

Glossaire IA & RGPD

À destination de tous les agents et étudiants de l'école


Les définitions suivantes reprennent pour l'essentiel les définitions posées par les règlements applicables en matière de protection des données personnelles et d'IA, en particulier le RGPD (règlement UE 2016/679) et l'IA Act (règlement UE 2024/1689).


Elles sont réparties sur la base du code couleur ci-dessous :


 RGPD	 IA Act	 Les deux	 Technique
Protection des données	Réglementation de l'IA	RGPD + IA Act	Concepts IA

Notions RGPD — Protection des données personnelles


AIPD — Analyse d'Impact sur la Protection des Données


 Une étude approfondie, obligatoire pour certains traitements à risque élevé, qui évalue les risques pour la vie privée et prévoit des mesures pour les limiter.


 **Exemple concret** : Mise en place d'un système de surveillance, utilisation d'un outil d'IA traitant des données sensibles d'étudiants.

 **Pourquoi c'est important** : *C'est une obligation légale pour les traitements à fort risque. Elle documente la réflexion et les précautions prises.*


Anonymisation


 Technique qui supprime définitivement tout lien entre une donnée et une personne. La ré-identification doit être impossible.


 **Exemple concret** : Des statistiques globales sur les résultats scolaires sans aucune information permettant d'identifier un étudiant.

 **Pourquoi c'est important** : *Des données vraiment anonymisées ne sont plus soumises au RGPD. Cela facilite leur usage dans des outils d'IA — mais attention, une anonymisation imparfaite reste risquée.*

Classification des données

 Système de catégorisation des données selon leur niveau de sensibilité et l'impact potentiel (public, interne, sensible, confidentiel).

 **Exemple concret** : Une plaquette de présentation de l'école = publique. Un dossier disciplinaire d'étudiant = confidentiel.

 **Pourquoi c'est important** : *Classer ses données permet d'adapter les mesures de protection et de contrôle d'accès à ce qui est vraiment critique.*

Consentement

💡 Manifestation de volonté libre, spécifique, éclairée et univoque par laquelle la personne accepte un traitement, par une déclaration ou un acte positif clair.

🔗 **Exemple concret** : Une case pré-cochée NE constitue PAS un consentement valable. La personne doit faire un geste actif.

⚠️ **Pourquoi c'est important** : *Le consentement n'est qu'une des bases légales possibles. Il doit être vérifiable et la personne peut le retirer à tout moment.*

Destinataire

💡 Toute personne ou organisme qui reçoit des données personnelles.

🔗 **Exemple concret** : Un partenaire académique à qui on envoie des données d'étudiants, un service interne, une administration.

⚠️ **Pourquoi c'est important** : *Cartographier les flux de données (qui reçoit quoi) est indispensable pour maîtriser leur circulation.*

Données à caractère personnel

💡 Toute information se rapportant à une personne physique identifiée ou identifiable, directement ou indirectement.

🔗 **Exemple concret** : Nom, adresse e-mail, numéro de téléphone, photo, adresse IP, données de localisation...

⚠️ **Pourquoi c'est important** : *Dès qu'on manipule ce type de données, le RGPD s'applique. C'est le point de départ de toute démarche de conformité à ce règlement*

Données sensibles

💡 Catégorie spéciale de données dont le traitement est en principe interdit, sauf exception légale stricte.

🔗 **Exemple concret** : Origine ethnique, opinions politiques, santé, données biométriques, orientation sexuelle, appartenance syndicale, convictions religieuses.

⚠️ **Pourquoi c'est important** : *Ces données bénéficient d'une protection renforcée. Avant tout traitement, vérifier qu'une exception légale s'applique (ex. consentement explicite).*

DPO — Délégué à la Protection des Données

💡 La personne chargée, au sein de l'établissement, de veiller au respect du RGPD et de conseiller l'organisation sur tout sujet lié à la protection des données.

🔗 **Exemple concret** : Le DPO de l'école est votre interlocuteur si vous avez un doute sur l'utilisation d'un outil numérique, sur la collecte de données étudiants, ou en cas de violation de données.

⚠️ **Pourquoi c'est important** : *Le DPO n'est pas un policier : c'est une ressource et un conseil. Sa désignation est obligatoire pour les organismes publics comme nous. Il est le point de contact avec la CNIL.*

Finalité

💡 La raison précise pour laquelle on collecte et utilise des données. On ne peut pas les réutiliser pour autre chose sans justification.

🔗 **Exemple concret** : Des données collectées pour la scolarité d'un étudiant ne peuvent pas être réutilisées à des fins de marketing.

⚠️ **Pourquoi c'est important** : *Ce principe interdit le détournement de données. Il faut définir le but avant de collecter, et s'y tenir.*

Pseudonymisation

💡 Technique qui remplace les informations identifiantes par un code ou un pseudonyme. La ré-identification reste possible avec une clé séparée.

🔗 **Exemple concret** : Remplacer le nom « Marie Dupont » par un identifiant « ETUD-4821 » dans une base de données de recherche.

⚠️ **Pourquoi c'est important** : *La pseudonymisation réduit les risques mais les données restent des données personnelles. Le RGPD continue de s'appliquer.*

Registre des activités de traitement

💡 Document officiel de l'établissement qui liste tous les traitements de données personnelles réalisés, leurs finalités, les catégories de données et personnes, les destinataires, les transferts de données, les durées de conservation, les mesures de sécurité.

🔗 **Exemple concret** : Gestion des dossiers étudiants, suivi des ressources humaines, utilisation d'outils numériques, vidéosurveillance...

⚠️ **Pourquoi c'est important** : *C'est la colonne vertébrale de la conformité RGPD. Il permet de prouver qu'on sait ce qu'on fait avec les données.*

Responsable de traitement

💡 L'entité (entreprise, école, administration) qui décide pourquoi et comment des données personnelles sont utilisées.

🔗 **Exemple concret** : L'école est responsable de traitement pour les données de ses étudiants et personnels — pas l'agent individuel qui tape dans l'outil.

⚠️ **Pourquoi c'est important** : *C'est l'établissement qui est juridiquement responsable devant la CNIL, pas l'employé. Cela ne dispense pas chacun d'agir avec soin.*

Sous-traitant

💡 Personne ou entité externe qui traite des données personnelles pour le compte de l'école, selon ses instructions, sans en déterminer les finalités.

🔗 **Exemple concret** : Un fournisseur de logiciel SaaS, un hébergeur cloud, une plateforme d'IA utilisée par l'établissement.

⚠️ **Pourquoi c'est important** : *La relation avec le sous-traitant doit être encadrée par un contrat spécifique (article 28 RGPD). L'école reste responsable.*

Traitement

💡 Toute action réalisée sur des données personnelles, par procédé automatisé ou non : les collecter, les stocker, les consulter, les partager, les effacer...

🔗 **Exemple concret** : Enregistrer des données dans un tableur, utiliser un outil d'IA avec des données d'élèves.

⚠️ **Pourquoi c'est important** : *Toute manipulation de données personnelles — même banale — déclenche l'application du RGPD.*

Transfert vers un pays tiers

💡 Envoi de données personnelles vers un pays hors Union Européenne (ou hors Espace Économique Européen), soumis à des garanties spécifiques.

🔗 **Exemple concret** : Utiliser un outil d'IA dont les serveurs sont aux États-Unis avec des données d'étudiants.

⚠️ **Pourquoi c'est important** : *Ce type de transfert est strictement encadré. Il faut des garanties juridiques spécifiques (ex. clauses contractuelles types).*

Violation de données

💡 Tout incident de sécurité qui entraîne la perte, la destruction, l'altération, la divulgation non autorisée ou l'accès non autorisé à des données personnelles.

🔗 **Exemple concret** : Un e-mail envoyé par erreur à la mauvaise personne, une clé USB perdue, un piratage informatique.

⚠️ **Pourquoi c'est important** : *Certaines violations doivent être signalées à la CNIL dans les 72 heures. Il faut des procédures prêtes à l'avance.*

Notions IA Act — Règlement européen sur l'intelligence artificielle

Classification des systèmes d'IA par niveau de risque

💡 L'IA Act classe les systèmes d'IA en 4 catégories selon les risques qu'ils présentent pour les personnes.

🔑 **Exemple concret :**

⊘ Risque inacceptable (interdit) : notation sociale, manipulation subliminale.

⚠️ Risque élevé (très réglementé) : recrutement, évaluation scolaire.

🗨️ Risque limité (transparence obligatoire) : chatbots.

☑️ Risque minimal : filtres anti-spam.

⚠️ **Pourquoi c'est important :** *Le niveau de risque détermine quelles obligations s'appliquent. Un outil d'IA utilisé dans l'évaluation d'étudiants est considéré à risque élevé.*

Déployeur d'IA

💡 La personne ou l'entité qui utilise un système d'IA dans un contexte professionnel — c'est-à-dire nous, lorsque nous utilisons un outil d'IA.

🔑 **Exemple concret :** L'école qui utilise un chatbot pour répondre aux questions des étudiants, ou un enseignant qui utilise un outil d'IA pour corriger des copies.

⚠️ **Pourquoi c'est important :** *Le déployeur a des obligations propres : utiliser l'outil conformément aux instructions, assurer une supervision humaine quand nécessaire, gérer les risques.*

Fournisseur d'IA

💡 L'entreprise ou l'entité qui développe, met sur le marché ou met en service, un système d'IA.

🔑 **Exemple concret :** OpenAI (ChatGPT), Google (Gemini), Anthropic (Claude), Mistral AI.

⚠️ **Pourquoi c'est important :** *Le fournisseur est responsable de la conception, de la documentation et de la conformité du système. C'est lui qui doit s'assurer que l'outil respecte l'IA Act.*

Obligation de littératie en IA — Article 4, IA Act

💡 L'IA Act impose aux organisations qui utilisent l'IA de s'assurer que leurs équipes comprennent suffisamment bien ces technologies pour les utiliser de façon responsable.

🔑 **Exemple concret :** Former le personnel aux bases de l'IA, à ses limites et à ses risques. C'est exactement l'objet de cette campagne de sensibilisation !

⚠️ **Pourquoi c'est important :** *C'est une obligation légale (article 4 de l'IA Act). L'école, en tant que déployeuse d'IA, doit pouvoir justifier que ses agents ont les connaissances nécessaires.*

Système d'IA (et GPAI — IA à usage général)

💡 Un système automatisé capable de produire des résultats (textes, images, décisions, recommandations) à partir de données qu'il a traitées, avec un certain niveau d'autonomie.

🔗 **Exemple concret** : GPT-4 (OpenAI), Gemini (Google), Claude (Anthropic), Mistral, Llama (Meta). Ces outils peuvent rédiger, résumer, traduire, coder, analyser...

⚠️ **Pourquoi c'est important** : *L'IA Act s'applique dès qu'on utilise ou déploie ce type de système. Le niveau d'obligation dépend du niveau de risque.*

🌐 Notions communes — RGPD + IA Act

Données d'entraînement et de test

💡 Les données utilisées pour apprendre à un système d'IA à fonctionner (entraînement) et pour vérifier qu'il fonctionne correctement (test).

🔗 **Exemple concret** : Des milliers de textes, d'images ou d'exemples annotés qui ont permis à un LLM (Large Language Model) d'apprendre à produire du langage cohérent.

⚠️ **Pourquoi c'est important** : *Si ces données contiennent des données personnelles, le RGPD s'applique. Si elles sont biaisées, le système risque de produire des résultats discriminatoires.*

Gouvernance des données

💡 L'ensemble des règles, rôles et processus qui définissent comment les données sont gérées dans l'organisation : qualité, accès, sécurité, partage.

🔗 **Exemple concret** : Qui a le droit d'accéder à quelles données ? Comment s'assure-t-on de leur qualité ? Qui est responsable en cas de problème ?

⚠️ **Pourquoi c'est important** : *Une bonne gouvernance des données est le socle commun pour se conformer au RGPD, à l'IA Act et aux exigences de cybersécurité.*

🌐 Concepts techniques IA — Pour comprendre comment fonctionnent les outils

Hallucination

💡 Quand un système d'IA produit une information fautive ou inventée, mais présentée avec assurance comme si elle était vraie.

🔗 **Exemple concret** : Un LLM qui cite un article scientifique inexistant, invente une loi qui n'existe pas, ou donne une date incorrecte avec l'air parfaitement sûr de lui.

⚠️ **Pourquoi c'est important** : *Les outils d'IA peuvent se tromper de façon convaincante. Il est essentiel de vérifier les informations importantes, surtout dans un cadre professionnel ou académique.*

IA agentique

💡 Un système d'IA capable d'agir de façon autonome pour accomplir des tâches complexes : il planifie, prend des décisions et exécute des actions chaînées sans intervention humaine à chaque étape.

🔗 **Exemple concret** : Un agent IA qui reçoit la consigne « organise la réunion de la semaine prochaine » et qui, seul, consulte les agendas, rédige les invitations et envoie les e-mails.

⚠️ **Pourquoi c'est important** : *Plus l'IA agit de façon autonome, plus la supervision humaine devient critique. L'IA Act encadre strictement les systèmes à haut niveau d'autonomie, en particulier dans des contextes sensibles.*

IA générative

💡 Catégorie de systèmes d'IA capables de créer du contenu nouveau : textes, images, sons, vidéos, code... à partir d'une instruction ou d'un exemple.

🔗 **Exemple concret** : ChatGPT génère du texte, DALL-E génère des images, Suno génère de la musique. Tous sont des IA génératives. Les LLM comme Claude ou Gemini en sont une forme centrée sur le langage.

⚠️ **Pourquoi c'est important** : *L'IA générative peut produire des contenus très convaincants, mais aussi faux ou biaisés. Elle soulève des questions de droits d'auteur, de confidentialité et d'authenticité.*

LLM — Grand modèle de langage

💡 Un type de système d'IA entraîné sur d'immenses quantités de textes pour comprendre et produire du langage humain.

🔗 **Exemple concret** : ChatGPT, Claude, Gemini, Mistral sont des LLM. Ils peuvent répondre à des questions, rédiger, résumer, traduire, expliquer...

⚠️ **Pourquoi c'est important** : *Les LLM sont une forme d'IA générative spécialisée dans le texte. Comprendre ce qu'ils sont (et ce qu'ils ne sont pas) aide à les utiliser avec discernement.*

RAG (Retrieval-Augmented Generation) — Génération augmentée par récupération

💡 Une architecture d'IA qui connecte un LLM à une base documentaire : avant de répondre, le système va d'abord chercher des informations dans des documents fournis.

🔗 **Exemple concret** : Un assistant d'IA qui peut répondre à des questions sur le règlement intérieur de l'école en ayant accès au document PDF correspondant.

⚠️ **Pourquoi c'est important** : *Le RAG permet de limiter les hallucinations et d'ancrer les réponses dans des sources connues. Il soulève aussi des questions de droits d'accès aux documents utilisés.*