BODHIVRUKSHA JOURNAL OF DIVERSE DISCIPLINE
E-ISSN : 3139-1486
An Open Access, Peer-Reviewed, Multidisciplinary Scholarly Journal
Bi-monthly | Multilingual (English, Hindi, Marathi) | Academic Research
@ Volume 2, Issue 5 (September-October 2026) (Special Issue 1) ISSN: 3139-1486 (Online)

# Consent-Aware Machine Unlearning: A Verifiable Framework for Selective Influence Removal

**Aayush Ramesh Ingle**

Dept. of Artificial Intelligence and Data Science
Ajeenkya DY Patil School of Engineering
Pune, Maharashtra, India
aayushingle23@gmail.com

**Dr. Vishal Puranik**

Dept. of Artificial Intelligence and Data Science
Ajeenkya DY Patil School of Engineering
Pune, Maharashtra, India
vishalpura@gmail.com

**Abstract**

Machine unlearning is the question of how a trained model can be made to act as if some of its training instances were never encountered without having to retrain the model from scratch. Interest in the topic has increased as privacy regulations such as GDPR have been introduced and people's demands to have control over their data have grown. Removal methods, verification and proof-of-unlearning schemes and audit analyses exist in the literature. From the papers that we reviewed however, none combine the concepts of consent withdrawal, scoped removal, verification, and an auditable record in one workflow. We aim to propose a framework that we call *Consent-Aware Machine Unlearning* (CAMU) that initiates unlearning when a person withdraws consent, verifies that unlearning is actually taking place, and documents this process in an auditable record. We review recent proposals, compare them on the issues of consent handling and verifiability, and identify the gap created, before discussing three potential approaches to verifying removal: membership-inference auditing, influence-function auditing, and certified removal bounds. The design proposal is called CAMU. We do not report any experiments and implementation and evaluation is left for future research.

**Keywords :** Machine unlearning, consent management, data privacy, GDPR, right to be forgotten, verification, privacy-preserving AI, trustworthy AI.

## I. Introduction

The datasets that most of the machine learning (ML) models rely on these days are huge, and they often include personal information. Provide individuals control over their data under the General Data

BODHIVRUKSHA JOURNAL OF DIVERSE DISCIPLINE
E-ISSN : 3139-1486
An Open Access, Peer-Reviewed, Multidisciplinary Scholarly Journal
Bi-monthly | Multilingual (English, Hindi, Marathi) | Academic Research
@ Volume 2, Issue 5 (September-October 2026) (Special Issue 1) ISSN: 3139-1486 (Online)

Protection Regulation (GDPR) and other legislation, This includes a request to have it taken down. However, when somebody withdraws his consent, once it has been stored, it doesn't complete the job: the model trained on it may still represent it. Redoing the whole training run for each request costs too much to be realistic, particularly for large models.

A novel method to remove the influence of selected data from a completed model is provided by *Machine unlearning* which does not diminish the utility of the model [17], [18], [20], [24]. However, most previous studies make the assumption that forgotten data is just given. A couple of questions, which are important for practice are not discussed as much:

1) **Consent as a trigger.** How a withdrawal is recorded, tied to the correct data, and converted into an unlearning request.
2) **Verifiability within the workflow.** How a data subject, auditor, or regulator can confirm that the data's influence has really disappeared, with the check linked to the specific erasure request and stored for later review.

This paper contributes in the following aspects:

- A structured comparison of representative unlearning approaches in terms of consent handling and verification (Section III).
- A conceptual framework, CAMU, that links consent withdrawal, unlearning, verification, and audit logging (Section IV).
- A discussion of candidate verification strategies, evaluation criteria, open challenges, and future directions (Sections V–VIII).

To the best of our knowledge, and given that our search was not systematic, none of the works we reviewed combines these steps in one workflow (Section III). What CAMU contributes is this integration, along with its scope-resolution and escalation steps, and not any single component on its own.

*Scope note:* The framework is conceptual. This paper reports no implementation and no experimental evaluation.

## II. Background

### A. Machine Unlearning

The training set is denoted by $D$ and the learning algorithm by $\mathcal{A}$, the result of which is the trained model $\theta = \mathcal{A}(D)$. If some of the data has to be forgotten, we partition the data into $D_f \subseteq D$ (forget set) and $D_r = D \setminus D_f$ (retain set). An unlearning procedure $\mathcal{U}$, is given the trained model, as well as both sets, and outputs a new model,

$$\theta_u = \mathcal{U}(\theta, D_f, D_r). \tag{1}$$

The aim is for $\theta_u$ to behave as closely as possible to a model trained without $D_f$. Methods fall into two broad groups:

BODHIVRUKSHA JOURNAL OF DIVERSE DISCIPLINE
E-ISSN : 3139-1486
An Open Access, Peer-Reviewed, Multidisciplinary Scholarly Journal
Bi-monthly | Multilingual (English, Hindi, Marathi) | Academic Research
@ Volume 2, Issue 5 (September-October 2026) (Special Issue 1) ISSN: 3139-1486 (Online)

- **Exact unlearning** yields a model $\theta_u$ that is the same as $\mathcal{A}(D_r)$ or follows the same distribution, i.e., $\theta_u \overset{d}{=} \mathcal{A}(D_r)$; sharded training in the style of SISA [17] is a typical example.
- **Approximate unlearning**: $\theta_u$ only lands near $\mathcal{A}(D_r)$, which weakens the guarantee but saves effort [18]–[20].

Surveys [3], [4], [8], [9], [13] cover taxonomies, metrics, and application scenarios.

### B. Certified Removal

C. Guo et al. [19] proposed $\epsilon$-certified removal, a notion modelled on differential privacy that quantifies how thoroughly a model has lost the influence of particular training data. A removal mechanism $\mathcal{M}$ satisfies it when, for every measurable set $\mathcal{T}$ of models,

$$e^{-\epsilon} \leq \frac{\Pr(\mathcal{M}(\mathcal{A}(D), D, x) \in \mathcal{T})}{\Pr(\mathcal{A}(D \setminus \{x\}) \in \mathcal{T})} \leq e^{\epsilon}. \tag{2}$$

In other words, the output of the unlearning step should be hard to tell apart from a model retrained without the data point $x$. The smaller $\epsilon$ is, the stronger the guarantee. (This is the multiplicative, $\delta = 0$ form; [19] also give a relaxed $(\epsilon, \delta)$ version.)

### C. Regulatory Context

The GDPR [23] allows data subjects to withdraw their consent as well as to seek the deletion of their personal data. The problem is that these rights need to be turned into properties of a trained model that can be tested.

## III. Related Work

Work published from 2024 onward has leaned towards surveys, verification, and governance. On the governance side, Lin et al. [1] (2026) treat the auditing of unlearning compliance as a game between an operator and an auditor, while Hutson et al. [2] (2026) and Cooper et al. [6] (2024) look at how unlearning relates to privacy law and where technical capabilities fall short of what the law expects. Recent surveys by Nguyen et al. [3] (2025), Li et al. [4] (2025), Shaik et al. [8] (2024), and Y. Xu [9] (2024) review techniques, evaluation metrics, applications, and open problems. Li et al. give a structured overview of verification methods and metrics, and Y. Xu also covers large language models. Xue et al. [5] (2025) focus specifically on verification. On that front, Zhang et al. [7] (2024) show that existing verification strategies can be circumvented by a dishonest provider, Weng et al. [10] (2024) introduce proof of unlearning and instantiate it with trusted hardware, and Y. Guo et al. [11] (2024) verify unlearning with invisible backdoor triggers.

Previous literature (2019–2023) established a foundation of the primary concepts for removal and verification. Zhang et al. [12] (2023) suggest a consent-tagging pattern for online data: when consumers withdraw their consent, the model is retrained or unlearned; it is then recorded on a distributed ledger. It is demonstrated as a proof of concept and as far as we could determine, it lacks any step to determine if the unlearning took effect. Xu et al. (2023) [13] reviews unlearning methods and their validation methods. Thudi et al. [14] (2022) suggest that unlearning can only be properly defined at algorithm level, with the only verifiable assertion that a model owner can make being that they employed an algorithm

BODHIVRUKSHA JOURNAL OF DIVERSE DISCIPLINE
E-ISSN : 3139-1486
An Open Access, Peer-Reviewed, Multidisciplinary Scholarly Journal
Bi-monthly | Multilingual (English, Hindi, Marathi) | Academic Research
@ Volume 2, Issue 5 (September-October 2026) (Special Issue 1) ISSN: 3139-1486 (Online)

subject to external scrutiny. Data deletion is suggested to be verified probabilistically using backdoor methods in Sommer et al. (2022) [15]. From the point of view of Unlearning, Jose and Simeone [16] (2021) look at unlearning through PAC-Bayesian and information-risk-minimization theory. Bourtoule et al. [17] (2021) proposed SISA training, which shards the data and trains isolated sub-models so that only the data's shard needs to be retrained, while Graves et al. [18] (2021) proposed Amnesiac Machine Learning, which selectively rolls back the parameter updates associated with the data being deleted. C. Guo et al. (2020) [19] provided certifications of data removal and Ginart et al. (2019) [20] investigated efficient data removal.

The idea itself goes back to Cao and Yang [24] (2015), who introduced machine unlearning as a way of making systems forget.

Table I summarizes representative works against the two aspects we care about, and Fig. 1 groups the removal methods by approach.

**Research gap.** Existing work covers many pieces of the problem: removal methods, verification and proof-of-unlearning techniques, audit and governance analyses, and a consent-triggered retraining or unlearning step for web data. Across the papers we looked at, none covers all five of the steps below inside one workflow for a trained model: (i) a consent registry recording the purpose and lawful basis of processing, (ii) scoping a withdrawal down to the forget set, (iii) removal with a suitable method and escalation, (iv) verification, and (v) an audit record tied to that individual request. Our search was not a systematic review, so this is a gap among the papers we examined and not proof that no such paper exists. It motivates the framework that follows.

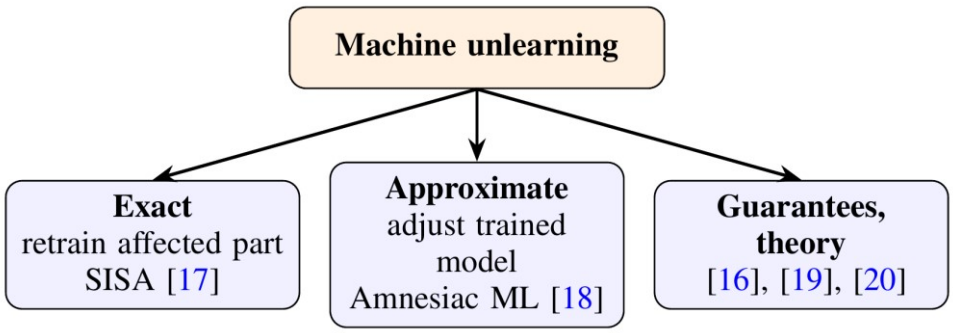



Fig. 1. Illustrative grouping of the unlearning works discussed in this section (not a formal taxonomy).

BODHIVRUKSHA JOURNAL OF DIVERSE DISCIPLINE
E-ISSN : 3139-1486
An Open Access, Peer-Reviewed, Multidisciplinary Scholarly Journal
Bi-monthly | Multilingual (English, Hindi, Marathi) | Academic Research
@ Volume 2, Issue 5 (September-October 2026) (Special Issue 1) ISSN: 3139-1486 (Online)

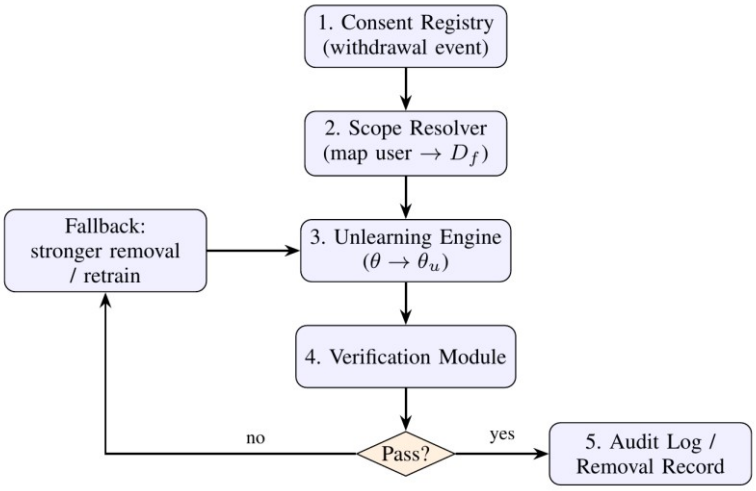



Fig. 2. Conceptual CAMU pipeline: consent withdrawal triggers scoped unlearning, followed by verification and audit logging.

## IV. Proposed Consent-Aware Unlearning Framework

CAMU is a conceptual pipeline with five components (Fig. 2).

1) **Consent Registry.** Stores each data subject's consent status, purpose, legal basis, and timestamps. A withdrawal event sets the pipeline going for data held under that consent. Data that must be kept under a different lawful basis [23] is a matter of policy and is outside the scope of this paper.

2) **Scope Resolver.** Maps the withdrawal to the concrete forget set $D_f$ (records, derived features, and, where applicable, data used in earlier model versions). Getting this right matters: remove too little and the request is not honoured, remove too much and the model loses utility.

3) **Unlearning Engine.** Applies an existing technique, e.g. shard-based retraining [17], update reversal [18], or certified removal [19], chosen to suit the model type and the cost budget.

4) **Verification Module.** Checks whether the influence of $D_f$ has been removed and whether utility on the retained data is preserved (Section V).

5) **Audit Log.** Documents the request, how it was done, the verification results and the version of the model that was created, for later review.

If verification fails, the pipeline escalates to a stronger removal strategy, up to full retraining, before the request is marked complete.

### A. Illustrative Walkthrough

Take a service that has trained a model on user interaction logs. Fig. 3 follows one request through CAMU. The data subject withdraws consent for a given purpose (step 1), and the registry emits a withdrawal event naming the user and the purpose (step 2). The scope resolver maps that event to the forget set $D_f$, derived features included, and passes it on (step 3). The unlearning engine picks a method that fits the model and cost budget, produces $\theta_u$, and hands it to the verification module (step 4). If a

BODHIVRUKSHA JOURNAL OF DIVERSE DISCIPLINE
E-ISSN : 3139-1486
An Open Access, Peer-Reviewed, Multidisciplinary Scholarly Journal
Bi-monthly | Multilingual (English, Hindi, Marathi) | Academic Research
@ Volume 2, Issue 5 (September-October 2026) (Special Issue 1) ISSN: 3139-1486 (Online)

check fails, the engine is asked to escalate to a stronger method (step 4b). After the checks succeed, the outcome and the chosen method are written to the audit log (step 5), producing a removal record that the data subject or an auditor can inspect (step 6). The walkthrough is only illustrative and does not come from an implementation.

TABLE I

COMPARISON OF EXISTING MACHINE UNLEARNING APPROACHES

| Work | Year | Method | Main finding | Limitation (w.r.t. this paper) |
|---|---|---|---|---|
| Lin et al. [1] (preprint) | 2026 | Game-theoretic audit model | Models operator–auditor interaction for unlearning compliance | Economic and regulatory focus; no technical pipeline |
| Nguyen et al. [3] | 2025 | Survey, taxonomy | Reviews techniques, removal requests, applications, challenges | No integrated consent-aware verification framework |
| Li et al. [4] | 2025 | Taxonomy, evaluation | Methods, verification and evaluation metrics, applications, attacks, future directions | Metrics are surveyed, but not linked to consent-triggered removal requests |
| Shaik et al. [8] | 2024 | Taxonomy, comparison | Classifies approaches; discusses metrics, datasets, challenges | Does not integrate consent withdrawal with verifiable removal |
| Y. Xu [9] | 2024 | Comparative survey | Traditional models vs. large language models; evaluation criteria | No unified consent-aware mechanism |
| Weng et al. [10] | 2024 | Proof of unlearning | Algorithm-level definitions with a trusted-hardware instantiation | Relies on trusted hardware; no consent management |
| Zhang et al. [12] (preprint) | 2023 | Consent-tagging framework | Consent withdrawal triggers retraining or unlearning, logged on a ledger; proof of concept | Web-data setting; no verification stage found |
| Xu et al. [13] | 2023 | Survey | Exact, approximate, model-specific unlearning; summarizes verification methods | Verification is reviewed, not embedded in a consent-driven workflow |
| Thudi et al. [14] | 2022 | Theory of auditable definitions | There is only an algorithm level definition of unlearning; auditable claims require an auditable algorithm | No consent-driven workflow |

BODHIVRUKSHA JOURNAL OF DIVERSE DISCIPLINE
E-ISSN : 3139-1486
An Open Access, Peer-Reviewed, Multidisciplinary Scholarly Journal
Bi-monthly | Multilingual (English, Hindi, Marathi) | Academic Research
@ Volume 2, Issue 5 (September-October 2026) (Special Issue 1) ISSN: 3139-1486 (Online)

| Work | Year | Method | Main finding | Limitation (w.r.t. this paper) |
|---|---|---|---|---|
| Sommer et al. [15] | 2022 | Backdoor-based verification | Probabilistic verification of data deletion in ML-as-a-service | Verification only; no consent trigger or audit workflow |
| Jose & Simeone [16] | 2021 | PAC-Bayesian framework | Theoretical framework for unlearning | Primarily theoretical; consent-driven practice not addressed |
| Bourtoule et al. [17] | 2021 | SISA framework | Lower unlearning cost via data/model partitioning | Does not address consent management |
| Graves et al. [18] | 2021 | Selective removal | Targeted removal of training-data influence | Possible compute/storage overhead; no explicit consent verification |
| C. Guo et al. [19] | 2020 | Certified removal | Theoretical guarantees limiting removed-data influence | Assumptions may not match real consent workflows |
| Ginart et al. [20] | 2019 | Data deletion methods | Efficient deletion of data influence | Focuses on efficiency rather than consent and verification |

### B. Objective

The unlearning step can be seen as a balance between three goals:

$$\min_{\theta_u} \underbrace{\mathcal{R}(\theta_u;\, D_f)}_{\text{residual influence}} + \lambda \underbrace{\mathcal{L}_{\text{util}}(\theta_u;\, D_r)}_{\text{utility loss}} + \mu \underbrace{\mathcal{C}(\mathcal{U})}_{\text{cost}}, \qquad \lambda, \mu \geq 0, \tag{3}$$

where $\mathscr{R}$ measures the influence of $D_f$ that remains, $\mathscr{L}_{\text{util}}$ the loss of performance on retained data, $\mathcal{C}$ the computational cost of the procedure $\mathcal{U}$ from Eq. (1), and $\lambda$, $\mu$ are trade-off weights. Equation (3) is a design abstraction and not a specific algorithm.

### C. Threat Model Sketch

Because CAMU produces records that other people may rely on, its failure modes matter. We assume an operator who may be careless or, in the worst case, dishonest, and we consider four concerns:

BODHIVRUKSHA JOURNAL OF DIVERSE DISCIPLINE
E-ISSN : 3139-1486
An Open Access, Peer-Reviewed, Multidisciplinary Scholarly Journal
Bi-monthly | Multilingual (English, Hindi, Marathi) | Academic Research
@ Volume 2, Issue 5 (September-October 2026) (Special Issue 1) ISSN: 3139-1486 (Online)

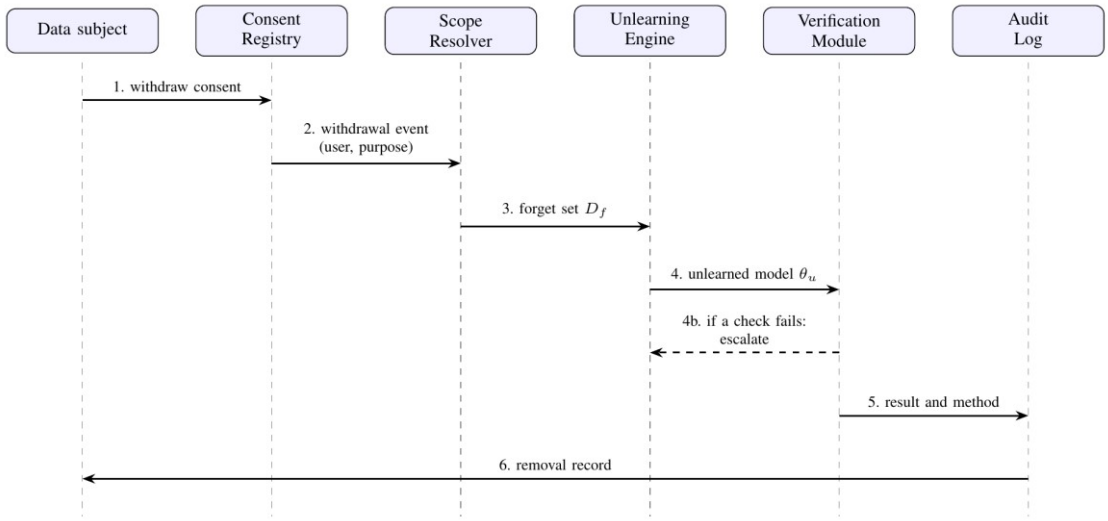



Fig. 3. Sequence of one consent-withdrawal request through the CAMU components (conceptual).

- **Forged or replayed requests:** someone submits a withdrawal on behalf of another person.
- **False claims of removal:** the operator reports success without actually performing or verifying the unlearning. Current verification schemes can be fooled by a dishonest provider [7].
- **Log tampering:** audit records are altered after the fact.
- **Leakage through the pipeline:** logs or verification outputs (e.g., membership scores) reveal information about the data being removed.

Possible design directions include authenticated withdrawal requests, append-only hash-chained audit records, logging identifiers and digests instead of raw data, and having a third-party auditor re-run the verification independently. These are directions for future work, not mechanisms we have evaluated.

## V. Candidate Verification Strategies

Verifying unlearning is hard because model parameters do not directly reveal which data shaped them. The strategies below are directions for future work, and none is implemented here. Fig. 4 shows how the verification module could combine them: a utility check on retained data first, then residual-influence audits, with escalation if either check fails. The tolerances are policy choices and would have to be set separately for each application.

### A. Membership Inference as a Residual-Influence Indicator

A membership inference attack (MIA) [22] attempts to determine if a record was in the training set. Once unlearned, the chance of success of the attack should approach the same rate as it would on records never encountered during the learning process, otherwise there is residual influence. This can be expressed in a simple format, the MIA advantage

$$\mathrm{Adv}_{\mathrm{MIA}} = \Big| \Pr\left[\mathrm{MIA}(x) = 1 \mid x \in D_f\right] - \Pr\left[\mathrm{MIA}(x) = 1 \mid x \notin D\right] \Big|, \tag{4}$$

BODHIVRUKSHA JOURNAL OF DIVERSE DISCIPLINE
E-ISSN : 3139-1486
An Open Access, Peer-Reviewed, Multidisciplinary Scholarly Journal
Bi-monthly | Multilingual (English, Hindi, Marathi) | Academic Research
@ Volume 2, Issue 5 (September-October 2026) (Special Issue 1) ISSN: 3139-1486 (Online)

where values near zero indicate little residual influence. The limitation is that a failed attack is evidence of removal, not proof.

### B. Influence-Function-Based Auditing

Influence functions [21] approximate how much a training point affects model parameters or predictions. Estimating the influence of $D_f$ on the unlearned model can show whether its contribution has shrunk. The drawbacks are the cost of the Hessian-related computation and the quality of the approximation for deep, non-convex models.

### C. Certified Removal Bounds

Where the learning setting allows it, certified removal [19] gives a quantifiable guarantee of the form in Eq. (2). Of the three strategies it is the strongest formally, but note that it certifies the removal mechanism under stated assumptions (for example, convex or well-behaved models) and does not verify an arbitrary deployed model.

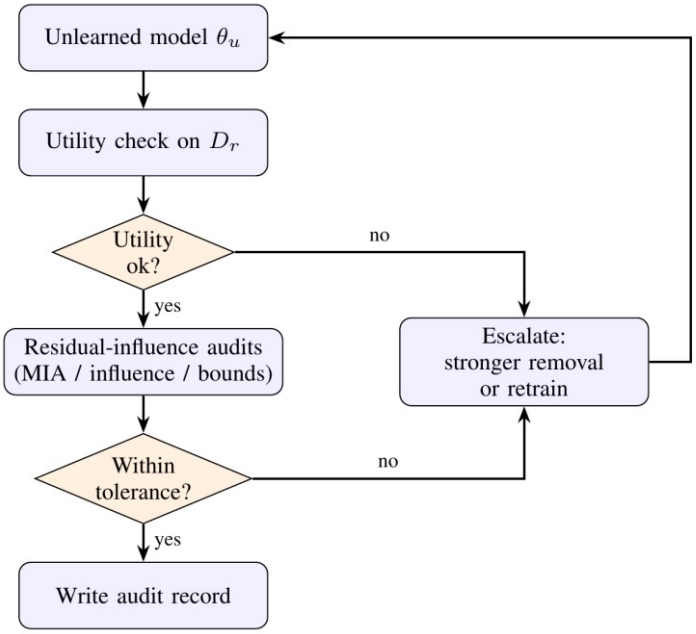



Fig. 4. Conceptual verification and escalation flow of the CAMU verification module.

TABLE II

CANDIDATE VERIFICATION STRATEGIES (CONCEPTUAL).

| Strategy | Strength | Weakness |
|---|---|---|
| MIA-based | Model-agnostic, empirical | Evidence only; attack-dependent |
| Influence-based | Targets specific points | Costly; approximation error |
| Certified bounds | Formal guarantee | Restrictive assumptions |

BODHIVRUKSHA JOURNAL OF DIVERSE DISCIPLINE
E-ISSN : 3139-1486
An Open Access, Peer-Reviewed, Multidisciplinary Scholarly Journal
Bi-monthly | Multilingual (English, Hindi, Marathi) | Academic Research
@ Volume 2, Issue 5 (September-October 2026) (Special Issue 1) ISSN: 3139-1486 (Online)

### D. Comparison

## VI. Proposed Evaluation Criteria

A future implementation of CAMU could be assessed along these axes:

- **Forgetting quality:** how much better a membership-inference attack performs on $D_f$ than chance (Eq. (4)), and how far accuracy on $D_f$ differs from that of a model retrained on $D_r$.
- **Utility:** accuracy or task performance on retained and held-out data.
- **Efficiency:** time and memory compared with full retraining.
- **Consent fidelity:** correctness of scope resolution (no missed or over-deleted data).
- **Auditability:** completeness and reproducibility of the audit record.

## VII. Challenges and Limitations

- **No universal verification.** Empirical tests can miss residual influence, and formal guarantees rest on strong assumptions. A dishonest provider can also get around verification [7], and only algorithm-level claims are formally auditable [14]. For this reason the verification module in CAMU supplies evidence for an audit, not proof.
- **Scoping.** Data can be duplicated, transformed, or aggregated, which makes it hard to find all the influence of one user's data.
- **Scale.** With large models, including large language models [9], repeated unlearning and verification get costly.
- **Sequential requests.** Handling many deletions one after another could erode accuracy or loosen the guarantees.
- **Legal interpretation.** What counts as sufficient removal under regulations such as the GDPR [23] has not been fully pinned down in technical terms.
- **Security of the pipeline.** The consent registry and the audit log need protection against tampering, and they must not leak the very data being removed.

This work is also limited by being a literature-based proposal with no experiments behind it.

## VIII. Future Directions

- Implement a prototype of CAMU on benchmark datasets and compare the verification strategies of Section V.
- Combine empirical audits with certified guarantees to obtain layered assurance.
- Study cryptographically verifiable or tamper-evident audit records.
- Extend consent-aware unlearning to large language models and federated settings.
- Establish common norms and indicators for consent-based de-programming.

BODHIVRUKSHA JOURNAL OF DIVERSE DISCIPLINE
E-ISSN : 3139-1486
An Open Access, Peer-Reviewed, Multidisciplinary Scholarly Journal
Bi-monthly | Multilingual (English, Hindi, Marathi) | Academic Research
@ Volume 2, Issue 5 (September-October 2026) (Special Issue 1) ISSN: 3139-1486 (Online)

## IX. Conclusion

Machine unlearning is an important tool for bringing ML systems in line with privacy regulation, yet none of the works we reviewed connects consent withdrawal, scoped removal, verification, and an auditable record in a single workflow. This paper reviewed representative approaches, identified that gap, and proposed the conceptual CAMU framework, which couples consent-triggered unlearning with verification and auditing. We also outlined candidate verification strategies and evaluation criteria to guide future implementation and empirical study.

## Acknowledgment

It is a great honor for me to express my sincere thanks to Dr. Vishal Puranik for his valuable advice and guidance throughout the completion of this project.

## References

[1] Q. Lin, N. Ding, L. Duan, and J. Huang, "Governing AI Forgetting: Auditing for Machine Unlearning Compliance," arXiv:2602.14553, 2026.

[2] J. Hutson, C. Whitney, and J. T. Conrad, "Forget Me Not? Machine Unlearning's Implications for Privacy Law," *Columbia Sci. Technol. Law Rev.*, vol. 27, no. 1, 2026, doi:10.52214/stlr.v27i1.14547.

[3] T. T. Nguyen, T. T. Huynh, Z. Ren, P. L. Nguyen, A. W.-C. Liew, H. Yin, and Q. V. H. Nguyen, "A Survey of Machine Unlearning," *ACM Trans. Intell. Syst. Technol.*, 2025.

[4] N. Li, C. Zhou, Y. Gao, H. Chen, A. Fu, Z. Zhang, and Y. Shui, "Machine Unlearning: Taxonomy, Metrics, Applications, Challenges, and Prospects," *IEEE Trans. Neural Netw. Learn. Syst.*, 2025.

[5] L. Xue *et al.*, "Towards Reliable Forgetting: A Survey on Machine Unlearning Verification," arXiv:2506.15115, 2025.

[6] A. F. Cooper *et al.*, "Machine Unlearning Doesn't Do What You Think: Lessons for Generative AI Policy, Research, and Practice," arXiv:2412.06966, 2024.

[7] B. Zhang, Z. Chen, C. Shen, and J. Li, "Verification of Machine Unlearning is Fragile," in *Proc. Int. Conf. Machine Learning (ICML)*, PMLR 235, 2024, pp. 58717–58738.

[8] T. Shaik, X. Tao, H. Xie, L. Li, X. Zhu, and Q. Li, "Exploring the Landscape of Machine Unlearning: A Comprehensive Survey and Taxonomy," *IEEE Trans. Neural Netw. Learn. Syst.*, 2024.

[9] Y. Xu, "Machine Unlearning for Traditional Models and Large Language Models: A Short Survey," arXiv:2404.01206, 2024.

[10] J. Weng, S. Yao, Y. Du, J. Huang, J. Weng, and C. Wang, "Proof of Unlearning: Definitions and Instantiation," *IEEE Trans. Inf. Forensics Security*, vol. 19, pp. 3309–3323, 2024.

[11] Y. Guo, Y. Zhao, S. Hou, C. Wang, and X. Jia, "Verifying in the Dark: Verifiable Machine Unlearning by Using Invisible Backdoor Triggers," *IEEE Trans. Inf. Forensics Security*, vol. 19, pp. 708–721, 2024.

BODHIVRUKSHA JOURNAL OF DIVERSE DISCIPLINE
E-ISSN : 3139-1486
An Open Access, Peer-Reviewed, Multidisciplinary Scholarly Journal
Bi-monthly | Multilingual (English, Hindi, Marathi) | Academic Research
@ Volume 2, Issue 5 (September-October 2026) (Special Issue 1) ISSN: 3139-1486 (Online)

[12] D. Zhang *et al.*, "Tag Your Fish in the Broken Net: A Responsible Web Framework for Protecting Online Privacy and Copyright," arXiv:2310.07915, 2023.

[13] H. Xu, T. Zhu, L. Zhang, W. Zhou, and P. S. Yu, "Machine Unlearning: A Survey," *ACM Comput. Surv.*, vol. 56, no. 1, Art. 9, 2023.

[14] A. Thudi, H. Jia, I. Shumailov, and N. Papernot, "On the Necessity of Auditable Algorithmic Definitions for Machine Unlearning," in *Proc. USENIX Security Symp.*, 2022, pp. 4007–4022.

[15] D. M. Sommer, L. Song, S. Wagh, and P. Mittal, "Athena: Probabilistic Verification of Machine Unlearning," *Proc. Privacy Enhancing Technologies (PoPETs)*, vol. 2022, no. 3, pp. 268–290, 2022.

[16] S. T. Jose and O. Simeone, "A Unified PAC-Bayesian Framework for Machine Unlearning via Information Risk Minimization," in *Proc. IEEE Int. Workshop Machine Learning for Signal Processing (MLSP)*, 2021.

[17] L. Bourtoule *et al.*, "Machine Unlearning," in *Proc. IEEE Symp. Security and Privacy (SP)*, 2021, pp. 141–159.

[18] L. Graves, V. Nagisetty, and V. Ganesh, "Amnesiac Machine Learning," in *Proc. AAAI Conf. Artificial Intelligence*, vol. 35, 2021, pp. 11516–11524.

[19] C. Guo, T. Goldstein, A. Hannun, and L. van der Maaten, "Certified Data Removal from Machine Learning Models," in *Proc. Int. Conf. Machine Learning (ICML)*, 2020, pp. 3832–3842.

[20] A. Ginart, M. Guan, G. Valiant, and J. Zou, "Making AI Forget You: Data Deletion in Machine Learning," in *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 32, 2019.

[21] P. W. Koh and P. Liang, "Understanding Black-box Predictions via Influence Functions," in *Proc. Int. Conf. Machine Learning (ICML)*, 2017, pp. 1885–1894.

[22] R. Shokri, M. Stronati, C. Song, and V. Shmatikov, "Membership Inference Attacks Against Machine Learning Models," in *Proc. IEEE Symp. Security and Privacy (SP)*, 2017, pp. 3–18.

[23] European Parliament and Council of the European Union, "Regulation (EU) 2016/679 (General Data Protection Regulation)," *Official Journal of the European Union*, L 119, pp. 1–88, 2016.

[24] Y. Cao and J. Yang, "Towards Making Systems Forget with Machine Unlearning," in *Proc. IEEE Symp. Security and Privacy (SP)*, 2015, pp. 463–480.