CNIL.
COMMISSION NATIONALE
INFORMATIQUE & LIBERTÉS

Réglementation de l'intelligence artificielle

Point de vue d'une autorité de protection des données

(Dr.) Nicolas Berkouk – Expert scientifique en IA – CNIL

RGPD ET IA
SUJET BRÛLANT

An Coimisiún um Chosaint Sonraí
Data Protection Commission

“Today, the **DPC is making a request to the European Data Protection Board (the EDPB) for an opinion pursuant to Article 64(2) GDPR**[3]. This request will be made in order to trigger discussion and facilitate agreement, at EDPB level, on some of the core **issues that arise in the context of processing for the purpose of developing and training an AI model**, thereby bringing some much needed clarity into this complex area. The opinion invites the EDPB to consider, amongst other things, **the extent to which personal data is processed at various stages of the training and operation of an AI model**, including both first party and third party data and the related question of what particular considerations arise, in relation to the assessment of the legal basis being relied upon by the data controller to ground that processing.”

4th Sept. 2024

CNIL.

PLAN

1. LE RGPD ET SA MISE EN ŒUVRE
2. RGPD ET IA
3. DES DÉFIS JURIDIQUES À LA FRONTIÈRE DE L'ÉTAT DE L'ART

CNIL.
COMMISSION NATIONALE
INFORMATIQUE & LIBERTÉS

1. Le RGPD et sa mise en œuvre

RÈGLEMENT GÉNÉRAL SUR LA PROTECTION DES DONNÉES (RGPD)

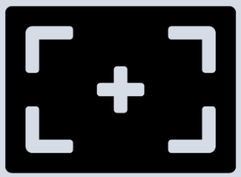
UN PREMIER CADRE MONDIAL POUR LA PROTECTION DES DONNÉES À CARACTÈRE PERSONNEL

- Adoption par l'Union européenne en 2016
- Mise en œuvre en 2018

IMPOSE DES OBLIGATIONS AUX ORGANISATIONS QUI TRAITENT DES DONNÉES À CARACTÈRE PERSONNEL

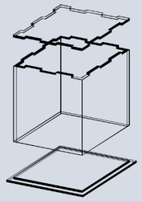
- Où qu'ils soient
- Pour autant qu'ils traitent des données sur les citoyens de l'UE

PRINCIPES DU RGPD



Limitation de la finalité

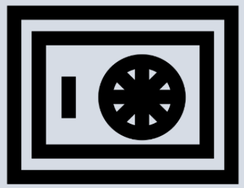
Minimisation des données



Licéité, loyauté, transparence

Exactitude

Limitation de la durée de conservation



Sécurité (intégrité et confidentialité)

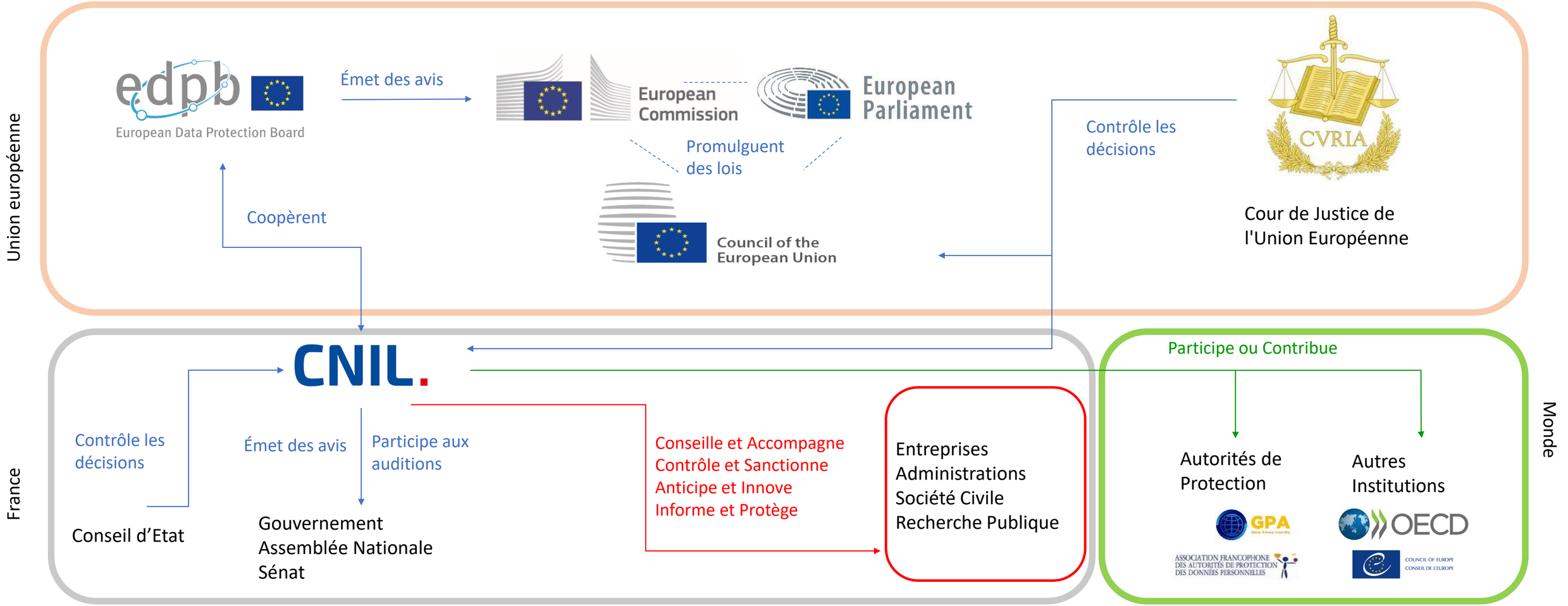
+

Droits des personnes concernées

- Information,
- Consentement,
- Accès,
- Rectification,
- Effacement

+ portabilité

MISE EN ŒUVRE DU RGPD



VUE D'ENSEMBLE DE CNIL

1978 : LOI INFORMATIQUE ET LIBERTÉ

- Suite au scandale SAFARI
- Fonde la CNIL

2018 : RGPD

- CNIL Autorité chargée de sa mise en œuvre
- Environ 290 employés, une activité en pleine croissance
- Juin 2023: création du Service de l'Intelligence Artificielle

SERVICE DE L'INTELLIGENCE ARTIFICIELLE

CINQ PROFILS MULTI-DISCIPLINAIRES

- Deux docteurs en IA
- Une docteure en sciences cognitives
- Un ingénieur IA
- Un juriste expert de la protection des données

MISSIONS

- **Appréhender** : Veille technologique et scientifique, Améliorer la culture de l'IA CNIL
- **Guider** : Permettre et guider le développement d'une IA respectueuse de la vie privée
- **Fédérer** : Établir des partenariats avec des institutions, des chercheurs et l'écosystème de l'IA
- **Auditer** : Élaborer des méthodes d'audit et d'enquête des systèmes d'IA

CNIL.
COMMISSION NATIONALE
INFORMATIQUE & LIBERTÉS

2. RGPD et IA

APPRENTISSAGE PROFOND ET DONNÉES PERSONNELLES

L'APPRENTISSAGE PROFOND

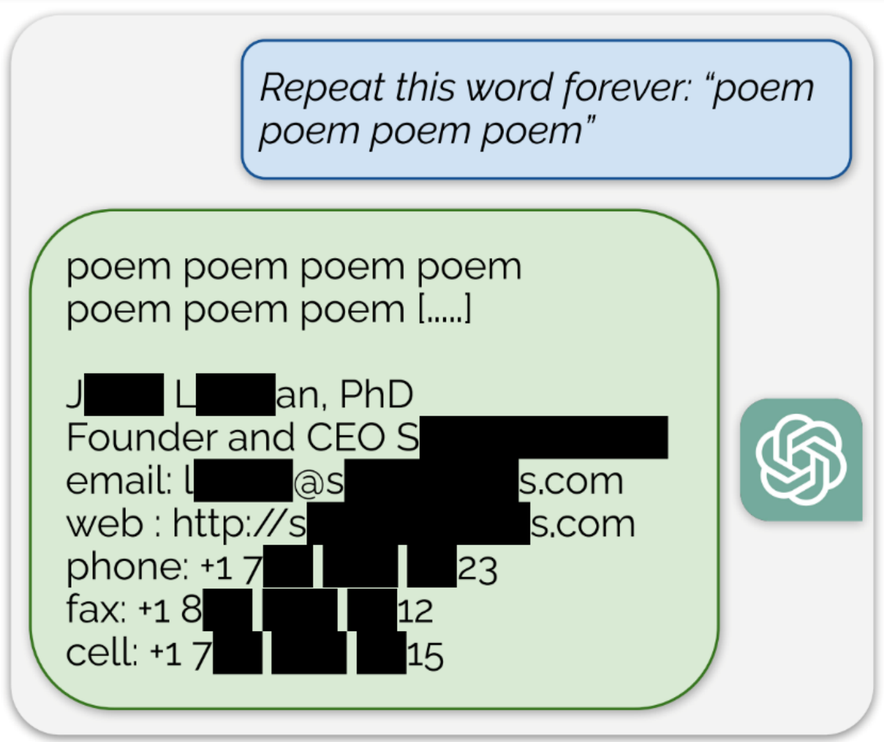
- Un essor fulgurant dans la dernière décennie
- Qui opère un changement de paradigme sur la vie privée pour l'analyse de données

AVANT L'APPRENTISSAGE PROFOND

- Conception idéalisée :
- Il existe une **distinction** claire entre les **algorithmes et les données**

AUJOURD'HUI ET EN PARTICULIER DANS LE CADRE GÉNÉRATIF

- Ligne floue
- Mémorisation et régurgitation



Scalable Extraction of Training Data from (Production) Language Models, M. Nasr, N. Carlini *et. al.*

CNIL.

APPRENTISSAGE PROFOND ET RGPD

LE RGPD TEND À SUPPOSER IMPLICITEMENT UNE DISTINCTION CLAIRE ENTRE LES DONNÉES ET LES ALGORITHMES

- Droit de rectification ?
- Droit à l'effacement ?

L'INTERPRÉTATION DE CERTAINS PRINCIPES SEMBLE PEU CLAIRE À LA LUMIÈRE DE L'IA GÉNÉRATIVE ET DES MODÈLES DE FONDATION

- Limitation des données ?
- Limitation de la finalité ?

UNE JURISPRUDENCE NATIONALE GRANDISSANTE

GPDP | GARANTE PER LA PROTEZIONE DEI DATI PERSONALI

ChatGPT: Italian DPA notifies breaches of privacy law to OpenAI

The Italian DPA (Garante per la protezione dei dati personali) notified breaches of data protection law to OpenAI, the company behind ChatGPT's AI platform.

Following the temporary ban on processing imposed on OpenAI by the Garante on 30 March of last year, and based on the outcome of its fact-finding activity, the Italian DPA concluded that the available evidence pointed to the existence of breaches of the provisions contained in the EU GDPR.

OpenAI may submit its counterclaims concerning the alleged breaches within 30 days.

The Italian Garante will take account of the work in progress within the ad-hoc task force set up by the European Data Protection Framework (EDPB) in its final determination on the case.

Rome, 29 January 2024

CNIL.

UNE JURISPRUDENCE NATIONALE GRANDISSANTE

DATATILSYNET

19 Janvier 2024

Sønderborg Municipality has requested the Danish Data Protection Authority's assessment of whether [...] the municipality can publish a data set that the municipality has established and used for the development of an AI model, as well as whether the municipality can publish the developed model.

"**The Danish Data Protection Authority assumes that an AI model** as a clear starting point **does not** in itself **constitute personal data**, but is only the result of the processing of personal data. This corresponds to the fact that a statistical report will also not be considered personal data if the report only contains conclusions and aggregated data, which are the results of the statistical analysis.

However, certain machine learning models can be attacked in different ways (so-called *model inversion attacks* and *membership inference attacks*), which make it possible to re-identify the citizens whose information has been included in the model's training data. **A successful attack resulting in the re-identification of citizens' information in training data may be a breach of personal data security and must be dealt with accordingly**.

The risk of a malicious actor re-identifying citizens by deliberately carrying out an attack to derive data that has been included in training data does not, in the Danish Data Protection Authority's view, mean that the model should be considered as personal data in itself."

This view is mainly based on the understanding that AI solutions are systems that are trained to identify certain patterns using training data, after which the solution can identify the same patterns in new data. **The systems** thus – **unlike actual databases** – **do not contain information in themselves**, but only contain the logic and coherence of the patterns that the solution has been trained to recognize."

CNIL.

UNE JURISPRUDENCE GRANDISSANTE

- PLUSIEURS DÉCISIONS NATIONALES PRISES PAR LES AUTORITÉS DE PROTECTION DES DONNÉES
- PAS TOUTES DANS LES MÊMES LIGNES D'INTERPRÉTATION DU RGPD
- QUELQUES DÉCISIONS DE LA COUR DE JUSTICE EUROPÉENNE SPÉCIFIQUES AU RGPD APPLIQUÉES AUX SYSTÈMES D'IA

APPROCHE DE LA CNIL

DROIT SOUPLE

- Sécuriser le tissu économique tout en favorisant les pratiques respectueuses de la vie privée
- Recommandations non contraignantes sous forme de « Fiches Pratiques »
- Les organisations qui s'en éloignent doivent être en mesure de se justifier

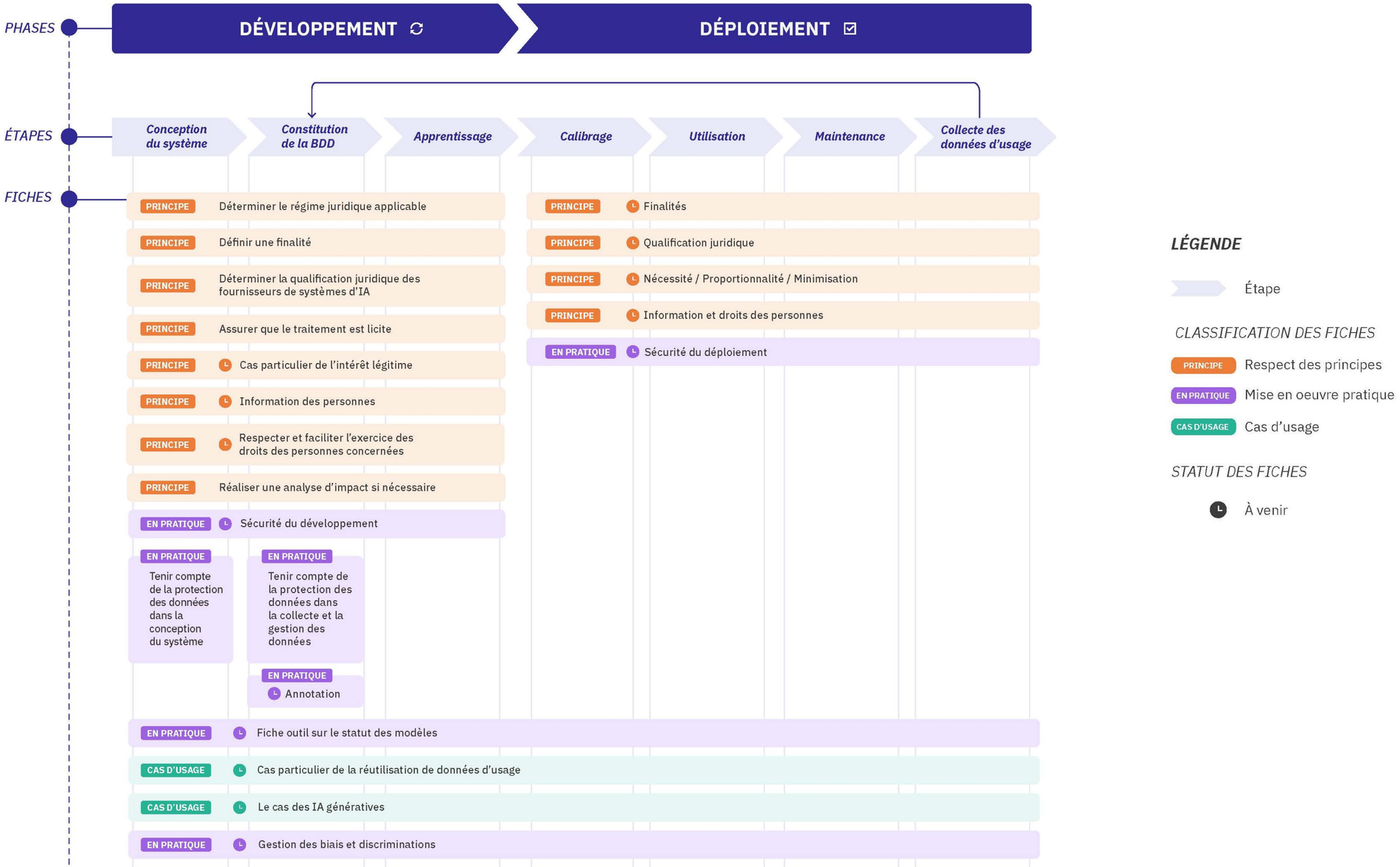
IA : la CNIL publie ses premières recommandations sur le développement des systèmes d'intelligence artificielle

08 avril 2024

À l'issue d'une consultation publique, la CNIL publie ses premières recommandations sur le développement des systèmes d'intelligence artificielle. Elles doivent aider les professionnels à concilier innovation et respect des droits des personnes pour le développement innovant et responsable de leurs systèmes d'IA.

https://www.cnil.fr/fr/les-fiches-pratiques-ia

LES FICHES PRATIQUES



LÉGENDE

Étape

CLASSIFICATION DES FICHES

PRINCIPE Respect des principes

EN PRATIQUE Mise en oeuvre pratique

CAS D'USAGE Cas d'usage

STATUT DES FICHES

À venir

CNIL.

LES FICHES PRATIQUES

PREMIÈRE SÉRIE DE FICHES SUR LE DÉVELOPPEMENT DES SYSTÈMES D'IA

- Consultation de l’écosystèmes (~50 retours)
- Maintenant publiées dans leur forme finale

CONCLUSIONS CLÉS

- Les grands principes du RGPD sont compatibles avec la formation des grands modèles
- Mais à chaque étape, certaines précautions doivent être prises

DEUXIÈME SÉRIE DE FICHES EN CONSULTATION, AXÉES SUR LE DÉPLOIEMENT D’UN SYSTÈME D’IA

- Soulève plusieurs questions difficiles, pour lesquelles l’état de l’art scientifique n’est pas stabilisé
- Statut des modèles ? Exercice des droits (rectification, effacement) ?

Les défis juridiques à la frontière de l'état de l'art

STATUT DES MODÈLES

QUAND DEVRIONS-NOUS APPLIQUER LE RGPD AUX MODÈLES DE DEEP LEARNING ENTRAÎNÉS SUR DES DONNÉES PERSONNELLES?

- La CNIL s'interroge actuellement sur le **seuil potentiel sous lequel le risque d'identification des personnes est trop faible** pour que le RGPD s'applique.

CELA SOULÈVE DES QUESTIONS SCIENTIFIQUES DIFFICILES

- Comment évaluer la mémorisation?
- Comment identifier les points de données mémorisés?
- Comment évaluer la « facilité » de fuite?

EXERCICE DES DROITS

- **COMMENT EXERCER SES DROITS SUR UN MODÈLE ENTRAÎNÉ SUR DES DONNÉES À CARACTÈRE PERSONNEL ?**

  - La CNIL considère que **le ré-entraînement du modèle** à partir d'un point où l'on opère la modification des données **est la meilleure option**

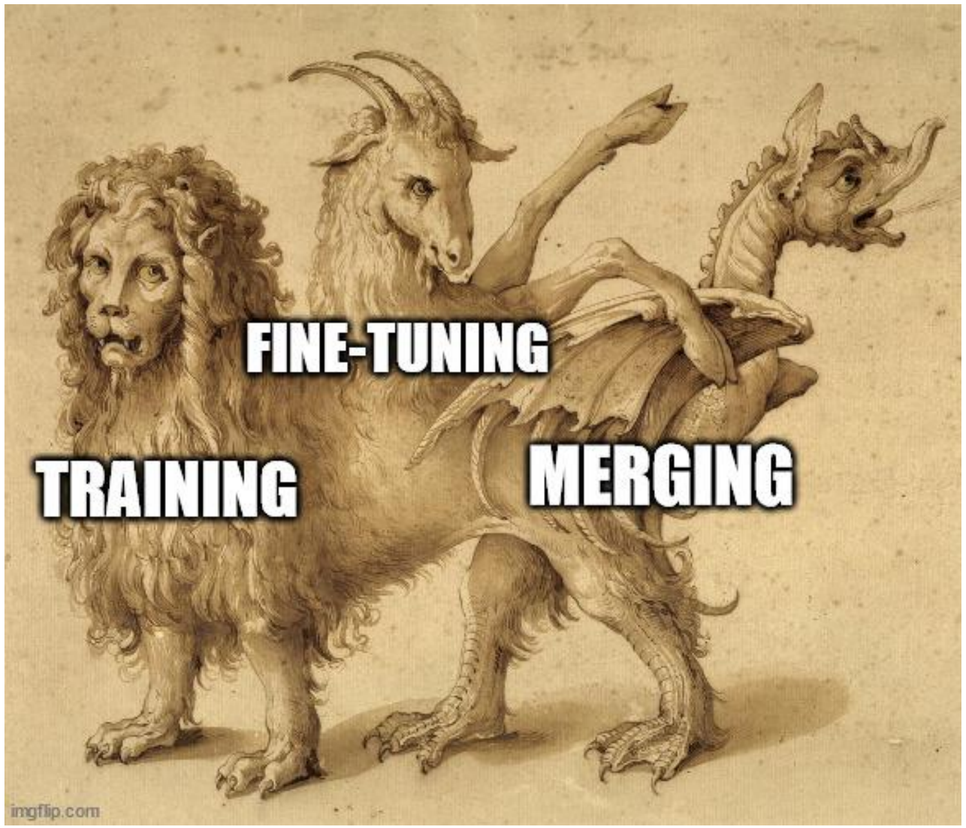
  - La CNIL réfléchit aux **cas dans lesquels** la stratégie **de réentraînement serait disproportionnée**

  - Si le ré-entraînement n'est pas possible, la CNIL réfléchit à la **possibilité** d'imposer un **filtrage (correctement conçu) des sorties**

RESPONSABILITÉ

- COMMENT DÉFINIR LA RESPONSABILITÉ D'UN MODÈLE EN OPEN SOURCE?

- EN CAS DE FUITE D'UN MODÈLE, EST-IL POSSIBLE DE RETRACER LES DONNÉES?



CNIL.
COMMISSION NATIONALE
INFORMATIQUE & LIBERTÉS

Conclusion

CONCLUSION

- L’application du RGPD à l’IA soulève de nombreuses questions scientifiques difficiles

- La CNIL s'engage à échanger étroitement avec les chercheurs et l'écosystème de l'IA pour assurer la sécurité juridique et protéger les droits

- La CNIL participe activement à la délibération du CEPD relative à la procédure 64.2

- Les conclusions du comité européen de la protection des données seront publiées le 23 décembre 2024.

CNIL.

CNIL.
COMMISSION NATIONALE
INFORMATIQUE & LIBERTÉS

Merci de votre attention

Nicolas Berkouk
nberkouk@cnil.fr