Memory Undone: Between Knowing and Not Knowing in Data Systems

Viktoriia Makovska
Ukrainian Catholic University
makovska.pn@ucu.edu.ua

George Fletcher
Eindhoven University of Technology
g.h.l.fletcher@tue.nl
gfletche.win.tue.nl

Julia Stoyanovich
New York University
stoyanovich@nyu.edu
stoyanovich.org

Tetiana Zakharchenko
Ukrainian Catholic University
t.herasymova@ucu.edu.ua

Memory in machine learning and data systems. Machine learning systems increasingly operate as infrastructures of memory. They collect, store, and learn from various datasets, many of which contain traces of personal, political, or cultural life. This raises a question: while data systems are designed for remembering, what role does a capability to forget play? The European Union's General Data Protection Regulation (GDPR) introduced the "right to be forgotten," granting individuals the legal ability to request deletion of their personal data.[1] Permanently removing information from large-scale systems is technically challenging, since deletion must propagate through copies, caches, and derived data products, and most infrastructures were originally designed to accumulate rather than erase data [1]. Moreover, deleting records from database storage does not remove their influence on models trained on that data. In modern deep learning systems, once information is integrated into the weights of a neural network, its traces remain entangled in billions of parameters. This mismatch between legal frameworks and technical capabilities has made machine unlearning an urgent research problem. But the challenge goes beyond deletion. We need mechanisms to evaluate whether data traces persist or have been effectively unlearned inside these systems. As an analogy, in quantum computation the quantum no-deleting theorem shows that an unknown quantum state cannot be deleted against a copy [17]. Likewise, in machine learning, "deletion" is often not a complete removal, but of redistributing or attenuating influence across representations. Even after erasure or unlearning interventions, traces of prior data may persist indirectly, through learned correlations, semantic dependencies, or downstream artifacts.

To reduce ambiguity, we distinguish several related practices. *Erasure* is an operational act that removes or disables access to data artifacts (e.g., rows, files, indexes, caches) [16]. *Unlearning* is an intentional technical intervention aimed at removing (or bounding) the influence of specific data on learned parameters and downstream outputs [7]. *Exclusion* refers to upstream decisions not to collect, label, or include certain people, events, or categories in datasets; it is a form of "forgetting" by omission [6, 13]. *Forgetting*, in contrast, is an umbrella term for sociotechnical processes through which something becomes unavailable, varying in agency (intentional vs. incidental), temporality (momentary vs. durable), reversibility (recoverable vs. irreversible), and scale (individual records vs. population-level) [10]. On this view, both GDPR-backed deletion and the systematic non-inclusion training data are forms of "forgetting," yet they operate through different mechanisms and raise distinct questions of governance.

Technical advances in unlearning demonstrate partial progress: exact unlearning in recommender systems [28], class forgetting without gradients [15], erasure in semantic web and description logic ontologies [9, 11] and selective unlearning in large language models [31]. Still, fundamental questions remain unresolved. What does it mean for a system to truly forget? How can erasure be ensured in the presence of semantic dependencies, where one record influences others? And at what point does forgetting protect rights, and when may it instead threaten them? For example, forgetting biased or harmful data may reduce harm, but forgetting underrepresented perspectives risks reproducing epistemic injustice [26, 27].

---

[1] https://eur-lex.europa.eu/eli/reg/2016/679/oj

Forgetting is not an anomaly in data science but a consistent practice. At every stage of the pipeline, from data collection to feature design, to labeling, choices are made about what to exclude, ignore, or treat as noise. These acts of exclusion represent forms of forgetting. Prior research shows that data never exist in a neutral form: they become 'data' only when transformed by human decisions shaped by social, political, and economic priorities [5, 30]. This opens the possibility of data silences [13], or even systematic erasures where certain groups or perspectives are excluded altogether [2, 8]. At the infrastructure level, preservation is often treated as the default ("keep everything"), while deletion and unlearning appear as exceptional interventions motivated by rights (privacy) or necessity (harm reduction).

The balancing of knowing and not knowing. Work in participatory data modeling further underscores the critical balances between knowing and not knowing. Studies of entity–relationship design have revealed the presence of invisible entities—objects, actors, and relations that shape data models without being explicitly represented [22]. Designers often leave such elements undescribed to reduce complexity or maintain convenience. Their absence is not incidental but engineered, and it mirrors broader forgetting practices in data science. What is excluded can structure outcomes as powerfully as what is included. Similarly, data infrastructures contain nonrandom exclusions: the absence of data about certain populations comes from systematic marginalization or perceived lack of economic or political value [18]. Forgetting, in this sense, is never neutral but an act of governance.

The stakes become clearer when considering information manipulation at scale. Research on Russian disinformation networks demonstrates how propaganda content, once statistically dominant in training corpora, can distort model outputs [21]. Large language models reproduce such narratives, sometimes without disclaimers, illustrating a form of epistemic poisoning. Here, forgetting as a tool is not only about privacy but about resisting the integration of manipulative content into the "memory" of AI systems. Whether unlearning can provide remedies for such epistemic distortions remains an open question. What is clear, however, is that forgetting must be understood as both a technical safeguard and a political practice.

Modern epistemic infrastructures tend to centralize knowledge, reinforcing popular or canonical ideas and narrow the range of what is considered to be stored and what is erased. Epistemic decentralization, by contrast, enables participation in knowledge production from historically excluded actors. It can bring up unseen or shadowed perspectives, challenge existing assumptions, and correct what has been conventionally accepted as objective knowledge [23]. Francis Bacon warned through the Idols of the Theater, dominant intellectual frameworks can hide alternative ways of knowing and limit our possibilities to be open to unknown knowledge [3].

Towards memory as a sociotechnical practice. In our work-in-progress we are investigating memory from a sociotechnical perspective, towards more socially responsive and just data and machine learning systems. Forgetting can be beneficial: it enables focus, reduces complexity, and removes harmful influence [30]. Yet it can also marginalize, silence, or overwrite. Researchers of memory remind us that remembering and forgetting are intersected, to understand one requires acknowledging the other [2, 4, 12]. This framing is aligned with work in social sciences and memory studies that treats remembering/forgetting as institutionally organized practices, not merely individual cognitive events [24]. It also resonates with research in cognitive psychology showing that human memory is reconstructive and susceptible to post-event framing effects [20], and with evidence that post-decision attitude change can occur even without explicit memory for the decision (a phenomenon sometimes discussed as "synthetic" preference or satisfaction) [19]. In data systems, the database is often imagined as a fortress [25], designed to protect and preserve memory. But memory is naturally selective, it becomes effective when it retains certain information while skipping others. Forgetting is not an exception to knowledge, but its condition. To know is always to begin from a position of not knowing [8, 14]. Knowing itself is never neutral: it assumes the power to select and keep authority over what is known, establishing a hierarchical relation between knower and known. In this context, not knowing can operate as a productive epistemic tool instead of being considered as a failure to know [12].

This perspective reframes machine unlearning not simply as a compliance tool, but as an opening to reimagine what knowing means in computational systems. If large language models are machines of remembering (aggregating vast corpora into probabilistic patterns) then forgetting must become a first-class capability. At the same time, critical media theory argues that technology functions primarily as a retentional device, and is therefore structurally misaligned with "forgetting" [29]; we take this as a challenge to clarify what kinds of forgetting are possible (and which are not) in computational infrastructures. This suggests developing models that are capable of unlearning by design, where forgetting is intentional, transparent, and accountable. Rather than treating forgetting as a limitation, unlearning can become a principle for healthier knowledge infrastructures.

Machine unlearning illustrates the fragile boundary: erasure can protect individuals, but it can also function as a tool of silencing. It highlights the dual role of forgetting and unlearning as both safeguard and risk. By reframing unlearning as a foundational step in the machine learning pipeline, this work aims to contribute to thinking about knowledge infrastructures: forgetting must be considered not as failure, but as a necessary, designed, and acknowledged practice.

Goals of the talk.   We propose to view forgetting in AI as a sociotechnical practice rather than a purely technical feature. This involves analyzing how unlearning operates within structures of power, rights, epistemic justice, and governance, and how it differs from forgetting and not knowing. Crucially, we treat forgetting/unlearning not as an internal optimization goal, but as a sociopolitical intervention with consequences beyond the data system—for example for privacy and redress, for the amplification or suppression of narratives, and for the potential silencing of underrepresented perspectives.

We also point toward the development of forgetting machines — computational systems capable not only of remembering but of forgetting responsibly by design. By bringing together advances in machine unlearning with critical insights from philosophy, ethics, memory studies, and human–computer interaction, we argue that forgetting should be repositioned as an intentional and accountable practice in the design of future data systems.

Our aims in this presentation are as follows: (1) to present the reframing of forgetting in data systems as a sociotechnical practice; (2) to illustrate how acts of unlearning intersect with power, rights, epistemic justice, and governance; and (3) to open a broader discussion about how computational and epistemic infrastructures might be designed to forget responsibly by design. We aim to engage the conference participants in critical dialogue, sharing ideas and perspectives and building new connections and networks of scholars in Computer Science and beyond.

Biographies.

Viktoriia Makovska. PhD researcher at the Ukrainian Catholic University (UCU), working on responsible data practices and machine unlearning. She holds a master's in data science from UCU and previously spent eight years in industry as a software engineer. She teaches decision support systems at the American University Kyiv and natural language processing at UCU.

George Fletcher. George studies data management systems with a particular focus on the social and human dimensions of query and schema languages. He has been on the faculty at Eindhoven University of Technology for the past 15 years and serves on the boards of the EDBT Association, ACM Transactions on Database Systems, ACM SIGMOD Record, and Transactions on Graph Data and Knowledge. He is a regular program committee member for EDBT, VLDB, ICDE, and SIGMOD, and has recently co-organized interdisciplinary Shonan and NIAS Lorentz Center seminars that bring together computer scientists, humanities scholars, and social scientists to explore algorithms and data as sociotechnical interventions. He holds a PhD in Computer Science from Indiana University Bloomington and completed undergraduate studies in mathematics and cognitive science at the University of North Florida.

Julia Stoyanovich. Dr. Julia Stoyanovich is the Institute Associate Professor of Computer Science and Engineering, Associate Professor of Data Science, and Director of the Center for Responsible AI at New York University. Her mission is to make "Responsible AI" synonymous with "AI." She pursues this goal through academic research, education, technology policy, and public engagement, regularly speaking about both the benefits and the risks of AI. Her research spans data management and AI systems, as well as the ethics and governance of AI. Julia holds an M.S. and Ph.D. in Computer Science from Columbia University, and a B.S. in Computer Science and in Mathematics & Statistics from the University of Massachusetts Amherst. She is a recipient of the Presidential Early Career Award for Scientists and Engineers (PECASE) and is a Senior Member of the Association for Computing Machinery (ACM).

Tetiana Zakharchenko. Tetiana is the Director of the Computer Science Program and the Center for Data, AI, and Society at the Ukrainian Catholic University. Her work lies at the intersection of technology and society, with a focus on responsible AI, algorithmic fairness, and data science in education. Tetiana holds a Ph.D. in Mathematics from Taras Shevchenko National University of Kyiv and a MicroMasters in Statistics and Data Science from MIT. She has been a Research Fellow at the Center for Responsible AI at New York University. Tetiana is a member of the ELLIS Community and serves on the Committee on Artificial Intelligence Development of the Ministry of Digital Transformation of Ukraine.

References

[1] Serge Abiteboul and Julia Stoyanovich. "Transparency, Fairness, Data Protection, Neutrality: Data Management Challenges in the Face of New Regulation". In: *Journal of Data and Information Quality* 11 (2019), pp. 1–9. DOI: 10.1145/3310231.

[2] Ariella Aïsha Azoulay. *Potential history: Unlearning imperialism*. Verso, 2019.

[3] Francis Bacon. *The New Organon*. Ed. by Lisa Jardine and Michael Silverthorne. Cambridge: Cambridge University Press, 2000.

[4] Geoffrey C Bowker. *Memory practices in the sciences*. MIT Press, 2008.

[5] Geoffrey C. Bowker and Susan Leigh Star. *Sorting things out*. MIT Press, 1999.

[6] D'Ignazio C. and Klein L. F. "Data Feminism". In: *The MIT Press* (2020). DOI: https://doi.org/10.7551/mitpress/12255.001.0001.

[7] Yinzhi Cao and Junfeng Yang. "Towards Making Systems Forget with Machine Unlearning". In: *2015 IEEE Symposium on Security and Privacy*. 2015, pp. 463–480. DOI: 10.1109/SP.2015.35.

[8] Michelle Caswell. *Urgent archives: Enacting liberatory memory work*. Routledge, 2021.

[9] Rada Chirkova and George H. L. Fletcher. "Towards Well-Behaved Schema Evolution". In: *12th International Workshop on the Web and Databases, WebDB 2009, Providence, Rhode Island, USA, June 28, 2009*. 2009. URL: http://webdb09.cse.buffalo.edu/papers/Paper37/ChirkovaFletcher-webDB09-camera.pdf.

[10] Paul Connerton. "Seven types of forgetting". In: *Memory Studies* 1.1 (2008), pp. 59–71. DOI: 10.1177/1750698007083889. eprint: https://doi.org/10.1177/1750698007083889. URL: https://doi.org/10.1177/1750698007083889.

[11] Giuseppe De Giacomo et al. "On Instance-level Update and Erasure in Description Logic Ontologies". In: *Journal of Logic and Computation* 19.5 (Aug. 2008), pp. 745–770. ISSN: 0955-792X. DOI: 10.1093/logcom/exn051. eprint: https://academic.oup.com/logcom/article-pdf/19/5/745/3460252/exn051.pdf. URL: https://doi.org/10.1093/logcom/exn051.

[12] Marisol de la Cadena. "Not Knowing: In the Presence of ..." In: *Experimenting with Ethnography: A Companion to Analysis*. Duke University Press, 2021.

[13] Mimi Ọnụọha. *On missing data sets*. https://github.com/MimiOnuoha/missing-datasets. 2016.

[14] Melanie Feinberg. *Everday adventures with unruly data*. MIT Press, 2022.

[15] Sangamesh Kodge, Gobinda Saha, and Kaushik Roy. "Deep Unlearning: Fast and Efficient Gradient-free Class Forgetting". In: *Transactions on Machine Learning Research* (2024). ISSN: 2835-8856. URL: https://openreview.net/forum?id=BmI5p6wBi0.

[16] Bert-Jaap Koops. "Forgetting Footprints, Shunning Shadows. A Critical Analysis Of The "Right To Be Forgotten" In Big Data Practice". In: *SCRIPTed* (2011). DOI: 10.2966/scrip.080311.229. URL: http://script-ed.org/?p=43.

[17] Arun Kumar Pati and Samuel L. Braunstein. "Impossibility of Deleting an Unknown Quantum State". In: *Nature* 404.6774 (2000), pp. 164–165. DOI: 10.1038/404130b0.

[18] Jonas Lerman. "Big Data and Its Exclusions". In: *Stanford Law Review Online* 66 (2013). Published September 3, 2013, pp. 55–63. URL: https://review.law.stanford.edu/wp-content/uploads/sites/3/2016/08/66_stanlrevonline_55_lerman.pdf.

[19] Matthew D. Lieberman et al. "Do Amnesics Exhibit Cognitive Dissonance Reduction? The Role of Explicit Memory and Attention in Attitude Change". In: *Psychological Science* 12.2 (2001), pp. 135–140. DOI: 10.1111/1467-9280.00323.

[20] Elizabeth F. Loftus and John C. Palmer. "Reconstruction of Automobile Destruction: An Example of the Interaction Between Language and Memory". In: *Journal of Verbal Learning and Verbal Behavior* 13.5 (1974), pp. 585–589. DOI: 10.1016/S0022-5371(74)80011-3.

[21] Mykola Makhortykh et al. "Stochastic lies: How LLM-powered chatbots deal with Russian disinformation about the war in Ukraine". In: *Harvard Kennedy School Misinformation Review* 5 (4 2024). DOI: 10.37016/mr-2020-154.

[22]   Viktoriia Makovska, George Fletcher, and Julia Stoyanovich. "ONION: A Multi-Layered Framework for Participatory ER Design". In: *Proceedings of the Workshop on Human-In-the-Loop Data Analytics*. HILDA '25. Intercontinental Berlin, Berlin, Germany: Association for Computing Machinery, 2025. ISBN: 9798400719592. DOI: 10.1145/3736733.3736736. URL: https://doi.org/10.1145/3736733.3736736.

[23]   Leandro Rodriguez Medina and Sandra G. Harding, eds. *Decentralizing Knowledges: Essays on Distributed Agency*. Durham: Duke University Press, 2025, pp. 1–30.

[24]   *Memory Studies Review, Issue 2*. Brill: Memory Studies Review. URL: https://brill.com/view/journals/mesr/1/2/article-p195_001.xml?language=en.

[25]   Richard Monson-Haefel. "Database As a Fortress". In: *97 Things Every Software Architect Should Know*. O'Reilly Media, 2009.

[26]   Michael Muller. "Data silences: How to unsilence the uncertainties in data science". In: *Proceedings of the 2024 SIAM International Conference on Data Mining (SDM)*. SIAM. 2024, pp. 388–391.

[27]   Michael Muller and Angelika Strohmayer. "Forgetting Practices in the Data Sciences". In: *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 2022.

[28]   Sebastian Schelter, Mozhdeh Ariannezhad, and Maarten de Rijke. "Forget Me Now: Fast and Exact Unlearning in Neighborhood-based Recommendation". In: *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval*. SIGIR '23. Taipei, Taiwan: Association for Computing Machinery, 2023, 2011–2015. ISBN: 9781450394086. DOI: 10.1145/3539618.3591989. URL: https://doi.org/10.1145/3539618.3591989.

[29]   Bernard Stiegler. "Memory". In: *Critical Terms for Media Studies*. Ed. by W. J. T. Mitchell and Mark B. N. Hansen. University of Chicago Press, 2010.

[30]   Angelika Strohmayer and Michael Muller. "Data-ing and Un-Data-ing". In: *Interactions* 30.3 (May 2023), 38–42. ISSN: 1072-5520. DOI: 10.1145/3587240. URL: https://doi.org/10.1145/3587240.

[31]   Lingzhi Wang et al. "Selective Forgetting: Advancing Machine Unlearning Techniques and Evaluation in Language Models". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. 2025.