EXISTE-T-IL UN DROIT A L'OUBLI AVEC CHATGPT ?

À mesure que les assistants conversationnels deviennent nos confidents du quotidien, une question essentielle s'impose : les IA peuvent-elles réellement oublier ce que nous leur avons confié ? Entre opacité technique, limites du RGPD et absence de garanties probatoires, Me Alexandra Iteanu montre pourquoi le droit à l'effacement reste largement théorique face aux modèles de langage, et appelle à un véritable réarmement juridique.

Avez-vous déjà demandé à votre chatbot préféré (ChatGPT, Claude, Llama, Perplexity…) ce qu'il savait de vous ? L'exercice est déroutant. La réponse, elle, est souvent glaçante.

Car ces grands modèles de langages (« LLM ») sont désormais devenus des confidents du quotidien : nous leur livrons nos inquiétudes, nos questions les plus personnelles, nos hésitations professionnelles, parfois même nos émotions les plus intimes. Mais ces « journaux intimes numériques » ont une particularité que l'on oublie trop souvent : rien de ce que nous leur confions ne reste sagement enfermé dans un tiroir.

Au contraire, ces données personnelles sont hébergées sur des serveurs externes, analysées, et — sauf exception — réutilisées pour entraîner les modèles. Cette externalisation massive de notre intimité expose les utilisateurs à plusieurs risques : compromission en cas de cyberattaque, exploitation à des fins d'amélioration algorithmique, mais aussi communication à des partenaires technologiques et, dans certains cas, aux autorités publiques, comme le prévoit la politique de confidentialité de certains éditeurs américains.

Reste alors une question essentielle, presque vertigineuse : peut-on vraiment effacer ces informations et reprendre le contrôle sur ce que ces IA savent de nous ?

Quand l'opacité des modèles rend l'utilisateur… introuvable

Les données à caractère personnel, c'est-à-dire toute information qui permet d'identifier directement ou indirectement une personne physique[1], sont traitées à très grande échelle par ces LLM.

Elles sont collectées par les éditeurs de ces services via plusieurs canaux : (1) les prompts et les informations que l'utilisateur fournit volontairement dans le cadre de l'utilisation du service ; (2) le web scraping ; (3) l'achat de bases de données personnelles auprès de data brokers.

C'est ce que reconnait l'éditeur OpenAI dans sa Politique de confidentialité disponible sur son site web[2].

Au moment de l'entrainement de ces modèles, on parlera alors de « données d'entrainement » ou « d'apprentissage », d'importants jeux de données sont utilisés. La quantité et la qualité de ces informations détermineront directement la puissance de ces modèles. Autrement dit, plus ces modèles sont nourris d'informations, y compris privées, plus ils seront performants.

C'est précisément à ce niveau que le droit commence à se heurter au mur technique.

Dans sa dernière publication de novembre 2025, l'EDPB, l'autorité de protection européenne, distingue les modèles de langages dont les bases de données sont « structurées » de celles dont les bases sont « non structurées ». Pour ces derniers, elle admet qu'il peut être difficile, voire impossible, d'identifier au sein de ces jeux de données la personne qui sollicite l'exercice de ses droits[3]. Autrement dit, le droit est là, mais l'identification technique de la personne fait défaut ; ce qui, pour une autorité de protection des données, équivaut à reconnaître l'impuissance du cadre juridique face à l'architecture même de ces systèmes.

L'EDPB souligne expressément la « complexité » de ces modèles d'IA, en particulier les deep neural networks et les LLM : opacité des couches de calcul, enchevêtrement des jeux de données, variabilité des réponses pour un même prompt… tout concourt à rendre le lien entre un individu et les données qui ont servi à entraîner le modèle pratiquement insaisissable.

La CNIL, dans ses fiches pratiques à destination des développeurs de modèles d'IA, entérine ce constat : identifier la personne concernée au sein d'une base d'entraînement peut s'avérer extrêmement difficile. Elle admet que le concepteur du modèle devra, en pratique, réclamer des « informations complémentaires » *[4]* à la personne qui exerce ses droits, afin de tenter de la rattacher à ces données. Et, au titre du *privacy by design*, la CNIL se contente de recommander d'anticiper ces difficultés en définissant dès la conception quelles informations supplémentaires (par exemple, en cas d'homonymie) pourraient permettre une identification fiable. Une recommandation qui, en creux, révèle surtout que l'effectivité du droit dépend aujourd'hui de choix techniques que les éditeurs restent libres… ou non d'opérer.

Un droit à l'effacement théorique, face à des modèles qui « n'oublient » pas

À titre d'illustration, l'éditeur OpenAI permet aux utilisateurs de demander l'effacement de leurs données à caractère personnel, via leur compte, une plateforme dédiée ou une adresse email spécifique.

En pratique, plusieurs obstacles majeurs se dressent immédiatement.

D'abord, l'asymétrie d'information : il est très difficile, pour une personne concernée, de connaître la nature exacte et l'étendue des données effectivement détenues par l'éditeur. Une partie de ces données est collectée en ligne, via des sources publiques ou des tiers, sans que la personne n'en soit informée de manière précise ; par ailleurs, dans le cadre du service, des données relatives à une personne peuvent être communiquées par un autre utilisateur, par exemple lorsqu'un tiers décrit, identifie ou interroge le modèle au sujet d'une personne déterminée.

Ensuite, même lorsque l'utilisateur parvient à cibler les données qu'il souhaite effacer, les LLM traitent et combinent des volumes de données tels qu'ils conservent des « structures ou fragments d'information » difficilement isolables. L'EDPB lui-même reconnaît qu'il peut être extrêmement complexe, techniquement, d'éliminer de manière fiable toute trace de ces données dans les paramètres du modèle.

Par conséquence, l'utilisateur n'a, en pratique, aucune certitude que ses données personnelles ont été intégralement effacées. L'opacité de ces systèmes ne permet pas de vérifier de manière concrète et transparente l'effectivité de cet effacement.

Il aurait été particulièrement opportun de donner aux citoyens des moyens concrets de s'assurer de la bonne application du droit de l'Union, en imposant, par exemple, des obligations de production d'éléments de preuve pertinents à la charge des fournisseurs de modèles. Or, le projet de directive sur la responsabilité en matière d'IA qui devait précisément mettre en place cet arsenal probatoire a été suspendu par la Commission européenne en février dernier. Le signal envoyé est clair : le législateur recule là où la charge de la preuve devrait, au contraire, être renforcée[5].

Droit à l'oubli : les mesures proposées par les autorités de contrôle insuffisantes aujourd'hui

Face à ces impasses, les autorités de contrôle avancent des pistes, tout en en reconnaissant elles-mêmes les limites.

La CNIL propose notamment deux types de mesures pour rendre, au moins partiellement, le droit à l'effacement opérationnel :

- Le réentraînement (ou *machine unlearning*) du modèle afin d'effacer les données au cœur même de l'architecture. La CNIL admet cependant que cette solution peut être « disproportionnée » au regard des coûts techniques et organisationnels.

- Le filtrage des sorties (*output filtering*), qui vise à empêcher le modèle de restituer certaines informations, même si celles-ci subsistent dans les paramètres internes.

Mais l'autorité française qualifie elle-même ces mesures de « *soit coûteuses et délicates, soit imparfaites* »…

L'EDPB, de son côté, se contente de reprendre ces deux leviers — *machine unlearning* et *output filtering* — sans formuler de garanties supplémentaires sur leur capacité réelle à restaurer l'effectivité du droit à l'effacement. On est davantage dans le registre des pistes exploratoires que dans celui d'un cadre juridiquement robuste et opérationnel.

Le droit à l'oubli à l'épreuve des LLM – des pistes concrètes pour pallier ses limites

Le constat qui se dessine est sans détour : un utilisateur juridiquement titulaire d'un droit à l'effacement, des modèles entraînés sur des masses de données opaques, des autorités qui reconnaissent la difficulté, voire l'impossibilité, de rattacher techniquement une personne à ces données, et l'absence d'un véritable régime probatoire opposable aux éditeurs.

En l'état, le droit à l'oubli appliqué aux LLM se limite le plus souvent à une procédure déclarative, sans garantie technique que les données aient effectivement disparu des jeux d'entraînement et des paramètres du modèle.

Le prochain mouvement ne pourra pas se contenter de « recommandations » techniques. Il devra consister à réarmer juridiquement ce droit, en :

- imposant des exigences minimales de traçabilité des jeux de données et des capacités d'*un Learning* vérifiables ;
- renforçant les pouvoirs d'enquête et de sanction des autorités de contrôle et des juges, spécifiquement adaptés aux LLM ;
- et en mettant à la charge des fournisseurs de modèles une véritable obligation de preuve, permettant aux personnes concernées et aux autorités de constater l'effectivité des effacements réalisés.

À défaut, le risque est clair : laisser s'installer une zone de non-droit technique au cœur même d'un droit fondamental pourtant consacré par le RGPD, en offrant aux éditeurs la possibilité de se retrancher derrière la « complexité » ou « l'impossibilité technique » pour ne pas donner pleinement effet aux demandes d'effacement. Cette dérive ferait progressivement du droit à l'oubli l'exception plutôt que la règle, et ne manquerait pas, à terme, de contaminer l'ensemble des autres droits et requêtes adressés à ces modèles.

Pour une vie de citoyen harmonieuse et épanouie, nous avons besoin de conserver la mémoire mais la société nous a également concédé un droit à l'oubli, parfois nécessaire pour nous permettre d'avancer dans la vie. C'est un équilibre que seul le droit dans un Etat de droit démocratique, peut réaliser. Alors, n'oublions pas le droit à l'oubli.

[1] Article 4 du règlement UE n°2016/679 dit « RGPD »

[2] « *Nous recueillons également des informations d'autres sources, comme des informations qui sont accessibles au public sur Internet, pour développer les modèles qui alimentent nos Services*. »

[3] EDPB, 11 Novembre 2025, Guidelines, Guidance for Risk Management of Artificial Intelligence systems

[4] CNIL, Fiche pratique IA n°10, Respecter et faciliter l'exercice de droits des personnes concernées

[5] https://www.solutions-numeriques.com/elon-musk-presente-grok-3-son-logiciel-dia-pour-rivaliser-avec-chatgpt-et-deepseek/

.