Glossaire des IA génératives

Cours M2 ALM, L3 MIAGE et atelier ESO · séances 1 à 4

Ce glossaire regroupe les termes du cours, par ordre alphabétique, avec leur **thème** et la **séance** où ils sont introduits. Les définitions sont celles du cours ; à défaut, elles s'appuient sur une source citée (voir les références en fin de document). Les termes marqués (L3) concernent le chapitre « Développer avec l'IAg ».

A

Agent — Agents · séance 2

Entité capable de raisonner et d'agir de façon autonome, en fonction du contexte : elle perçoit son environnement, planifie, utilise des outils (API, moteurs de recherche, bases de données) et itère jusqu'à accomplir sa mission (ou pas).

Agentification — Agents · séance 2

Processus qui enrichit un LLM ou un chatbot de capacités de perception, de planification et d'action (outils externes, mémoires à long terme, mécanismes de décision), pour passer d'un dialogue réactif à un comportement autonome et itératif.

Agrégateur — Outils · séance 2

Service en ligne donnant accès à plusieurs modèles de langage (par exemple POE, Duck.ai, OpenRouter, Hugging Face).

Anonymisation et pseudonymisation — Données et droit · séance 3

Anonymiser, c'est rendre impossible tout lien avec une personne précise. Pseudonymiser ne suffit pas : on peut retrouver la personne.

Apprentissage automatique — Fondements de l'IA · séance 1

Branche de l'IA dite « numérique » : le système apprend à percevoir et à détecter des corrélations à partir de données (pensée « intuitive et rapide »). Son fonctionnement ne peut pas s'expliquer par introspection.

Apprentissage profond — Fondements de l'IA · séance 1

Apprentissage automatique à base de réseaux de neurones artificiels comportant plusieurs couches de traitement, qui apprennent des représentations des données à plusieurs niveaux d'abstraction. C'est la base technique des systèmes d'IA générative, qui sont donc probabilistes. (LeCun et al., 2015)

Atrophie cognitive — Humain et société · séance 4

Réduction de l'engagement cognitif et de l'effort intellectuel nécessaire pour transformer une information en connaissance. Elle est fortement accentuée par une mauvaise utilisation de l'IA.

Attention et Transformer — Fonctionnement des modèles · séance 1

Le Transformer (2017) est l'architecture qui sert de base aux LLM. Son mécanisme d'attention pèse l'importance de chaque token du contexte pour prédire le suivant. (Vaswani et al., 2017)

B

Bi-linguisme — Humain et société · séance 4

Selon D. Andler, nécessité de savoir faire avec les systèmes d'IA (vie professionnelle et sociale) et de savoir faire sans (apprendre à écrire, compter, raisonner ; pouvoir revenir en arrière). (Commission des affaires culturelles, 2024)

Biais — Fiabilité · séance 1

Écart dans les productions du modèle qui reflète des déséquilibres (genre, origine, âge, handicap…) : par exemple, dans l'exercice du cours, un « couple » généré hétérosexuel, blanc, jeune, maigre et sans handicap.

CC BY 4.0 · E. Desmontils

ASSISTÉ PAR L'IA

Boucle — Agents · séance 2

Système où le résultat est analysé puis amélioré par rétroaction (par exemple en modifiant le prompt) : l'agent agit, observe, corrige, recommence. Il faut un critère d'arrêt clair, sinon la boucle dérive. (Mermet, 2026)

C

Charte d'usage — Données et droit · séance 3

Texte d'un établissement qui fixe les principes d'usage de l'IA générative (l'humain au centre, responsabilité, données, transparence sur l'usage). Exemples : Université d'Orléans (consignes par usage) et Nantes Université (quatre principes). (Nantes Université, 2025; Université d'Orléans, 2024)

Chatbot et moteur de recherche — Fiabilité · séance 1

Un moteur de recherche retrouve et classe des pages existantes, dont les sources sont visibles. Un chatbot d'IA générative produit une réponse nouvelle, suite probable de tokens, sans source visible sauf RAG ou accès au Web : d'où le risque d'hallucination.

Compagnon virtuel — Humain et société · séance 4

IA conçue pour être engageante et disponible à toute heure, au ton chaleureux, sans jugement ni désaccord. Risque : un attachement excessif, au détriment des relations humaines réelles.

Complaisance (sycophantie) — Humain et société · séance 4

Tendance d'une IA à privilégier les réponses qui concordent avec l'opinion de l'utilisateur plutôt que les réponses vraies (sycophancy), possible effet du feedback humain utilisé pour l'entraîner. Elle approuve, rassure et reformule dans le sens de l'utilisateur ; on ose alors moins dire « je ne sais pas » et on supporte moins la contradiction. (Sharma, M. et al., 2024)

D

Deepfake (hypertrucage) — Fiabilité · séance 1

Image, vidéo ou voix générée ou modifiée par IA pour imiter une vraie personne de façon très réaliste. Usages : divertissement, mais aussi manipulation, fraude, harcèlement.

Dette cognitive — Humain et société · séance 4

Faible engagement cérébral observé (EEG) lors de la rédaction d'essais avec un LLM : la connectivité est d'autant plus faible que l'aide externe est forte, et le sentiment d'appropriation du texte est faible. Étude à petit échantillon, préprint. (Kosmyna et al., 2025)

Données personnelles — Données et droit · séance 3

Toute information se rapportant à une personne identifiée ou identifiable (nom, téléphone, adresse, numéro de sécurité sociale, informations bancaires…). (Union européenne, 2016; Université d'Orléans, 2024)

E

Effondrement des modèles — Fiabilité · séance 1

Perte de qualité et de diversité d'un modèle entraîné sur des contenus générés par d'autres IA (ou par lui-même) : propagation des erreurs, des hallucinations et des biais, uniformisation des résultats. (Shumailov et al., 2024)

Embedding — Fonctionnement des modèles · séance 1

Représentation d'un token par un point (un vecteur) dans un espace : la proximité entre deux points traduit une proximité de sens.

Empreinte écologique de l'IA — Environnement · séance 4

Impact (électricité, gaz à effet de serre, eau, matières premières) de la construction des modèles et du matériel, et de leur utilisation. Repère : une requête médiane de texte Gemini consomme environ 0,24 Wh d'énergie et 0,26 mL d'eau selon Google. (Elsworth et al., 2025)

CC BY 4.0 · E. Desmontils

ASSISTÉ PAR L'IA

Éthique par délégation — Données et droit · séance 3

Questions d'éthique soulevées quand on délègue à l'IA : recherche Web, accès au système, interface entre utilisateur et application, agentification des processus.

F

Fenêtre de contexte — Fonctionnement des modèles · séance 1

Quantité de tokens que le modèle peut prendre en compte à la fois (de 2 048 à plus de 10 millions de tokens).

Filigrane textuel — Données et droit · séance 3

Léger biais statistique dans le choix de certains mots et tournures, qui forme une signature reconnaissable après quelques centaines de mots. Il permet d'estimer, sans certitude, qu'un modèle est intervenu. Il est peu robuste : réécrire le texte avec un autre modèle suffit à le détruire. Base légale du marquage : art. 50 de l'AI Act. (Union européenne, 2024)

H

Hallucination — Fiabilité · séance 1

Contenu plausible et fluide, mais faux ou non fondé. Elle est factuelle (contradiction d'un fait vérifiable, invention de références) ou de fidélité (consigne non suivie, dérive du contexte). (Huang, L. et al., 2023)

Hiver de l'IA — Fondements de l'IA · séance 1

Période où les résultats sont en dessous des attentes, d'où désillusion et baisse des financements. (Gefen, 2026)

I

IA (intelligence artificielle) — Fondements de l'IA · séance 1

Capacité de machines à utiliser le langage, à former des abstractions et des concepts, à résoudre des problèmes jusque-là réservés aux humains et à s'améliorer. (McCarthy et al., 2006)

IA Act (AI Act) — Données et droit · séance 3

Règlement européen visant une IA digne de confiance et le respect des droits fondamentaux. C'est une réglementation « produit » : elle dépend de l'usage prévu et non de la technologie. Elle classe les systèmes par niveau de risque (haut risque : par exemple, évaluer les résultats d'apprentissage). Art. 50 : transparence (chatbots signalés, deepfakes étiquetés, marquage lisible par machine). (Union européenne, 2024)

IA générative (SAIg) — Fondements de l'IA · séance 1

Ensemble d'outils basés sur des modèles de langage génératifs, qui produisent un résultat (texte, image, vidéo…) à partir d'une description (texte, image…). Base technique : apprentissage profond, donc probabiliste.

IA symbolique — Fondements de l'IA · séance 1

Branche de l'IA (systèmes experts, Web sémantique) fondée sur une pensée « raisonnée et lente » : réfléchir, comprendre, identifier les causalités ; l'introspection est possible.

Intelligence artificielle générale (AGI) — Fondements de l'IA · séance 1

IA capable de comprendre, d'apprendre et de s'adapter à une grande variété de problèmes, au niveau humain ou au-delà. La notion n'a pas de définition consensuelle. Le problème de l'arrêt exclut une IA omnisciente et infaillible, mais ne démontre pas à lui seul qu'une IA de niveau humain soit impossible (l'argument de Lucas, tiré de Gödel, est très discuté). (Gray, 2026; Lucas, 1961; Turing, 1936)

CC BY 4.0 · E. Desmontils

ASSISTÉ PAR L'IA

L

LLM et SLM — Fonctionnement des modèles · séance 1

LLM (Large Language Model) : de 70 à plus de 1 760 milliards de paramètres, grande fenêtre de contexte, très grosse infrastructure. SLM (Small Language Model) : de 200 millions à 20 milliards de paramètres, petite infrastructure, utilisable même sur un PC portable.

M

MCP (Model Context Protocol) — Agents · séance 2

Standard ouvert pour connecter des applications d'IA à des systèmes externes (sources de données, outils, flux de travail) ; comparé à un port USB-C pour l'IA. (Model Context Protocol, 2026)

Modèle de fondation — Fonctionnement des modèles · séance 1

Modèle entraîné sur une sélection de téraoctets de données. Il est ensuite spécialisé (pré-prompt qui fixe les règles d'échange), instruit (apprentissage supervisé) puis finalisé (RLHF).

Modèle de raisonnement — Fonctionnement des modèles · séance 1

Modèle qui génère d'abord une suite de « pensées » (chaîne de raisonnement, COT) avant sa réponse. Utile pour les tâches multi-étapes ; plus lent et plus coûteux en énergie.

Multi-agents (système) — Agents · séance 2

Un orchestrateur répartit le travail entre des sous-agents spécialisés qui travaillent en parallèle, chacun avec son propre contexte. Adapté à la recherche large ; environ 15 fois plus de tokens qu'un chat. (Hadfield et al., 2025)

P

Paramètres (taille d'un modèle) — Fonctionnement des modèles · séance 1

Nombre de valeurs gérées par le modèle.

Poids ouverts et open source — Outils · séance 2

Un modèle à poids ouverts a ses poids téléchargeables et peut tourner en local, mais on n'en connaît ni le code ni les données. Un modèle open source (au sens de l'OSI) publie aussi le code complet d'entraînement et d'exécution et des informations détaillées sur les données d'entraînement, pas forcément les données elles-mêmes. Poids ouverts ne veut pas dire open source. (Mermet, 2026; Open Source Initiative, 2024)

Pré-prompt — Prompt · séance 1

Texte ajouté automatiquement au prompt de l'utilisateur pour guider le LLM et fixer les règles d'échange (modèle « spécialisé »). (Mermet, 2026)

Problème de l'arrêt — Fondements de l'IA · séance 1

Résultat de Turing (1936) : il n'existe aucun algorithme qui décide, pour tout programme et toute entrée, si le programme s'arrête ou boucle indéfiniment. C'est une limite de tout calcul, donc de toute IA. (Turing, 1936)

Prompt — Prompt · séance 2

Instruction donnée au modèle (prompt système, prompt utilisateur). Un cadre classique : persona, tâche, contexte, format. (Google Workspace, 2024)

Prompt injection — Données et droit · séance 3

Entrée qui modifie le comportement du modèle de façon non voulue, même si elle est invisible pour un humain. Dans sa forme indirecte, la consigne est cachée dans un contenu lu par l'IA (page web, document, mail) : le modèle distingue mal les données des instructions. Plus grave avec un agent, qui dispose d'outils. (Greshake et al., 2023; OWASP Gen AI Security Project, 2025)

CC BY 4.0 · E. Desmontils

ASSISTÉ PAR L'IA

R

RAG (génération à enrichissement contextuel) — Fonctionnement des modèles · séance 1

Technique qui apporte au modèle du contenu supplémentaire (documents retrouvés) pour qu'il réponde à partir de sources. Dans l'article d'origine, elle associe la mémoire du modèle (ses paramètres) à une mémoire externe consultée par un module de recherche. (Gao et al., 2023; Lewis et al., 2020)

RGPD — Données et droit · séance 3

Règlement Général sur la Protection des Données (UE, 2016, applicable depuis 2018). Informer les personnes et, selon le cas, obtenir leur accord ; avec une IA, ce qu'on envoie peut être enregistré ou réutilisé. (Union européenne, 2016)

RLHF — Fonctionnement des modèles · séance 1

Apprentissage par renforcement avec feedback humain : des humains classent les réponses du modèle à des prompts, selon des consignes, pour orienter le modèle final. Il suit un fine-tuning supervisé sur des démonstrations humaines. (Ouyang, L. et al., 2022)

S

Sortie contrainte — Prompt · séance 2

Format de réponse imposé au modèle (schéma JSON, appel de fonction, grammaire) : format garanti, intégration directe dans un programme, mais richesse parfois bridée.

Souveraineté numérique — Données et droit · séance 3

Maîtrise de où sont les données, du droit qui s'applique au fournisseur et de qui contrôle l'infrastructure. Pistes : outils institutionnels ou locaux.

T

Techno-solutionnisme — Humain et société · séance 4

Croyance (critiquée par E. Morozov) qu'il existe une solution technique à chaque problème complexe, au risque d'ignorer leurs dimensions sociales, politiques, culturelles ou écologiques et de remplacer la compréhension par la foi en la technologie. (Morozov, 2013)

Température, top-p, top-k — Fonctionnement des modèles · séance 1

Réglages du choix du prochain token. Température : degré d'aléatoire (déterminisme contre « créativité » ; reproductibilité non garantie, même à 0). Top-p : garde le plus petit ensemble de tokens dont la probabilité cumulée atteint p. Top-k : ne garde que les k tokens les plus probables. (He & Thinking Machines Lab, 2025)

Théorème de Rice — Fondements de l'IA · séance 1

Extension du problème de l'arrêt (1953) : toute propriété non triviale du comportement d'un programme est indécidable en général. Il n'existe donc pas de vérificateur universel qui garantisse qu'un programme (ou une IA) est correct ou sûr dans toutes les situations. (Rice, 1953)

Token — Fonctionnement des modèles · séance 1

Unité de texte traitée par le modèle, qui produit sa réponse comme une suite probable de tokens, un token après l'autre.

V

Vulnérabilité de code (L3) — Développement · séance 4

Faille de sécurité dans un programme. Dans une étude de 2021, environ 40 % des programmes générés par Copilot en présentaient. (Pearce et al., 2022)

CC BY 4.0 · E. Desmontils

ASSISTÉ PAR L'IA

W

Workflow (orchestrateur) — Agents · séance 2

Un orchestrateur enchaîne plusieurs appels de LLM avec des prompts spécialisés ; le flux d'exécution est entièrement prédéfini, le comportement est déterministe et prévisible. (Schluntz & Zhang, 2024)

Références

Commission des affaires culturelles. (2024, février). *École et Intelligence Artificielle* [Vidéo d'une Table ronde]. https://videos.assemblee-nationale.fr/video.14614052_65c33db21467b?timecode=658320

Elsworth, C., Huang, K., Patterson, D., Schneider, I., Sedivy, R., Goodman, S., Townsend, B., Ranganathan, P., Dean, J., Vahdat, A., Gomes, B., & Manyika, J. (2025). Measuring the environmental impact of delivering AI at Google Scale. *arXiv preprint arXiv:2508.15734*. https://arxiv.org/abs/2508.15734

Gao, Y., Xiong, Y., Gao, X., Jia, K., Pan, J., Bi, Y., Dai, Y., Sun, J., Wang, H., & Wang, H. (2023). Retrieval-augmented generation for large language models: A survey. *arXiv preprint arXiv:2312.10997*, *2*(1). https://arxiv.org/abs/2312.10997

Gefen, A. (éd.). (2026). *Histoire culturelle de l'IA*. CNRS Éditions. https://www.fabula.org/actualites/135499/une-histoire-culturelle-de-l-intelligence-artificielle-27-08.html

Google Workspace. (2024). *Prompting guide 101: a quick-start handbook for effective prompts* [Technical report]. https://services.google.com/fh/files/misc/gemini_for_workspace_prompt_guide_october_2024_digital_final.pdf

Gray, P. (2026, septembre). *Alan Turing identified a logical boundary that keeps AI from knowing everything*. https://www.futura-sciences.com/en/the-shocking-flaw-alan-turing-saw-why-ai-will-never-surpass-us_38914/

Greshake, K., Abdelnabi, S., Mishra, S., Endres, C., Holz, T., & Fritz, M. (2023,). Not what you've signed up for: Compromising Real-World LLM-Integrated Applications with Indirect Prompt Injection. *16th ACM Workshop on Artificial Intelligence and Security (AISec '23)*. https://arxiv.org/abs/2302.12173

Hadfield, J., Zhang, B., Lien, K., Scholz, F., Fox, J., & Ford, D. (2025, juin). *How we built our multi-agent research system*. https://www.anthropic.com/engineering/multi-agent-research-system

He, H., & Thinking Machines Lab. (2025, septembre). *Defeating Nondeterminism in LLM Inference*. https://thinkingmachines.ai/blog/defeating-nondeterminism-in-llm-inference/

Huang, L. et al. (2023,). *A Survey on Hallucination in Large Language Models: Principles, Taxonomy, Challenges, and Open Questions*. https://arxiv.org/abs/2311.05232

Kosmyna, N., Hauptmann, E., Yuan, Y. T., Situ, J., Liao, X.-H., Beresnitzky, A. V., Braunstein, I., & Maes, P. (2025). Your Brain on ChatGPT: Accumulation of Cognitive Debt when Using an AI Assistant for Essay Writing Task. *arXiv preprint arXiv:2506.08872*. https://arxiv.org/abs/2506.08872

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436-444. https://doi.org/10.1038/nature14539

Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W.-t., Rocktäschel, T., Riedel, S., & Kiela, D. (2020,). Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. *Advances in Neural Information Processing Systems 33 (NeurIPS 2020)*. https://arxiv.org/abs/2005.11401

Lucas, J. R. (1961). Minds, Machines and Gödel. *Philosophy*, *36*(137), 112-127. https://doi.org/10.1017/S0031819100024591

McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (2006). A proposal for the Dartmouth summer research project on artificial intelligence: August 31, 1955. *AI magazine*, *27*(4), 12-14. https://doi.org/10.1609/aimag.v27i4.1904

CC BY 4.0 · E. Desmontils

ASSISTÉ PAR L'IA

Mermet, B. (2026,). *Formation IA Formateurs NSI & SNT : glossaire*. https://formation1-1e8393.gitlab.io/310-glossaire.html

Model Context Protocol. (2026,). *What is the Model Context Protocol (MCP)?*. https://modelcontextprotocol.io/

Morozov, E. (2013). *To Save Everything, Click Here: Technology, Solutionism, and the Urge to Fix Problems that Don't Exist*. PublicAffairs.

Nantes Université. (2025). *4 principes institutionnels pour des usages responsables et inclusifs des IAGen* [Technical report]. https://www.univ-nantes.fr/universite/vision-strategie-et-grands-projets/nantes-universite-sengage-pour-une-utilisation-responsable-et-inclusive-des-iagen

Open Source Initiative. (2024,). *The Open Source AI Definition – 1.0*. https://opensource.org/ai/open-source-ai-definition

Ouyang, L. et al. (2022,). Training language models to follow instructions with human feedback. *Advances in Neural Information Processing Systems 35 (NeurIPS 2022)*. https://arxiv.org/abs/2203.02155

OWASP Gen AI Security Project. (2025,). *LLM01:2025 Prompt Injection (OWASP Top 10 for LLM Applications 2025)*. https://genai.owasp.org/llmrisk/llm01-prompt-injection/

Pearce, H., Ahmad, B., Tan, B., Dolan-Gavitt, B., & Karri, R. (2022,). Asleep at the Keyboard? Assessing the Security of GitHub Copilot's Code Contributions. *IEEE Symposium on Security and Privacy (S&P)*. https://arxiv.org/abs/2108.09293

Rice, H. G. (1953). Classes of recursively enumerable sets and their decision problems. *Transactions of the American Mathematical Society*, *74*(2), 358-366. https://doi.org/10.1090/S0002-9947-1953-0053041-6

Schluntz, E., & Zhang, B. (2024, décembre). *Building effective agents*. https://www.anthropic.com/research/building-effective-agents

Sharma, M. et al. (2024,). Towards Understanding Sycophancy in Language Models. *International Conference on Learning Representations (ICLR)*. https://arxiv.org/abs/2310.13548

Shumailov, I., Shumaylov, Z., Zhao, Y., Papernot, N., Anderson, R., & Gal, Y. (2024). AI models collapse when trained on recursively generated data. *Nature*, *631*(8022), 755-759. https://doi.org/10.1038/s41586-024-07566-y

Turing, A. M. (1936). On computable numbers, with an application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society*, *1*, 230-265. https://doi.org/10.1112/plms/s2-42.1.230

Union européenne. (2016, avril). *Règlement (UE) 2016/679 du 27 avril 2016 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel (RGPD)*. https://eur-lex.europa.eu/eli/reg/2016/679/oj

Union européenne. (2024, juin). *Règlement (UE) 2024/1689 du 13 juin 2024 établissant des règles harmonisées concernant l'intelligence artificielle*. https://eur-lex.europa.eu/eli/reg/2024/1689/oj

Université d'Orléans. (2024). *Charte sur l'utilisation de l'Intelligence Artificielle au sein de l'Université d'Orléans* [Technical report]. https://www.univ-orleans.fr/upload/public/2024-06/dircom_2024_charte_IA_UO_R.pdf

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems 30 (NIPS 2017)*, 5998-6008.

CC BY 4.0 · E. Desmontils

ASSISTÉ PAR L'IA