy; whether a mere human right or a basic right, protected by the Constitution, as a corollary to the right to life and personal liberty. The earliest instances exploring the right to privacy and granting it a status, not of a basic right, were MP Sharma v. Satish Chandra[11] and Kharak Singh v. State of Uttar Pradesh[12]. Justice

[7] GDPR Commission Regulation 2016/679, 2016 O.J. (L 119)
[8] Charter of Fundamental Rights of the European Union, art. 8, 2012 O.J. (C 326) 397; see International Conference of Data Protection & Privacy Commissioners, International Resolution on Privacy as a Fundamental Human Right and Precondition for Exercising Other Fundamental Rights
[9] Oskar J. Gstrein, Right to Be Forgotten: European Data Imperialism, National Privilege, or Universal Human Right?
[10] Hugh Tomlinson, Case Law, Belgium: Olivier G v Le Soir. "Right to Be Forgotten" Requires Anonymisation of Online Newspaper Archive, INFORRM'S BLOG (2016), https://inform.org/2016/07/19/case-law-belgium-olivier-g-v-lesoir-right-to-be-forgotten-requires-anonymisation-of-online-newspaper-archive-hugh-tomlinson-qc/; Hof van Cassatie [Cass.], AR C150052F, http://www.cass.be (Belg.) available at https://inforrm.files.wordpress.com/2016/07/ph-vog.pdf (last accessed on 9th May, 2025)
[11] M.P. Sharma v. Satish Chandra, 1954 SCR 1077
[12] Kharak Singh v. State of UP, 1964 1 SCR 332

 ISSN: 2582-6433

Subba Rao, in his minority judgment, planted the seeds of its acknowledgment as essential right by saying that rights in Part III of the Constitution have a "overlapping area." The next parts will now cover the Indian courts' view on the 'right to be forgotten' as a byproduct of privacy.

Even prior to the formal recognition of the right to privacy as a fundamental right in India, the judiciary had already begun grappling with the contours of the "right to be forgotten." Indian courts, however, have expressed divergent views on the matter. Notably, a legal inconsistency emerges from the contrasting stances of the Karnataka and Gujarat High Courts, with the Kerala High Court offering more implicit acknowledgment of the right.

In the case of Sri Vasunathan v. The Registrar General & Ors.[13], the bench of High Court, Karnataka, for the very first time in the judicial precedent history, affirms the right to be forgotten. The petitioner, a woman, sought redaction of her name from a judicial order previously issued by the same court, on the grounds that online search results displaying her association with the case could significantly harm her personal and marital life. Her appeal invoked the "right to be left alone," essentially seeking erasure of her digital footprint. Justice Anand Bypareddy, while adjudicating the matter, noted that adopting such a measure would align with international practices, particularly in Western jurisdictions, where the right to be forgotten is upheld, especially in sensitive cases involving women, including instances related to sexual violence or matters impacting dignity and reputation.[14] In this instance, the recognition of the right stemmed from the deeply personal and sensitive nature of the information involved. The petitioner was a party in a matrimonial annulment case and sought to move forward without the continued burden of a publicly accessible digital record. Thus, the court's reasoning subtly incorporated global jurisprudence to accommodate evolving demands for informational autonomy in the Indian context.

The Delhi High Court has also been confronted with issues implicating the contours of the right to be forgotten. In the ongoing case of Zulfiqar Ahman Khan v. M/s Quintillion Business Media Pvt. Ltd. & Ors.[15], the plaintiff, a prominent public figure, sought the removal of certain online articles published during the #MeToo movement. These publications, based on anonymous

[13] Writ Petition Number 62038 of 2016
[14] Ibid
[15] Zulfiqar Ahman Khan v. M/s Quintillion Business Media Pvt. Ltd. & Ors., 2019 (175) DRJ 660

 ISSN: 2582-6433

allegations, were alleged to have caused significant harm to the plaintiff's professional standing and public image. In response, the Court issued an interim restraining order, recognizing that the constitutional right to privacy encompasses both the "right to be forgotten" and the "right to be left alone."[16] The Court thereby prohibited any further dissemination of the contested material until the case is resolved. Similar petitions have been brought before various other High Courts, seeking the removal or masking of judicial pronouncements that continue to appear in search engine results under the names of the individuals involved. Such requests underscore a growing concern over the persistence of digital records and their implications for personal dignity.

These judicial developments collectively reveal a lack of consistency and clarity in the Indian legal landscape regarding the full and formal recognition of the right to be forgotten. The jurisprudence remains fragmented, with High Courts adopting varying interpretations and degrees of enforcement, thereby highlighting the need for a uniform statutory or constitutional framework to address these emerging challenges in the digital age.

THE PRIVACY JUDGMENT: S.N. KAUL'S OBSERVATIONS

It is now increasingly evident that the "right to be forgotten" forms an integral component of an individual's right to privacy, warranting protection in both physical and digital domains. While the Indian judiciary has a well-documented history of grappling with privacy as a constitutional concern, the definitive elevation of privacy to the status of a fundamental right came with the Supreme Court's landmark judgment in Justice K.S. Puttaswamy (Retd.) v. Union of India[17]. This ruling not only recognized privacy as a constitutionally protected right but also aligned Indian jurisprudence with global human rights standards.

Of particular significance to this study is the concurring opinion of Justice Sanjay Kishan Kaul, who offered a nuanced analysis of privacy in the context of emerging technologies. Emphasizing the need to protect personal data from unwarranted access, especially by non-state actors, Justice Kaul highlighted the extent to which digital platforms like Facebook, Uber, and Alibaba collect and process user information. He drew attention to the challenges posed by big data, calling for robust privacy safeguards in the digital ecosystem and advocating for

[16] Kunal Garg, Right to be forgotten in India: A Hustle over Protecting Personal Data, INDIA LAW JOURNAL https://indialawjournal.org/a-hustle-overprotecting-personal-data.php (Last ACCESSED on 9th May, 2025),
[17] K. S. Puttaswamy v. Union of India, [2012] Writ Petition (Civil) No. 494

 ISSN: 2582-6433

individual autonomy over personal information.

Justice Kaul further invoked European legal principles to propose a framework akin to the "right to be forgotten," suggesting that individuals should be able to request the erasure of their data when it is no longer relevant, necessary, or accurate.[18] His approach reflected the French notion of the "right to oblivion," rooted in the human capacity for error, reform, and the desire to start anew.

This judicial recognition of the right to be forgotten is of considerable jurisprudential value, particularly in the absence of explicit legislative provisions. The ruling fills a critical gap by reinforcing digital privacy protections and affirms the judiciary's proactive role in safeguarding constitutional rights in the evolving landscape of cyberspace.

LET US UNDERSTAND MACHINE LEARNING

Machine Learning (ML) technologies are increasingly deployed in critical domains, including facial recognition in visual media, personalized product recommendations, and criminal identification systems. Often referred to as "Software 2.0,"[19] ML programs do not rely on manually coded instructions. Instead, they learn autonomously by analyzing vast amounts of data. This data, collected from individuals, have delicate information of a person such as email addresses, financial details & employment records. Privacy regulations in various jurisdictions[20] grant individuals the Right to Be Forgotten (RTBF). This legal provision allows users to request the deletion of their personal data and associated records from service providers. In response to such requests, ML service providers may need to remove relevant data from their training datasets and retrain their models. Cases involving major entities like Clearview AI[21], Google[22], and Europol[23] exemplify the practical implications of these rights. With the growing prominence of data protection norms and increased public awareness, such demands are likely to rise. This requirement to erase personal data from trained models introduces a critical technical challenge. Researchers have responded by developing "machine

[18] Ibid
[19] Ratner, A.J., Hancock, B., Ré, C.: The role of massively multi-task and weak supervision in software 2.0
[20] EU GDPR, California Consumer Privacy Act, Canada's PIPEDA
[21] https://www.buzzfeednews.com/article/richardnieva/clearview-ordered-to-delete-in-france (last accessed on 9th May, 2025)
[22] https://www.reuters.com/article/us-eu-alphabet-privacy-idUSKBN1W90R5 (last accessed on 9th May, 2025)
[23] https://www.bleepingcomputer.com/news/security/europol-ordered-to-erase-data-on-those-not-linked-to-crime/ (last accessed on 9th May, 2025)

 ISSN: 2582-6433

unlearning" techniques that aim to ensure ML models can effectively forget the influence of specific data points used during training. The goal of machine unlearning is to eliminate the learned representation of deleted data without the costly and computationally intensive process of retraining the entire model from scratch—a naïve yet obvious solution that is rarely feasible in practice. Recent studies have explored more efficient unlearning strategies that circumvent these limitations.

However, much of the current focus in the field of machine unlearning has centered around computational efficiency and compliance with RTBF requirements, often neglecting other important ethical and legal dimensions—such as algorithmic fairness. Fairness in AI pertains to minimizing bias in ML models, particularly biases related to protected characteristics like race, gender, and familial status. Although fairness has been a central concern in AI ethics research, there remains a significant gap in understanding how machine unlearning affects these dimensions.

Notably, machine unlearning methods differ from conventional ML in their approach to data input and training processes, which may influence fairness outcomes. To date, there has been limited examination of how unlearning techniques impact algorithmic bias, raising concerns that such omissions could inadvertently reinforce discrimination. Consequently, unexamined use of machine unlearning might contravene anti-discrimination frameworks such as the U.S. Civil Rights Act.

WHY DO WE FORGET THINGS? DIFFERENCE OF HUMAN MEMORY AND MACHINE MEMORY

To critically examine the application of the Right to be Forgotten (RTBF) in the realm of artificial intelligence, it is essential to first explore how memory and forgetting are conceptualized in both human cognition and AI systems. Present legal frameworks often conflate human and machine memory, treating them as functionally equivalent. This approach is based on a misconception, as it fails to align with the realities of either domain. Scholars have already pointed out the problematic assumption that machines possess perfect recall, a quality that diverges significantly from human memory processes.[24]

[24] https://www.bbc.com/future/article/20150401-whats-the-most-we-can-remember (last accessed on 9th May, 2025)

 ISSN: 2582-6433

In cognitive psychology, we can find two parts of memory which exists in human mind: short-term and long-term. However, researchers have yet to reach a definitive consensus regarding the key distinctions between the two. The transfer of information to long-term memory may depend on various factors, including the perceived relevance or "meaningfulness" of an experience, though what constitutes "meaningfulness" remains ambiguous. Moreover, even basic estimates of the storage capacity of the human brain remain uncertain. As such, our scientific grasp of human memory remains limited, and in some respects, deeply flawed.

Conversely, the mechanisms underpinning artificial intelligence are generally better understood. This is primarily because AI systems are built upon logical frameworks and computational rules designed by humans. While advanced AI models may develop their own internal procedures without explicit instruction, a phenomenon often described as the "black box" problem, the foundational principles governing AI behaviour are still largely within the realm of human comprehension. This contrasts with the relative opacity surrounding the human brain's decision-making functions.

Importantly, AI systems offer a clearer model of how data is handled, from input to storage to deletion. While the inner workings of complex AI algorithms may not always be fully transparent, the general operations of data processing in artificial systems are better mapped than those of human cognition. Recognizing this fundamental difference is crucial when evaluating the adequacy of current privacy laws. Specifically, it highlights critical gaps in how laws such as the RTBF are structured, given their tendency to impose human-like forgetting standards on artificial entities that function very differently.

CHALLENGES TO FOLLOW GDPR COMPLIANCE IN THE ML ENVIRONMENT

Machine Learning (ML) has demonstrated its capability to handle complex tasks across a wide range of sectors, notably in healthcare and transportation, and continues to show promise for even broader adoption. Nevertheless, concerns related to the General Data Protection Regulation (GDPR) can arise, particularly when personal data is gathered and processed by ML systems without the explicit awareness or consent of the individuals involved.[25] As part of

[25] THE ROYAL SOC'Y, MACHINE LEARNING: THE POWER AND PROMISE OF COMPUTERS THAT LEARN BY EXAMPLE 34 (2017), https://royalsociety.org/~/media/policy/projects/machine-learning/publications/machinelearning-report.pdf (last accessed on 9th May, 2025)

 ISSN: 2582-6433

a deeper examination into GDPR compliance within ML environments, the following discussion will explore three specific ML subfields that are especially prone to triggering regulatory and data protection challenges.[26]

I. ALGORITHM

Various branches of machine learning (ML), such as supervised learning, unsupervised learning, reinforcement learning, and deep learning are extensively integrated into algorithmic systems. These technologies are increasingly utilized in commercial settings to interpret and manage massive datasets that are characterized by high volume, speed, and diversity, and to execute distinct, task-specific functions. While the operational mechanisms of these ML models are technically akin to software executing a sequence of computational instructions, they tend to yield comparable outcomes when fed with similarly structured input data. This is primarily because algorithms are designed to detect and leverage unique patterns within an individual's dataset, enabling them to infer connections with other individuals exhibiting similar traits.[27]

Once trained using designated "training data," such algorithms acquire the capacity to independently process new datasets without further human guidance, even if the new data has not been explicitly labelled. A case in point is facial recognition technology: once an algorithm is fed images tagged with a specific individual's identity, it becomes capable of recognizing that person in other photographs or visual inputs automatically.[28]

However, these algorithms are not infallible. In unsupervised learning, for instance, the absence of labelled data means the algorithm autonomously determines how to categorize and interpret information, making it difficult for developers to trace or validate the accuracy of the system's outcomes.[29] This introduces concerns around classification errors or anomalies in clustering. Unlike supervised learning, where the outcome is more predictable due to labelled inputs, unsupervised models operate in a more opaque manner, limiting the programmer's ability to interpret decision-making processes.[30]

[26] Stephen McJohn & Ian McJohn, Fair Use and Machine Learning
[27] Warren E. Agin, A Simple Guide to Machine Learning
[28] Harry Surden, Machine Learning and Law
[29] Patrick W. Nutter, Comment, Machine Learning Evidence: Admissibility and Weight
[30] Argyro P. Karanasiou & Dimitris A. Pinotsis, A Study into the Layers of Automated Decision-Making: Emergent Normative and Legal Aspects of Deep Learning

 ISSN: 2582-6433

Reinforcement learning models, on the other hand, aim to optimize outcomes over extended periods, often tolerating minor immediate inaccuracies in pursuit of a more robust and stable analytical model. Despite these challenges and the inherent unpredictability associated with ML technologies, their application continues to be justified and increasingly adopted across domains.

II. DATA COLLECTION

Algorithms function as structured instructions for sorting and interpreting datasets, necessitating vast quantities of data to operate at their highest potential. The United Kingdom's Information Commissioner's Office (UK ICO) has emphasized that the foundation of machine learning (ML) systems lies in the availability of substantial and diverse datasets.[31] Given that "big data" is characterized by its variety, velocity, and volume, it requires sophisticated and intelligent data processing mechanisms capable of extracting value from the entirety of the information collected.

Big data storage systems are generally categorized into operational databases (front-end) and long-term repositories such as archives or backups. However, in practice, these systems often prioritize specific, relevant data subsets for processing and preservation rather than handling the entire dataset. This selective approach, known as data summarization, is commonly seen in commercial contexts, where algorithms interpret behavioural patterns to customize advertisements or adjust pricing models based on customers' social or economic profiles.

Machine learning-driven big data analysis yields numerous advantages, extending beyond individual benefits to collective gains for communities and governance structures. The ability to harness big data can stimulate technological advancement, enhance societal communication, improve economic productivity, and support more responsive public administration. These benefits often serve as counterbalances when assessing concerns around privacy violations, such as the illegitimate gathering, handling, or disclosure of personal information, or discriminatory outcomes driven by automated decision-making systems. Furthermore, the autonomous nature of unsupervised learning, which operates independently of human guidance makes it difficult to determine the exact data it will process or predict the nature of the outputs.

[31] INFO. COMM'RS OFF., BIG DATA, ARTIFICIAL INTELLIGENCE, MACHINE LEARNING AND DATA PROTECTION 7 (2017), https://ico.org.uk/media/fororganisations/documents/2013559/big-data-ai-ml-and-data-protection.pdf (last accessed on 10th May, 2025)

 ISSN: 2582-6433

These outputs may themselves be used in future algorithmic workflows, further complicating accountability. Reinforcement learning, though distinct in using a trial-and-error learning mechanism, also processes data without human oversight. These autonomous models typically require broader access to external data sources, sometimes even involving third parties under the doctrine of data portability.[32]

Additionally, retaining some degree of data is often indispensable for ensuring the smooth functioning of ML systems or resolving technical issues that may arise during their operation.

A TECHNICAL ANALYSIS OF AI DATA DELETION ("FORGETTING")

The **Right to be Forgotten (RTBF)** is conceptually rooted in the metaphor of human forgetting—an individual's request to have their personal data deleted can be likened to asking others to forget specific information. However, this metaphor does not seamlessly extend to artificial intelligence and machine learning systems. Unlike human cognition, AI does not "forget" in a natural or organic manner. Instead, data deletion within AI-driven architectures presents substantial technical challenges.

One core difficulty lies in determining whether true deletion is even feasible in modern, data-intensive environments. In particular, **relational database management systems (DBMSs)**, which underpin many AI applications, complicate the implementation of RTBF. These systems prioritize data retrieval efficiency through advanced indexing structures—typically **B+ Trees**—that organize and store records for rapid search and access. Additional indexing layers are often created to further optimize specific query types. While such structures enable swift data access across billions of records, they also disperse and replicate data in ways that obscure complete erasure.

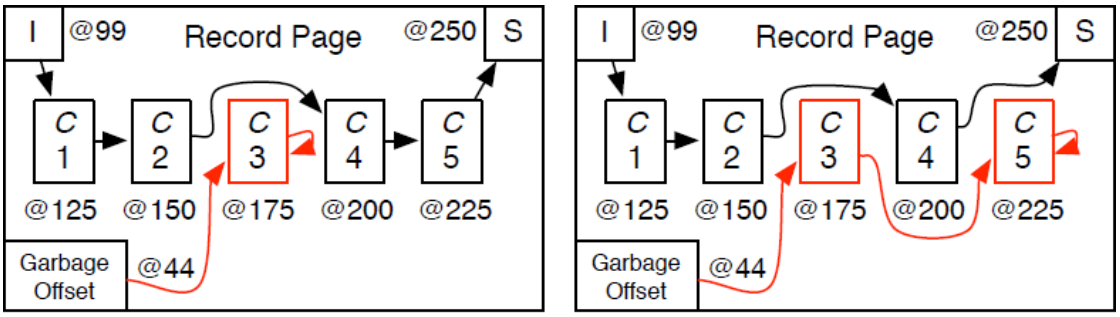
Moreover, user interactions with these systems occur through structured query languages like **SQL**, which abstract the complexities of data storage and retrieval. Consequently, ensuring the permanent and verifiable deletion of personal data—across all instances and derivations—becomes an intricate task, especially given the existence of backups, logs, and derived data.

[32] Jules Polonetsky & Omer Tene, Privacy and Big Data: Making Ends Meet, 66 STAN. L. REV. ONLINE 25, 28–29 (2013), https://review.law.stanford.edu/wpcontent/uploads/sites/3/2016/08/PolonetskyTene.pdf (last accessed on 10th May, 2025)

 ISSN: 2582-6433

These technical realities illustrate the **disjunction between legal mandates for data erasure and the operational realities of AI systems**, underscoring the need for rethinking data governance in machine learning contexts.

LIMITATIONS OF DATA DELETION IN RELATIONAL DATABASES



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

Figure 1(a) represents a simplified view of a relational database segment, focusing on a specific page rather than the full index tree. Within this page, five data records (C1–C5) are stored between nodes I (start) and S (end). One record, C3, has already been marked for deletion and added to the "garbage offset"—a list of reclaimable storage locations.

In the event of a deletion request, such as for record C5, the database traverses the search tree to locate and isolate the target data. Upon identification, the entry is not physically erased; instead, it is flagged for deletion. This involves reconfiguring pointers to exclude C5 from the active search path and linking it to the garbage offset, essentially transferring it from active to inactive status.

Importantly, the record remains on disk, and its removal from the index merely renders it invisible to regular queries. Actual data erasure only occurs if and when the system reuses that storage space—a process that may be significantly delayed, as many databases prefer appending new data over searching for reusable space due to performance optimization strategies.

This technical mechanism demonstrates that so-called "deletion" in relational databases often equates to de-indexing rather than true erasure, posing serious implications for the enforcement of the **Right to be Forgotten** in AI systems.

 ISSN: 2582-6433

CONCLUSION

While the Right to be Forgotten constitutes only a narrow segment of broader privacy legislation, its examination through a technical lens highlights the pressing need for deeper interdisciplinary collaboration between legal and technological domains. Although the right is rooted in well-meaning regulatory intent and is widely regarded as a critical safeguard for individual autonomy, its practical enforcement within AI and data-driven environments reveals significant friction between normative legal expectations and computational realities.

As privacy scholars have similarly observed in the context of Privacy by Design, the legal conceptions of data deletion often fail to translate into system-level operations due to divergent terminologies and conceptual frameworks. This disconnects fosters miscommunication and hampers implementation, particularly in complex data infrastructures where system integrity relies on the ability to revert to previous states—an essential feature for maintaining ACID compliance and operational resilience. Vint Cerf aptly noted the impracticality of universal content removal, underscoring the technical challenges of enforcing "forgetting" in a distributed and persistent digital ecosystem. In practice, AI systems and databases are not inherently equipped to "forget" in the human sense, making the fulfillment of legal mandates like the Right to be Forgotten technically elusive.

This inquiry into the doctrinal and technical contours of digital forgetting exemplifies the value of interdisciplinary engagement. However, a more expansive approach—including insights from cognitive science and neuroscience—is imperative to fully grasp and operationalize the nuanced interplay between human memory, artificial intelligence, and legal frameworks.