Université de Rennes

CHARTE PORTANT SUR L'UTILISATION DES IAg

Table des matières

Préambule

L'Université de Rennes est engagée sur le sujet des transitions environnementale et numérique. A ce titre nous portons de nombreux projets en IA : CMA Tiare, IA cluster SequoIA, RAGaRenn AIR DemoES. Comme pour certains pans de la société et dans de nombreux établissements, l'irruption de l'IA générative (IAg) présente des opportunités et risques que notre établissement souhaite maitriser.

La présente charte vise à limiter les risques et saisir les opportunités, lorsqu'elles s'avèrent pertinentes, en encadrant les usages des outils d'IA générative. L'objectif est d'établir un environnement de travail où les solutions d'IAg sont utilisées de manière éthique, responsable, durable et bénéfique pour tous, dans le respect des principes généraux de responsabilité, de transparence, de précaution et parcimonie.

La responsabilité de l'utilisateur est centrale : nous encourageons chacun à exercer sa curiosité scientifique, par l'exploration, la formation continue et le partage d'expériences, tout en veillant à ce que chaque utilisateur endosse la responsabilité des contenus produits à l'aide d'outils d'IAg qu'il utilise.

La transparence dans le déploiement et l'usage de ces outils est indissociable de notre intégrité académique, scientifique et professionnelle, tout en promouvant notre démarche de reproductibilité scientifique.

Le principe de précaution et parcimonie assure que les risques sont anticipés et gérés de manière proactive : plus spécifiquement la parcimonie tend vers une sobriété numérique, utilisant les ressources de manière judicieuse et responsable, afin de maitriser l'impact environnemental et social des outils d'IAg.

La charte s'adresse à l'ensemble de nos utilisateurs : usagers (étudiants, stagiaires de la formation continue, etc.) et personnels (enseignants, enseignants-chercheurs, chercheurs et fonctions support et soutien). L'université étudie, valide, voire développe et déploie si nécessaire des outils à l'intention de ses utilisateurs, en assurant l'hébergement et la maintenance lorsqu'il est pertinent de le faire (tel que pour RAGaRenn lancé début 2024) : tout déploiement d'outil numérique nécessite étude et validation, y compris des licences ou demandes d'abonnement pour des outils tels que ChatGPT ou Copilot.

La charte s'inscrit dans la législation française applicable (RGPD, LREN, code de l'éducation) et le règlement 2024/1689 du Parlement Européen et du Conseil Européen du 13 juin 2024, qui est le premier acte législatif sur l'intelligence artificielle (IA). Il est paru au Journal officiel de l'Union européenne (JOUE) le 12 juillet 2024.

1. Qu’est-ce qu’une IAg*

Selon la CNIL, l'intelligence artificielle « générative » désigne la classe des systèmes capables de créer des contenus (texte, code informatique, images, musique, audio, vidéos, etc.). Ces systèmes sont qualifiables de systèmes d’IA à usage général lorsqu’ils permettent de réaliser des tâches de différents domaines d’application : c’est le cas notamment des systèmes qui reposent sur des modèles de langage (LLM).

Ils reposent sur le formalisme de "réseaux de neurones artificiels" qui permet d'apprendre à partir de (très) nombreux exemples les caractéristiques statistiques des contenus à générer. Les robots conversationnels en langage naturel, comme ChatGPT d'OpenAI ou Le Chat de Mistral, constituent un exemple typique d'IAg apprise pour générer du contenu en réponse à des instructions exprimées en langage naturel, la réponse pouvant être entre autres choses du langage naturel, du code informatique, des images ou encore des vidéos.

La conception de ces systèmes nécessite de vastes quantités de données provenant de différentes sources (internet, sources tierces sous licence, conversations générées par des formateurs humains, interactions avec les utilisateurs, données synthétiques, etc.) pour apprendre les caractéristiques statistiques des contenus à générer et, le cas échéant, des instructions permettant leur génération. Les IAg sont avant tout des modèles statistiques, certes forts complexes et impressionnants, mais dénués de toute capacité de “réflexion”. Par exemple, les modèles de langues sont entraînés pour prédire le mot qui suit le plus probablement un énoncé, permettant ainsi de générer un mot après l'autre, sans aucune compréhension de la demande ou du texte généré. Leur réponse peut également être améliorée en s'appuyant sur un ensemble de documents fournis par l'utilisateur et dans lequel le programme d'IAg recherche des éléments de réponse appuyant sa génération de texte : on parle alors de 'Retrieval Augmented Generation' (RAG).

Ainsi si l'on demande à une IA conversationnelle de 2024 la recette du far breton au lait de canard, nous obtiendrons une réponse... absurde ! Si l'on demande « De quelle couleur est le ciel ? », il est très probable que la réponse soit « bleu » car c'est ce qui est statistiquement le plus probable. Mais si on utilise un système de RAG dont la base documentaire précise que le ciel est jaune fluo, alors le système pourra répondre que le ciel est jaune fluo puisque cela devient plus probable en raison des affirmations de la base documentaire.

2. Consignes générales

L'utilisation des IAg doit respecter les législations, règlements et recommandations, et en particulier ceux liées à la propriété intellectuelle, la sécurité des systèmes d'information et la protection des données à caractère personnel. Les grands principes à respecter lors de l'utilisation et de la mise en place d'une solution d'IAg sont énoncés ci-dessous :

Principes génériques :

Sécurité des IAg

- S'assurer de la sécurité des IAg utilisées en demandant l'appui des services compétents.
- Prévoir un mode dégradé des services métier sans système d'IAg (principe de réversibilité).
- Sensibiliser les utilisateurs aux risques et opportunités liés à l'usage de l'IAg.

Vérifier les sources

- Privilégier l'usage de modèles d'IAg transparents et compréhensibles afin de comprendre comment un contenu est généré et de pouvoir corriger les éventuelles erreurs tout en garantissant l'originalité de la production.
- Utilisez des contenus dont la légitimité, ou les sources ont été vérifiées et s'avèrent fiables. L'IAg peut produire des éléments de réponse faux ou approximatifs (cf. ci-après point de vigilance sur les résultats produits). L'IAg peut également produire des extraits issus de contenus protégés, il est donc nécessaire de vérifier les droits applicables avant d'utiliser ou de partager les contenus.
- Préférer les IAg entraînées sur des données libres de droits ou ouvertes, afin de limiter les risques juridiques et de favoriser le recours à des pratiques plus éthiques.

Mentionner le recours à l'IAg

- Mentionner explicitement ce recours lorsqu'il s'agit d'un élément d'appréciation nécessaire dans des travaux et productions (bibliographie, notes, mentions générales du document, etc.).
  Documenter les outils, données, prompts et choix technologiques dans toute génération par IAg. Il s'agit d'indiquer clairement les modèles utilisés, les conditions de licence, et des prompts, en particulier dans les contextes pédagogiques ou scientifiques.
  Faire systématiquement vérifier par un humain les résultats produits par des outils d'IAg.
- Contrôler les productions de l'IAg : Faire produire du code lisible par des IAg.
- Contrôler les résultats produits : Faire vérifier tous les résultats produits par un outil d'IAg dès lors qu'elle concerne une tâche susceptible d'entrainer des effets d'évaluation ou de prise de décision concernant une personne (recrutement, évaluation, etc.).
  Cette vérification humaine doit pouvoir être prouvée.

Utiliser des outils d'IA respectueux de la vie privée et des droits des personnes

- Utiliser des IAg respectueuses de la vie privée des personnes.
- Dans la mesure du possible, éviter de traiter dans les outils d'IAg des données nominatives ou directement identifiantes (nom, prénom, photo, enregistrement vocal, description physique, adresse personnelle, numéro de téléphone, etc.) ou sensibles (numéro de sécurité sociale, RIB, numéro de carte bleue, donnée de santé, etc.).
- N'utiliser que les données strictement nécessaires à la réalisation d'une tâche lorsqu'elle concerne directement ou indirectement des données de personnes.

Utiliser des outils d'IAg respectueux des lois et règlements en vigueur

- Interdiction d'utiliser des IAg grand public pour traiter des données sensibles ou confidentielles pouvant entrainer une fuite de données.
- Vérifier que le logiciel qui va traiter ou héberger les données est souverain et respectueux des lois et règlements européens et français sur la protection des données (RGPD, Loi informatique et libertés, etc). Préférer l'usage d'outils souverains.

Avoir des usages d'IAg respectueux des lois et règlements

- Ne pas s'attribuer en tant qu'auteur un contenu généré par l'IAg, par exemple du texte paraphrasé par l'IAg à partir de textes écrits par d'autres, cette pratique pouvant porter atteinte aux droits d'auteur.
- Ne pas présumer que tout ce qui est produit par une IA est librement réutilisable. Les conditions d'usage varient selon les outils. Éviter d'utiliser du contenu généré par une IAg sans définir à qui il appartient afin d''éviter les conflits d'usage.
- Former les utilisateurs d'IAg aux droits d'auteur, aux licences et aux usages numériques liés à l'IAg.
- Vérifier ce que disent vos statut, contrat d'engagement ou la réglementation sur les créations d'IAg à l'Université, certains droits peuvent revenir à l'établissement, d'autres à l'auteur.
- Respecter les licences des logiciels, IAg et bases de données utilisées, qu'elles soient libres, open source ou propriétaires. Leurs conditions d'usage sont toujours documentées.

IA et impact environnemental

- Être vigilant sur l'impact environnemental et la consommation énergétique liée à l'usage de l'IAg.

Points de vigilance sur les résultats produits

Le fonctionnement d'une IAg nécessite d'être vigilant quant au contenu des productions qui peuvent contenir des erreurs étant donné leur nature :

- Prendre en compte le fait qu'une IA
  - N'est pas "neutre" puisqu'elle est conçue par un humain avec ses propres biais (*exemple : certains algorithmes de reconnaissance faciale entraînés sur des ensembles de données où les personnes de certaines origines ethniques étaient en nombre insuffisant*.).
  - Qu'elle peut commettre des erreurs (*erreurs dans la conception ou entrainement réalisé sur un jeu de données trop restreint ou sur de mauvais critères par exemple*).
  - Fonctionne à partir de données ou de matériel qui peut être de mauvaise qualité : des banques de données (BDD) dont le contenu serait partial, erroné, basé sur de fausses informations, incomplet, volé (propriété intellectuelle), illégal, etc.
- Prendre en compte que l'IAg malgré son nom ne dispose pas d'aptitudes à la "réflexion",
- Résultats qui peuvent contenir des Hallucinations (cf. définition ci-dessous),
- Résultats qui peuvent contenir des Biais (cf. définition ci-dessous).

3. Consignes liées aux besoins des services administratifs

Les consignes générales du point 2 de cette charte s'appliquent pleinement à cette partie.

Processus projet et choix d'outil : Les services administratifs demandeurs d'outils d'IA générative doivent solliciter le pôle PMO de la DSI pour engager le processus projet et se référer au point 6 de la présente charte afin de s'orienter vers des outils validés par la gouvernance de l'établissement.

Exigé / Proscrit

√	Faire vérifier les documents administratifs générés par une IAg (courrier, mail, CCTP, ...) par les services compétents (DAJI, DRH, PMO, DAFPA, etc.).	×	Il est interdit de renseigner des données nominatives, sensibles et/ou stratégiques (cf. définition ci-après) dans un outil d'IAg.
√	Réaliser une déclaration de traitement et des mentions d'information pour chaque usage d'un outil d'IAg pour une "tâche" en particulier en lien avec la DPO, à partir des ressources disponibles sur l'intranet.	×	Il est interdit d'utiliser des outils grand public en renseignant des données de l'établissement.
√	Faire vérifier tous les résultats produits par un outil d'IAg par un humain (étude des CV, recrutement, etc.) et pouvoir apporter la preuve de cette vérification humaine.	×	Il est interdit de créer un document (texte, image, vidéo...) à partir d'un document non libre de droit.
√	Créer ou utiliser des IA respectueuses de la vie privée des personnes.		Ne pas traiter via un outil d'IAg des tableurs contenant des données à caractère personnel identifiantes ou sensibles (salaire, nom, prénom, etc.).

4. Consignes liées aux besoins des services pédagogiques

Les consignes générales point 2 de cette charte s'appliquent pleinement à cette partie.

Processus projet et choix d'outil : Les services pédagogiques ou composantes de l'université demandeurs d'outils d'IA générative doivent solliciter le pôle PMO de la DSI pour engager le processus projet et se référer au point 6 de la présente charte afin de s'orienter vers des outils validés par la gouvernance de l'établissement.

Deux types d'usages sont à étudier/encadrer dans les usages liés à la pédagogie : les usages à destination des enseignants et ceux à destination de la formation des étudiants à l'usage des outils d'IA.

4.1 Cas de l'usage des IA génératives par les enseignants

L'usage de l'IA par les enseignants pour les aider dans leurs missions est soumis aux mêmes règles que les consignes générales et le processus et choix d'outil reste celui du processus projet de la DSI (pôle PMO) et plus particulièrement le cadre de la comitologie SIFOP.

Attention, l'usage d'outil d'IA dans les missions d'enseignement ne doit pas donner lieu à des décisions automatiques pouvant augmenter le risque de recours. Toutes les tâches permettant l'évaluation des étudiants ou leur admission automatisée par IA doit obligatoirement faire l'objet d'une vérification humaine qui doit le cas échéant être démontrable.

Exigé / Proscrit

√	Utiliser l'IAg de façon optionnelle, comme une assistance réversible sur les gestes pédagogiques d'un enseignant (aide à la production de contenus, correction de copies, formulation de retours personnalisés (feedbacks) aux étudiants, etc.).	×	Il est proscrit d'utiliser d'outil d'IAg en substitution d'un geste enseignant, sans supervision ni vérification, pour des activités telles que la production de contenus, la correction des copies, la formulation de retours personnalisés (feedbacks) aux étudiants.
√	En tant qu'enseignant, communiquer explicitement les consignes d'usages ou d'interdictions de l'IAg auprès des étudiants durant les activités pédagogiques même celles donnant lieu à évaluation.	×	Ne pas traiter via un outil d'IAg des tableurs contenant des données à caractère personnel identifiantes ou sensibles (notes, résultats d'examens, nom, prénom, etc.).
√	Expliquer aux étudiants qu'il est essentiel de savoir évaluer les résultats d'un outil d'IAg pour l'utiliser de manière appropriée et éclairée.	×	Ne pas utiliser sur une présentation une image générée par IA sans vérifier, au préalable, ce qu'impose la licence d'utilisation.
√	Réaliser une déclaration de traitement et des mentions d'information pour chaque usage d'un outil d'IAg pour une « tâche » en lien avec la DPO, à partir des ressources disponibles sur l'intranet.	×	
√	Faire vérifier tous les résultats produits par un outil d'IAg par un humain (correction de copies, candidatures, etc.) et pouvoir apporter la preuve de cette vérification humaine.	×	

4.2 Cas de l'usage des IA génératives dans le cadre de la formation des étudiants à ces technologies

Il est nécessaire d'apprendre aux étudiants à comprendre et utiliser les outils d'IAg à l'Université. Pour ce faire, l'Université est engagée dans divers projets pédagogiques et des règles sont mises en place afin d'intégrer ces outils dans les processus pédagogiques.

Exigé / Proscrit

√	En tant qu'étudiant, dans le cadre d'une évaluation, lorsque l'usage est autorisé, déclarer l'utilisation de l'IAg dans les passages concernés.	×	Il est proscrit de rendre une production d'IAg sans le dire à la place d'un travail demandé à l'étudiant.
√	Créer ou utiliser des IA respectueuses de la vie privée des personnes et des droits de propriété intellectuelle.	×	Ne pas traiter via un outil d'IAg des tableurs contenant des données à caractère personnel identifiantes ou sensibles (listes de personnes, gestion des associations, etc.).
		×	Il est proscrit d'utiliser des systèmes d'IA extérieurs à l'établissement en les alimentant avec des informations ou des documents issus des travaux des enseignants et de l'université de manière générale.

5. Consignes liées aux besoins de la recherche

Les consignes générales du point 2 de cette charte s'appliquent pleinement à cette partie.

Processus projet et choix d'outil : Les unités de recherche demandeuses d'outils d'IA générative doivent solliciter le pôle PMO de la DSI pour engager le processus projet et se référer au point 6 de la présente charte afin de s'orienter vers des outils validés par la gouvernance de l'établissement.

Ne pas mettre en production des outils d'IAg sans préalablement nettoyer les bases d'apprentissages dès lors que des données personnelles, confidentielles, sensibles ou soumises à droit d'auteur ont été utilisées.

Exigé / Proscrit

√	Préserver la confidentialité et la valeur stratégique de l'information dans l'utilisation des outils qui les manipulent.	×	Dans la création de contenu ou dans la rédaction de publications, il est proscrit de dissimuler l'utilisation d'IAg ou d'outils automatisés, cette pratique étant considérée comme un manquement à l'intégrité scientifique (Code de conduite européen pour l'intégrité scientifique, juin 2023).
√	Vérifier si le contenu généré provient de sources existantes afin d'ajouter les références appropriées si nécessaire : s'il s'agit véritablement d'une innovation, mentionner l'IAg utilisée en cas d'utilisation directe de tout ou partie du contenu généré.	×	Il est proscrit de réutiliser dans un outil d'IAg des données identifiantes issues d'un projet de recherche (listes participants, données de santé, etc.).
√	Vérifier le cadre légal et le point de vue des revues, des éditeurs ou des financeurs concernant l'utilisation d'IAg, même s'il semble y avoir un certain consensus sur l'aide à la rédaction de textes scientifiques. Attention le cadre législatif sur ces outils est en cours d'élaboration.	×	Ne pas réutiliser sans vérification des sources et de leur "propriété" des œuvres, articles, contenus scientifiques etc. issues d'une IAg sans accord préalable de l'auteur.
√	Créer ou utiliser des IA respectueuses de la vie privée des personnes.		

Cas spécifique des données de santé

Les consignes générales du point 2 de cette charte s'appliquent pleinement à cette partie. Le cadre de cette partie concerne la recherche mais également les services de santé de l'Université de Rennes. Elle s'applique également aux personnels rattachés au CHU uniquement dans leurs missions liées à l'enseignement supérieur et à la recherche à l'Université de Rennes.

Processus projet et choix d'outil : Les unités de recherche ou les services de santé de l'Université de Rennes (SSE, SMUT) demandeurs d'outils d'IA générative doivent solliciter le pôle PMO de la DSI pour engager le processus projet et se référer au point 6 de la présente charte afin de s'orienter vers des outils validés par la gouvernance de l'établissement.

Aucune donnée de santé non anonymisée ne doit être renseignée dans les outils d'IAg.

La réidentification des personnes *via* l'utilisation d'outils d'IAg est proscrite.

Cet encadrement des données dans les outils d'IAg vaut pour tous les périmètres d'utilisation y compris la recherche.

Si, dans un cas spécifique, des données nominatives de santé ou pseudonymisées devaient être utilisées, une demande préalable devra être adressée aux DPO et RSSI de l'Université de Rennes.

Exigé / Proscrit

√	Réaliser une déclaration de traitement et des mentions d'information pour chaque usage d'un outil d'IAg pour une "tâche" » en lien avec la DPO, à partir des ressources disponibles sur l'intranet.	×	Traiter des données de santé non anonymisées (au sens de la définition du document en l'espèce) dans un outil d'IAg.
		×	Permettre la réidentification des patients dont les données ont été anonymisées via un outil d'IAg.

6. Choix des outils d'IA générative

IA générative « open source »

L'Open Source Initiative a publié une définition dédiée à l'IA (Open Source AI https://opensource.org/ai). Sur le fond la définition proposée s'inscrit dans les grands principes du logiciel libre, bien que la communauté scientifique s'interroge sur la reproductibilité qui est difficile à garantir.

Quels outils d'IA générative choisir et pourquoi ?

Le principe de précaution s'applique ici : il est interdit d'utiliser un outil d'IAg non validé.

Exemple 1 : sont proscrites les solutions grand public, gratuites ou payantes, dont le modèle économique repose sur la réutilisation des données fournies par l'utilisateur, sauf autorisation préalable et expresse de l'établissement sur des cas d'usages identifiés.

Exemple 2 : l'établissement publie une liste des outils recommandés en fonction des usages envisagés.

7. Définitions

- **Anonymisation** : L'anonymisation est un traitement de données personnelles qui consiste à utiliser un ensemble de techniques de manière à rendre impossible, toute réidentification de la personne, par quelque moyen que ce soit. Contrairement à la pseudonymisation, l'anonymisation est donc une opération irréversible.

- **Biais** : Le fait qu'une IA reproduise des distorsions dans le traitement de l'information, notamment des biais sociaux, présents dans les données qui ont servi à entrainer cette IA. Exemple : classifier en bus scolaire toute image contenant un véhicule jaune pour une IA entrainée sur des données américaines où tous les bus scolaires sont jaunes.

- **Données à caractère personnel, ou Données personnelles :** Conformément à l'article 4.1 du RGPD, désigne toute information se rapportant à une personne physique identifiée ou identifiable (ci-après dénommée «personne concernée») ; est réputée être une «personne physique identifiable» une personne physique qui peut être identifiée, directement ou indirectement, notamment par référence à un identifiant, tel qu'un nom, un numéro d'identification, des données de localisation, un identifiant en ligne, ou à un ou plusieurs éléments spécifiques propres à son identité physique, physiologique, génétique, psychique, économique, culturelle ou sociale.

- **Données confiées par un tiers nécessitant d'assurer le respect du droit d'auteur :** S'entend comme toutes données confiées par un tiers quel qu'il soit qui ne sont pas libres de droit au sens juridique du terme.

- **Données/informations confidentielles :** Désignent toutes les données, connaissances, informations d'ordre général ou particulier, savoir-faire, résultats, méthodologies, documents et échanges, sous quelque forme que ce soit (écrite, orale, électronique ou autre), qui sont communiquées entre les Parties dans le cadre de la présente Convention. Les Informations Confidentielles concernent toutes les informations scientifiques et techniques (résultats de recherche en cours ou finalisés ; découvertes, inventions, innovations et savoir-faire, qu'ils soient brevetables ou non ; modèles, schémas, algorithmes, codes sources et logiciels ; protocoles expérimentaux, méthodologies et procédures d'essai ; rapports de recherche, études, notes techniques et analyses), mais également toutes les données stratégiques et administratives (programmes de recherche, orientations scientifiques ; informations relatives au financement et à la gestion de la recherche ; plans d'affaires, propositions de collaboration et stratégies de propriété intellectuelle ; documents internes de gouvernance et d'organisation académique).

  Elles comprennent également toutes les informations pédagogiques et académiques (supports de cours, documents d''évaluation et examens non publiés ; contenus des formations et des méthodologies appliquées) mais également toutes les données financières et contractuelles (accords de partenariat de recherche et conventions de financement ; conditions financières, budgets prévisionnels et rapports financiers internes ; tarifications, offres et conditions commerciales des partenaires impliqués).

- **Données sensibles :** S'entend de toutes les données définies dans l'article 9 du RGPD (données à caractère personnel qui révèle l'origine raciale ou ethnique, les opinions politiques, les convictions religieuses ou philosophiques ou l'appartenance syndicale, ainsi que le traitement des données génétiques, des données biométriques aux fins d'identifier une personne physique de manière unique, des données concernant la santé ou des données concernant la vie sexuelle ou l'orientation sexuelle d'une personne physique) ainsi que les données dont le traitement peut engendrer un risque pour les personnes (RIB ou informations bancaires, numéros de CB, numéro de sécurité sociale, etc.) données RH.

- **Données stratégiques :** S'entend de toutes les données considérées comme stratégiques pour l'établissement. Il peut s'agir de données issues de la recherche, plans précis des bâtiments dont zones ZRR, architecture du SI de l'Université, données financières, informations considérées comme confidentielles notamment par la DSI, la DAJI, la DPR ou la gouvernance.

- **Hallucination :** Capacité d'une IAg à générer des données qui ne correspondent à aucun fait avéré, aucune réalité physique. Exemple : générer des mains à 7 doigts dans les images, inventer des faits dans la génération de texte.

- **Intelligence Artificielle – IA** : Conformément à la définition posée par la Commission Européenne, "l'IA regroupe les approches d'apprentissage automatique ; les approches fondées sur la logique et les connaissances et les approches statistiques, l'estimation bayésienne, et les méthodes de recherche et d'optimisation".

- **IA générative :** Selon la CNIL, l'intelligence artificielle « générative » désigne la classe des systèmes capables de créer des contenus (texte, code informatique, images, musique, audio, vidéos, etc.). Ces systèmes sont qualifiables de systèmes d'IA à usage général lorsqu'ils permettent de réaliser tout un ensemble de tâches : c'est le cas notamment des systèmes qui reposent sur des Modèles de langage (LLM).
  La conception de ces systèmes nécessite de vastes quantités de données provenant de différentes sources (internet, sources tierces sous licence, conversations générées par des formateurs humains, interactions avec les utilisateurs, données synthétiques, etc.).

- **Modèle de langage :** Modèle statistique de la distribution d'unité linguistiques (par exemple : lettres, phonèmes, mots) dans une langue naturelle. Un modèle de langage peut par exemple prédire le mot suivant dans une séquence de mots. On parle de modèles de langage de grande taille ou « *Large Language Models* » *(LLM)* en anglais pour les modèles possédant un grand nombre de paramètres (généralement de l'ordre du milliard de poids ou plus) comme GPT-3, BLOOM, Megatron NLG, Llama ou encore PaLM.

- **Personne concernée :** Désigne toute personne physique faisant l'objet d'un traitement de ses Données personnelles.

- **Phase d'entrainement** : L'entraînement est le processus de l'apprentissage automatique pendant lequel le système d'intelligence artificielle construit un modèle à partir de données.

- **Phase de déploiement** : Le déploiement fait référence au processus d'installation, de configuration et de mise en service d'un logiciel ou d'une application dans un environnement opérationnel.

- **Pseudonymisation** : Traitement de données personnelles réalisé afin qu'on ne puisse plus attribuer les données à une personne physique identifiée sans information supplémentaire.
  En pratique, la pseudonymisation consiste à remplacer les données directement identifiantes (nom, prénoms, etc.) d'un jeu de données par des données indirectement identifiantes (alias, numéro séquentiel, etc.). La pseudonymisation permet ainsi de traiter les données d'individus sans pouvoir identifier ceux-ci de façon directe. Contrairement à l'anonymisation, la pseudonymisation est une opération réversible : il est possible de retrouver l'identité d'une personne si l'on dispose d'informations supplémentaires.
  En général elle est utilisée quand un jeu de données est fourni à un tiers (B) par un organisme (A). Cela permet à (A) de réidentifier les données traitées par (B), sans que (B) ne puisse identifier les personnes.

- **Responsable du traitement :** Désigne la personne physique ou morale, l'autorité publique, le service ou un autre organisme qui, seul ou conjointement avec d'autres, détermine les finalités et les moyens du traitement ; lorsque les finalités et les moyens de ce traitement sont déterminés par le droit de l'Union ou le droit d'un État membre, le responsable du traitement peut être désigné ou les critères spécifiques applicables à sa désignation peuvent être prévus par le droit de l'Union ou par le droit d'un État membre.

- **Services ou composantes de l'Université :** désignent les unités de formation et de recherche (UFR), les écoles et les instituts qui forment, avec les services administratifs, la base de l'organisation universitaire. Sont également inclues dans cette définition les unités de recherche qu'elles soient multi-tutelles ou non.

- **Sous-traitant :** Désigne la personne physique ou morale, l'autorité publique, le service ou un autre organisme qui traite des données à caractère personnel pour le compte du responsable du traitement.

- **Traitement :** Désigne toute opération ou tout ensemble d'opérations effectuées ou non à l'aide de procédés automatisés et appliquées à des données ou des ensembles de données à caractère personnel, telles que la collecte, l'enregistrement, l'organisation, la structuration, la conservation, l'adaptation ou la modification, l'extraction, la consultation, l'utilisation, la communication par transmission, la diffusion ou toute autre forme de mise à disposition, le rapprochement ou l'interconnexion, la limitation, l'effacement ou la destruction.

8. Références documentaires

CNIL :

- https://www.cnil.fr/fr/intelligence-artificielle/intelligence-artificielle-de-quoi-parle-t-on
- https://linc.cnil.fr/dossier-intelligence-artificielle
- https://www.cnil.fr/fr/ia-la-cnil-publie-ses-premieres-recommandations-sur-le-developpement-des-systemes-dintelligence
- https://www.cnil.fr/fr/intelligence-artificielle-la-cnil-poursuit-ses-travaux
- https://www.cnil.fr/fr/ia-et-rgpd-la-cnil-publie-ses-nouvelles-recommandations-pour-accompagner-une-innovation-responsable
- Lire la méthodologie de la CNIL: Déployer une IA générative et Utilisation d'un système d'IA générative

ANSSI :

- https://cyber.gouv.fr/publications/recommandations-de-securite-pour-un-systeme-dia-generative

Autre

- https://ai-infrastructure.org/understanding-types-of-ai-attacks/

- https://meta.stackoverflow.com/questions/421831/policy-generative-ai-e-g-chatgpt-is-banned
- https://www.amue.fr/publications/actualites/details/les-rdv-de-lamue-2024-les-replays-sont-en-ligne

IA ACT et législation :

- https://digital-strategy.ec.europa.eu/fr/policies/regulatory-framework-ai
- https://www.cnil.fr/fr/entree-en-vigueur-du-reglement-europeen-sur-lia-les-premieres-questions-reponses-de-la-cnil
- https://www.entreprises.gouv.fr/la-dge/actualites/le-reglement-europeen-sur-lintelligence-artificielle-publics-concernes-dates-cles

Expérimentation Université de Rennes:

https://personnel.univ-rennes.fr/luniversite-de-rennes-se-lance-dans-lexperimentation-doutils-dintelligence-artificielle-generative