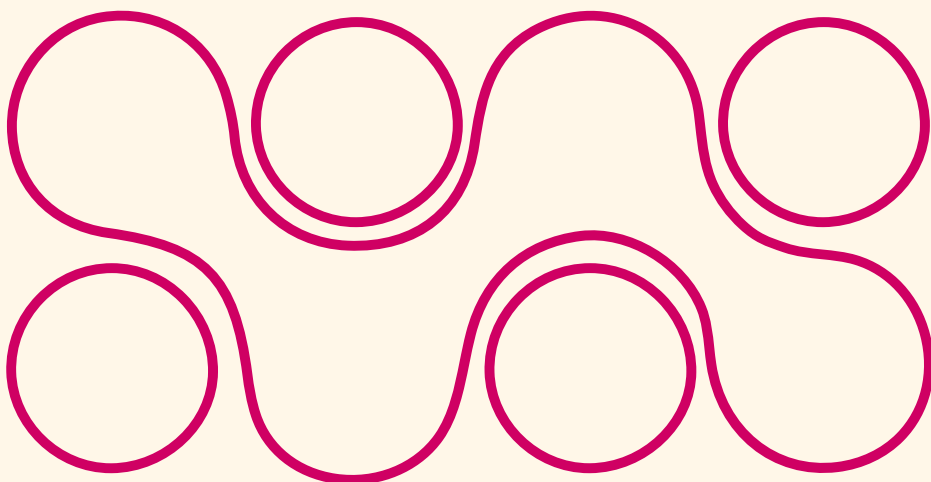
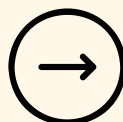


GUIDE À L'INTENTION DE LA COMMUNAUTÉ UNIVERSITAIRE

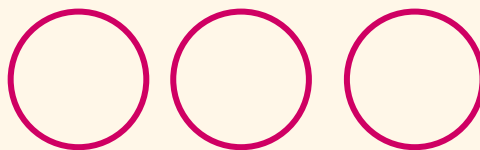
V1.0 - septembre 2024

INTELLIGENCE ARTIFICIELLE GÉNÉRATIVE



UNIVERSITÉ
DE GENÈVE

À propos



Le lancement de GPT-3 par OpenAI en novembre 2022 a été marquant pour la démocratisation de technologies d'intelligence artificielle (IA) générative auprès du grand public. Son adoption rapide – 100 millions de personnes en deux mois – a poussé les grandes entreprises technologiques à entrer dans une compétition effrénée, chacune cherchant à développer des modèles de langage toujours plus avancés et spécialisés.

Face à cette évolution et consciente de son impact, l'Université de Genève (UNIGE) a choisi d'être proactive et de s'adapter plutôt que de résister. Le 6 juillet 2023, le rectorat a publié une prise de position officielle et a établi un cadre évolutif pour l'utilisation de ces technologies. Ce cadre était accompagné de recommandations et de mesures de soutien. Rapidement, des ressources ont été mises à la disposition des enseignant-es et des étudiant-es. Cette prise de position a été actualisée dans une nouvelle version publiée le 1er juillet 2024.

L'IA générative a le potentiel de révolutionner les processus administratifs en optimisant les tâches répétitives et en redéfinissant la gestion des charges de travail. L'UNIGE a ainsi également reconnu le besoin d'accompagner le personnel administratif et technique (PAT) dans l'adoption des outils accessibles en ligne afin d'éviter un développement inégal des compétences et une possible réticence au changement.

S'appuyant sur des entretiens et enquêtes menés auprès du PAT, la rédaction de ce guide a progressivement pris une forme transversale, loin d'une catégorisation habituelle par métier, inadéquate pour de tels outils. Dès lors, il a été décidé de le publier à l'intention de toute la communauté.

La première partie du document présente des connaissances de base sur les technologies d'IA génératives. De plus, elle expose les risques inhérents et formule des recommandations. La seconde est un guide pratique orienté sur le recours aux outils disponibles en ligne de manière autonome. La troisième partie, résultat d'une enquête en septembre 2023, fait un état des lieux de l'utilisation de l'IA générative à l'UNIGE, ainsi que des craintes et attentes qui y sont associées.

Principes d'utilisation

Établis lors de la prise de position du 1er juillet 2024, les principes suivants s'appliquent à l'usage de l'IA générative dans la recherche, dans l'enseignement et l'apprentissage, ainsi que dans l'administration à l'Université de Genève:

Responsabilité

Les étudiants et étudiantes, collaborateurs et collaboratrices de l'UNIGE sont personnellement responsables de leur usage de l'IA générative.

Il est nécessaire d'avoir une compréhension de base des caractéristiques (i.e. modèle probabiliste) et des risques inhérents aux IA génératives (biais, hallucinations, données personnelles, etc.) avant d'en faire usage.

Les usagers et usagères doivent avoir la capacité de comprendre, de juger et d'évaluer la validité des réponses produites par l'IA générative pour ne pas en devenir dépendant-es.

La validation finale de chaque texte, image, vidéo ou autre travail ou décision générés ou assistés par un outil d'IA générative incombe à l'utilisateur ou l'utilisatrice.

Légalité

Les lois applicables doivent être respectées ainsi que, le cas échéant, les spécificités des clauses contractuelles (i.e. contrats d'édition).

Une attention particulière doit être portée à la protection des données, au droit d'auteur/à la propriété intellectuelle, au secret de fonction et à la confidentialité.

Un manquement aux principes de l'UNIGE ne peut pas être imputé à l'outil ou à une méconnaissance des risques et problèmes liés à l'IA. Des ressources, des guides et des formations sont recensées sur le portail IA de l'UNIGE pour accompagner la communauté dans la compréhension et la mise en application de ces principes.

Intégrité académique et transparence

Dans leurs publications et créations scientifiques, les chercheurs et chercheuses, ainsi que les étudiants et étudiantes, sont tenu-es de respecter les règles et principes gouvernant l'intégrité scientifique, notamment l'interdiction de plagiat, et de respecter les bonnes pratiques scientifiques ainsi que, le cas échéant,

Dans un contexte d'utilisation non-scientifique et publique, chaque texte, image, vidéo ou autre travail majoritairement généré ou modifié par une IA générative devrait être clairement identifié comme tel.

Les collaborateurs et collaboratrices sont tenu-es de signaler à leur hiérarchie l'utilisation d'outils IA générative dans le cadre de leurs fonctions.

Économie et écologie

L'IA n'est appliquée que dans les cas où elle est pertinente et sa valeur ajoutée est avérée.

Une attention particulière doit être portée à l'impact environnemental de l'utilisation d'outils d'IA génératives (e.g. consommation d'énergie).

Sommaire

1.

- 5 Introduction à l'IA générative
- 6 Connaissances de base
- 10 Une histoire de données
- 11 Risques
- 15 Bonnes pratiques

2.

- 17 Guides d'utilisation
- 18 Comment communiquer?
- 19 Structure d'une requête
- 22 Styles de travail
- 24 Choisir un usage

3.

- 29 Compte rendu
- 30 Adoption à l'UNIGE
- 31 Écosystème
- 34 Attentes
- 35 Craintes et freins

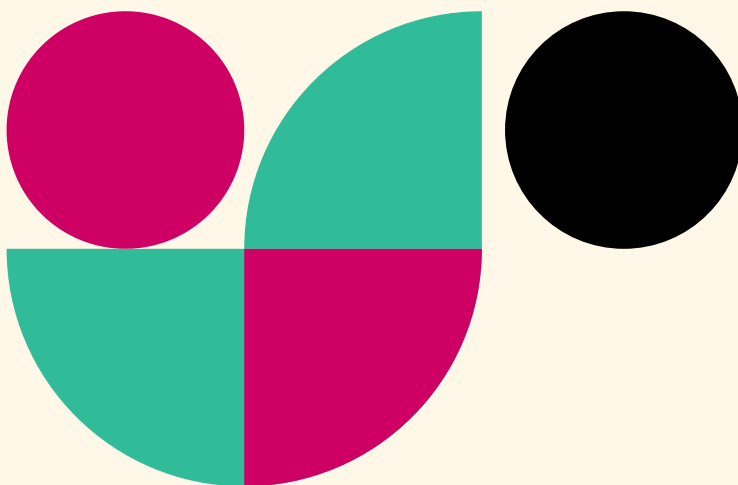
4.

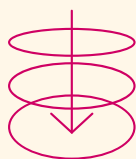
- 37 Conclusion
- 40 Remerciements
- 41 Références & ressources



PARTIE 1

INTRODUCTION À L'IA GÉNÉRATIVE





Connaissances de base

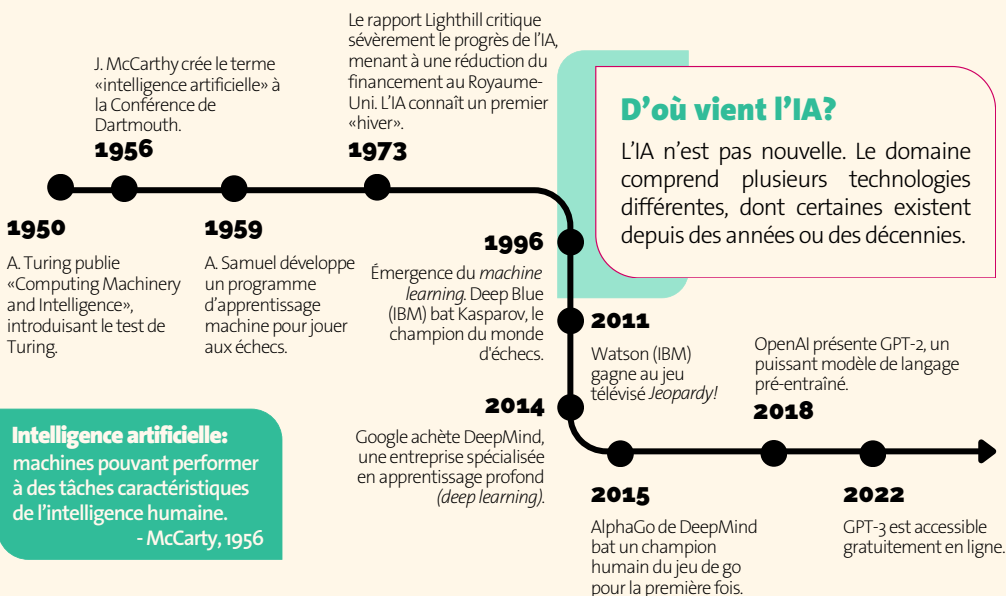
Parler d'intelligence artificielle (IA) induit le biais de se représenter l'intelligence selon notre perception humaine de ce qu'elle peut être. Ses contours sont mal définis, mais nous nous accordons généralement pour en avoir une idée commune qui se rapporte à notre capacité à effectuer certaines tâches avec un certain degré de compétence. L'analogie informatique par l'ajout du qualificatif «artificielle» ne permet pas de comprendre ce qu'est une IA, ni ce qu'elle peut faire, car elle ne partage pas la même cartographie de compétences.

Qu'est-ce que l'IA?

Cette locution désigne un domaine informatique. Les technologies qu'il englobe, souvent utilisées en combinaison, permettent aux ordinateurs d'exécuter un certain nombre de tâches spécialisées. La particularité principale, comparativement à un programme informatique classique, est l'autonomie de l'IA. Elle est conçue pour apprendre et s'améliorer avec le temps, elle peut faire un raisonnement, mais elle peut aussi percevoir des stimuli externes et interagir avec son environnement. Pour ces raisons, une IA est capable de réaliser un certain nombre de tâches complexes, habituellement propres à l'être humain, et généralement associées à notre conception de l'intelligence.

D'où vient l'IA?

L'IA n'est pas nouvelle. Le domaine comprend plusieurs technologies différentes, dont certaines existent depuis des années ou des décennies.



Que veut dire «généralive»?

La locution IA généralive se réfère à un ensemble de techniques (i.e. algorithmes) issu du domaine de l'intelligence artificielle. Ce que l'on entend par le terme de «généralive» est la capacité de l'outil à créer de l'information de manière automatique à partir de grands volumes de données existantes sur lesquelles il est entraîné. L'IA ne fait pas de simple copier-coller de ce qu'elle a analysé: elle imite, améliore et crée un résultat entièrement nouveau, basé sur une recombinaison statistique des motifs et structures qu'elle a identifiés pendant son apprentissage. Ces résultats, que l'on appelle **contenus**, peuvent être des textes, des images, de la musique ou du code informatique.

data set

Une grande quantité de données: des images, du texte, de la musique, des vidéos, etc.

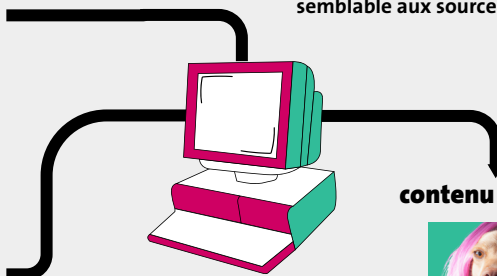


prompt

Dall-E: A photographic portrait, long shot, of a dog with long pink colored human hair, bright white studio lighting



Le processus est une création statistiquement probable et semblable aux sources



contenu généré



L'IA apprend les règles en analysant des milliers de textes, images ou morceaux de musique (**data set**). Elle étudie la manière dont ils sont arrangés et détecte des motifs récurrents qu'elle utilise ensuite pour générer de nouveaux contenus en fonction d'une requête (**prompt**).

Les **contenus générés** par l'IA sont des recombinaisons statistiques de tout ce qui a été appris. Ils ont du sens parce qu'ils respectent les règles grammaticales, visuelles ou musicales extraites des *data sets* d'entraînement. Ces règles sont mémorisées sous forme d'un *large Language Model* (LLM, voir p.9). Cela permet, entre autres, de synthétiser des informations ou de proposer des contenus originaux.

OpenAI

Fondée en 2015 en association à but non lucratif, OpenAI devient une entreprise à but lucratif en 2019 et établit un partenariat majeur avec Microsoft. Elle développe GPT-3, un modèle de langage, en 2020 et annonce DALL-E pour la génération d'images en 2021. En 2022, GPT-3 est accessible en ligne, marquant un tournant majeur. Grâce aux investissements de Microsoft, GPT-4 est publié en 2023.

L'outil ChatGPT repose sur GPT-4, un modèle puissant capable de traiter du texte et des images et de faire une recherche sur Internet en direct.

Microsoft Research

Créé en 1991 avec un focus sur les calculs d'ordinateur, la filiale s'investit dans l'IA dans les années 2010 avec divers rachats et possède maintenant 20% des brevets. Les innovations sont régulièrement ajoutées à ses produits (i.e. Bing, Holo Lens, Cortana, 365, Azure, etc.). En 2016, la division Microsoft Research AI est créée et LinkedIn est racheté pour alimenter de données ses modèles. Le partenariat en 2019 avec Open AI (i.e. ressources cloud) a été crucial pour le développement des modèles GPT.

En 2023, Microsoft publie Copilot, un assistant IA qui s'intègre à tous ses services.

DeepMind (Google)

Créée en 2010, DeepMind est acquise en 2014 par Google pour devenir sa filiale IA. Ses modèles s'inspirent des neurosciences pour développer des algorithmes d'apprentissage. DeepMind se concentre sur les systèmes capables de jouer à des jeux. AlphaGo a été le premier à battre le champion du monde de go en 2017. À la suite de GPT-3, Google panique et publie rapidement son prototype d'agent conversationnel.

L'outil Bard est vite publié avec le modèle LaMDA, rapidement changé pour PaLM2, avant d'être remplacé par Gemini.



IBM

Pionnière, notamment avec DeepBlue qui bat le champion d'échecs Kasparov en 1996. En 2011, le programme Watson fait sensation en remportant le jeu télévisé *Jeopardy!* ce qui nécessite de comprendre l'énoncé, d'utiliser le buzzer et de répondre vocalement. Après un élargissement à Internet, l'usage est restreint, le constat ayant été fait qu'il lisait Wikipedia et utilisait régulièrement le mot *bullshit* dans ses réponses.

Watson est monétisé en 2012 pour la santé, la finance et la recherche. En 2022, considéré comme un échec, le programme est vendu à un fonds d'investissement.

Meta AI

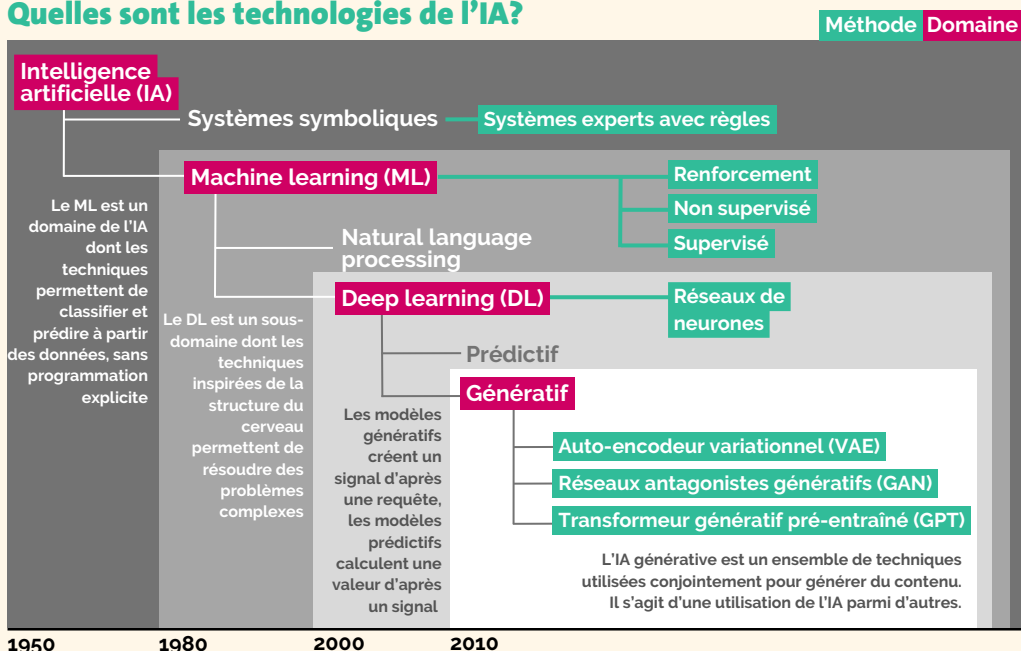
Créé en 2013, Facebook Artificial Intelligence Research (FAIR) publie PyTorch en 2017, un catalogue open source de modèles de *machine learning* (i.e. utilisés par Tesla ou Uber). FAIR est rebaptisé Meta AI après la réorganisation de Facebook et, en 2023, publie son modèle de langage Llama. Le modèle est initialement accessible sur demande aux chercheurs et chercheuses pour un usage non commercial avant d'être fuité un mois plus tard.

Llama 2 est une famille de modèles de langage open source. Meta n'a pas d'outil accessible publiquement.

Qui est derrière?

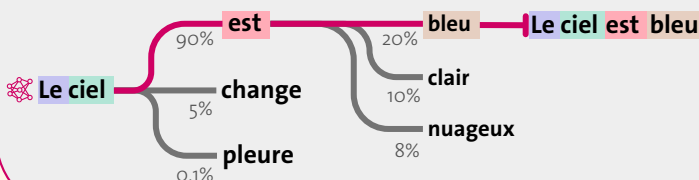
Plusieurs entreprises sont actuellement des actrices majeures dans le domaine de l'IA. Leurs outils, accessibles en ligne, souvent gratuitement, sont actuellement les plus utilisés.

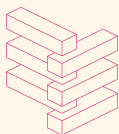
Quelles sont les technologies de l'IA?



Qu'est-ce qu'un «large language model» (LLM)?

Il est large, car il possède énormément de paramètres (de l'ordre de milliards), qui sont des pièces d'information. C'est un modèle, car il s'agit d'un réseau de neurones entraîné sur une grande quantité de textes pour produire des tâches non spécifiques. Il est linguistique, car il reproduit la syntaxe et la sémantique du langage naturel humain en prédisant la suite probable pour une certaine entrée. Cela lui permet également d'avoir une «connaissance» générale sur la base des textes d'entraînement.





Une histoire de données

Les outils d'IA génératives sont alimentés par d'immenses quantités de données pour entraîner les algorithmes à l'origine des modèles de fonctionnement. Ces données peuvent avoir une origine variée :



Elles sont massivement collectées dans ce qui est disponible en ligne.



Les requêtes et leurs réponses constituent aussi des données qui vont servir à alimenter le modèle et à l'améliorer.

À la suite des requêtes formulées par les utilisateurs/trices, ces outils vont générer davantage de données à partir de cet ensemble initial. Dès lors, la donnée apparaît comme l'élément crucial, la matière première dans le fonctionnement et l'usage des IA génératives. La qualité de la donnée sera un élément important de sa valeur.

Si la donnée est un objet possédant une grande valeur pour les outils, elle comporte également des risques d'utilisation.

La donnée produite est **faillible** avec des conséquences pour la vérification de l'information et sa transparence.

La donnée insérée possède une **sensibilité** dont découlent des enjeux de sécurité, de protection de la vie privée, de propriété intellectuelle, etc.

L'IA générative nous donne l'illusion du contrôle. L'outil reste maître d'énormément de paramètres que nous ne comprenons pas ou ne maîtrisons pas et il est délicat d'institutionnaliser son usage. Une étude récente (Dell'Acqua et al., 2023) évoque l'idée d'une frontière «dentelée» ou «irrégulière» pour se référer aux limites imprévisibles de l'IA générative. Non linéaire, incohérente, cette limite rend l'outil ambivalent. Il peut soit améliorer, soit entraver la performance d'une tâche. Par exemple, il produit des textes complexes, comme des poèmes, mais peine à donner une liste de mots commençant par une même lettre.

L'erreur serait d'anthropomorphiser l'IA générative comme un assistant dont les erreurs et bizarreries seraient dues à un manque d'intelligence ou un atypisme.

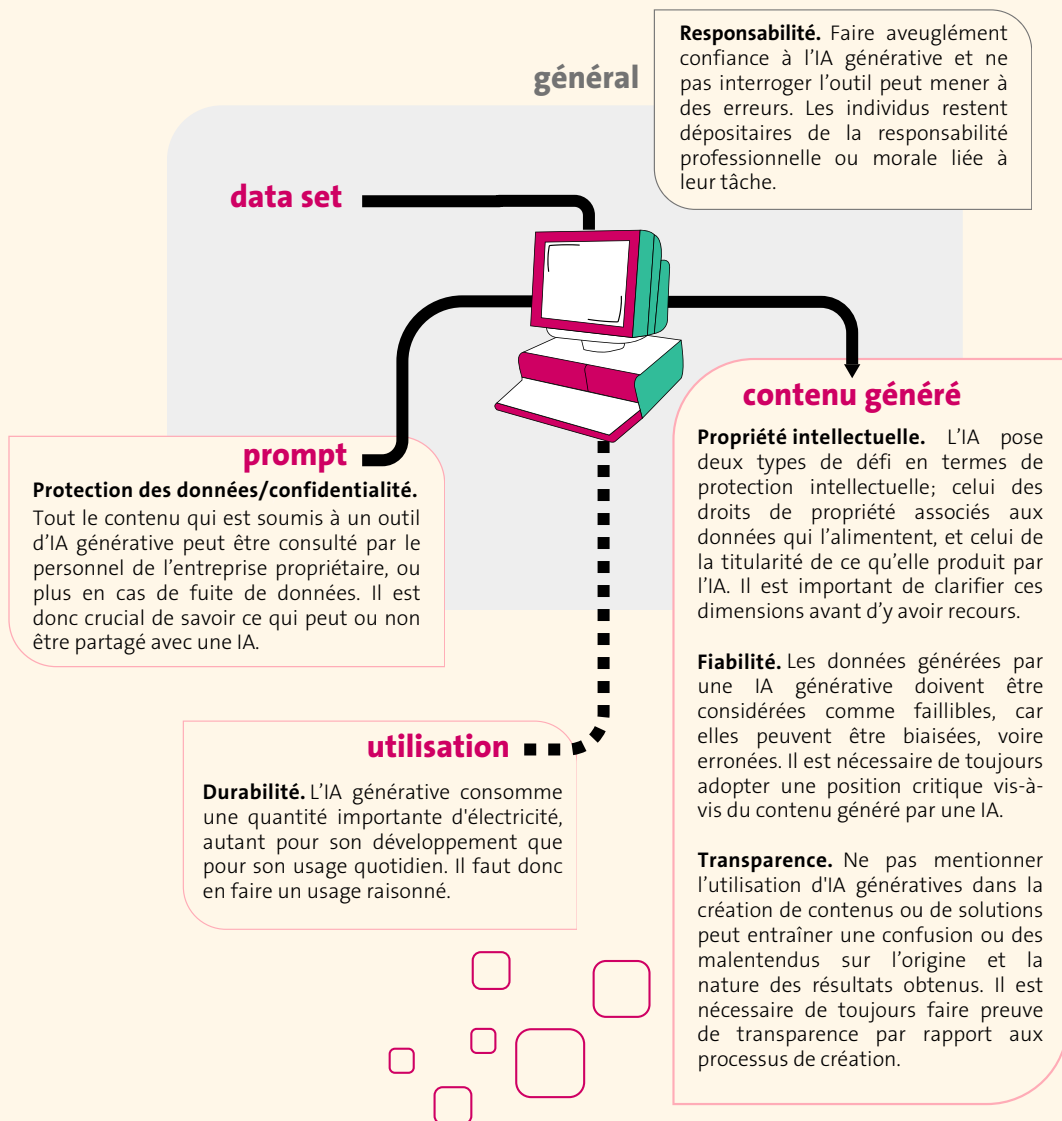
Il s'agit d'un objet entièrement nouveau et nous ne pouvons pas, contrairement aux individus, extrapoler sur le niveau de compétence général à partir de quelques tâches, car il ne repose pas sur un ensemble de ressources cognitives corrélées ou transversales. Nous n'avons pas le recul pour en définir les contours.

Un enjeu particulier pour l'UNIGE a trait au droit des personnes, notamment en matière de données personnelles, qui n'est pas traité par le règlement de l'Université ou par une directive et ne peut l'être en raison de sa complexité. La décision de fournir néanmoins des recommandations s'appuie sur le risque dommageable pour l'institution et les individus de ne pas accompagner l'usage d'IA génératives dont les outils, approuvés ou non, sont tout de même accessibles.

Risques

CARTOGRAPHIE

L'IA est associée à une série de risques pour les usagers/ères qui impliquent de prendre en compte un certain nombre de précautions et de recommandations, pouvant être schématiquement rattachées à différentes étapes de l'utilisation d'une IA générative.



EXPLICATION

Cette section présente un approfondissement de la notion de risques pour comprendre les enjeux associés et les recommandations qui en découlent (p. 13).

PROMPT

Protection des données/confidentialité

Et si c'était moi?

Je travaille sur le budget de mon service. Pour gagner du temps, je copie des données financières de l'institution ou des données sur les employé-es dans une requête, en retirant toute mention de l'Université. Plus tard, l'outil subit une cyberattaque, toutes les données sont publiées sur Internet. Comme j'ai utilisé une adresse UNIGE pour ouvrir mon compte IA, et donc pour faire les requêtes, n'importe qui peut déduire que les données financières sont celles de l'UNIGE, exposant ainsi l'institution.

D'un point de vue technique, tout le contenu soumis à une IA générative est accessible et consultable par l'entreprise propriétaire de l'outil. Il est peu probable que celle-ci s'intéresse au contenu individuel des requêtes et la plupart d'entre elles s'engagent à ne pas le faire. Les utilisateurs/trices doivent toutefois envisager que toutes leurs données ou requêtes deviendront des informations publiques. D'une part, il existe un risque de fuite de données dans le cas d'une faille de sécurité. D'autre part, il n'existe aucune assurance vérifiable sur la gouvernance des données par ces entreprises, par exemple concernant le stockage et l'utilisation des requêtes pour développer le modèle.

À ce titre, il est important de comprendre que toutes les données n'ont pas la même sensibilité, et que celle-ci dépend de leur nature.



Les données peuvent être classées selon quatre catégories principales:

Anonymes: elles ne posent pas de problème car elle sont indépendantes, soit elles sont intrinsèquement dissociées d'un individu en particulier ou ce lien a été rompu lors d'une anonymisation.

Ordinaires: ce sont les données personnelles classiques, souvent demandées lorsque l'on remplit un formulaire ou lorsque l'on traite le dossier d'un ou d'une employée ou d'une ou d'un étudiant (e.g. âge, sexe, nom, etc.). Elles permettent d'identifier une personne sans discrimination possible.

Sensibles: ce sont des données personnelles qui portent une information qui, si elle était connue, pourrait être utilisée au détriment de la personne (e.g. religion, congé maladie, assurances, etc.).

Secrètes ou confidentielles: ce sont des données possédant un enjeu stratégique. Plus rarement associées à des individus, elles concernent généralement des informations politiques, militaires, industrielles ou financières.

Chacune de ces catégories de données exige un niveau de protection approprié pour éviter les risques potentiels.

CONTENU GÉNÉRÉ

Transparence

Et si c'était moi?

Je suis chargée de recrutement et veux rédiger des lettres de refus personnalisées pour humaniser le processus, en offrant des retours plus détaillés et spécifiques. Sans le mentionner, j'utilise un outil d'IA. Un-e candidat-e reconnaît la structure et le style typiques d'une IA générative et soulève des questions sur un forum public. Mon intention initiale se noie dans une polémique engageant la transparence de l'Université dans ses communications.

Le devoir de transparence lors de l'utilisation d'intelligences artificielles génératives est une responsabilité incontournable dans les pratiques professionnelles. Cette transparence suppose une communication claire sur la manière dont ces technologies fonctionnent, leurs capacités ainsi que leurs limitations. Elle implique aussi de signaler l'utilisation d'IA génératives dans la création de contenus ou de solutions pour éviter toute confusion sur l'origine et la nature des résultats obtenus.

Fiabilité

Et si c'était moi?

J'utilise ChatGPT pour répondre aux questions fréquentes des étudiant-es, gérer les courriels et même rédiger des documents administratifs. Au bout de quelques semaines, il apparaît que ChatGPT produit parfois des réponses inexactes ou inappropriées. Il fournit des informations erronées sur les dates limites de dépôt de certains dossiers académiques ou rédige des réponses confuses qui donnent lieu à de nombreux recours.

Les données générées par une IA générative doivent être considérées comme faillibles (biaisées ou fausses). Cela s'explique par plusieurs facteurs.

Limite dans le temps. La plupart des modèles d'IA génératives ont été entraînés sur un ensemble de données correspondant à une période particulière sans connaissance des faits qui se situent avant ou après.

Hallucination. Les modèles d'IA utilisent des algorithmes pour détecter des motifs, combiner des éléments et produire un contenu susceptible de correspondre à la requête. Or, il existe des motifs imperceptibles ou dénués de sens pour l'humain. Incidemment, cela permet la création de contenus originaux et créatifs mais, parce que l'IA n'est pas critique, cela peut induire l'apposition d'informations individuellement correctes qui, mises ensemble, sont fausses. Au gré des mises à jour, cela devient plus rare, mais le risque reste présent.

Biais algorithmique. La qualité des contenus générés dépend de la qualité des données à partir desquelles le modèle a été entraîné. Si les données sont biaisées, les résultats le refléteront. Des biais dans la réponse formulée par l'IA peuvent également émerger en fonction de qui pose la question. Une utilisation irréfléchie peut entraîner involontairement le renforcement des stéréotypes ou des discriminations sociales.

Mode collapse. Plutôt que de générer des résultats variés, le modèle commence à produire de manière répétée un nombre très restreint de résultats similaires. Soit il n'a pas su capturer la richesse des données du modèle, soit il en reflète le manque de diversité. Cela crée un point aveugle dans l'interprétation des résultats; quand quelque chose manque, on s'en rend compte moins facilement que si quelque chose est faux.

Propriété intellectuelle

Et si c'était moi?

Je dois réaliser des illustrations pour la communication d'un événement et fournis les idées et les directives à l'IA qui conçoit les images. Je considère ces créations comme le fruit de mon travail créatif et de ma direction artistique. Après mon événement, je découvre les mêmes images réutilisées dans une campagne publicitaire. Je contacte l'entreprise qui a développé l'IA. Conformément aux termes d'utilisation, elle détient les droits de propriété intellectuelle sur les images générées.

L'IA se base sur des données, notamment des œuvres et des textes qui ne sont pas nécessairement libres de droits, pour générer de nouveaux contenus. Cela peut concerner les données soumises lors d'une requête, mais également le modèle de base. Cette capacité à reproduire ou à transformer des œuvres protégées sans le consentement explicite des détenteurs/trices de droits soulève des interrogations sur les violations potentielles du droit d'auteur-e. De plus, l'utilisation de l'IA pose la question de la titularité de ce qu'elle produit. Cette titularité est normalement assurée lors d'un usage payant.

UTILISATION

Durabilité

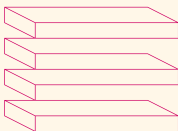
Et si c'était moi?

Lors des réunions d'équipe, les idées et les discussions sont représentées de manière simple et efficace sur un tableau blanc. Séduit-e par l'IA, je remplace ce système et génère des images d'illustration pour chaque nouvelle idée. Après plusieurs essais, j'obtiens une image complexe, censée représenter visuellement cette idée, pour rendre les discussions plus dynamiques et visuelles. Mais cette méthode s'avère rapidement plus distrayante qu'utile et aura eu un impact environnemental non négligeable.

L'IA générative est une grosse consommatrice d'électricité, tant pour son développement que pour son usage quotidien. Chaque requête ChatGPT équivaut à la consommation énergétique d'une lampe LED de 5W pendant une heure. La création de GPT-3 a généré plus de 500 tonnes de CO₂, soit l'équivalent de 123 véhicules à essence pendant un an, et 1287 MWh d'électricité, soit la consommation mensuelle de 1467 foyers américains. GPT-4 a été créé avec 500 fois plus de paramètres que GPT-3. Le développement de l'IA et son utilisation généralisée ne sont pas sans conséquences en termes de durabilité. Il faut prendre en compte la consommation énergétique de l'IA générative et limiter son utilisation à des usages spécifiques à forte valeur ajoutée.

RESPONSABILITÉ

Toutes les recommandations énoncées dans ces pages peuvent se résumer à une notion: les IA génératives ne sont que des outils informatiques puissants à utiliser pour assister les individus dans leurs tâches. Les individus restent dépositaires de la responsabilité professionnelle ou morale associée à leur tâche; elle ne peut être déléguée à l'IA générative qui ne peut être considérée comme autre chose qu'un outil.



BONNES PRATIQUES

Protection des données et confidentialité

Toutes les précautions nécessaires possibles doivent être prises pour assurer le respect de la confidentialité.

- ✦ Limiter l'usage à des données accessibles publiquement ou, le cas échéant, anonymes ou anonymisées. Dans certains cas définis et validés par la hiérarchie, les données ordinaires peuvent être utilisées en silo. L'usage d'autres types de données présente un risque majeur.
- ✦ Vérifier les paramétrages des outils en ligne : désactiver ou supprimer son historique, bloquer l'accès au reste de l'appareil, refuser toute réutilisation des données, etc.
- ✦ Ne pas permettre une identification personnelle qui pourrait, par déduction, mener à contextualiser une donnée. Cela inclut de ne pas utiliser son adresse courriel professionnelle, de ne pas décrire précisément sa profession, son rôle ou sa situation géographique.
- ✦ Mieux vaut ne pas employer un outil d'IA générative s'il y a le moindre doute sur la nature des données qui sont utilisées ou sur ce qu'il adviendra des données insérées.

Transparence

L'usage des IA génératives ne devrait pas être fait à l'insu de celles et ceux qui sont exposé-es au contenu généré.

- ✦ Toujours déclarer à sa hiérarchie ou son enseignant-e l'usage de l'IA générative dans son travail.
- ✦ Dans un document public, généré par l'IA générative, possédant une bibliographie, citer l'outil selon les règles en vigueur. S'il s'agit d'un communiqué, une image, ou autre document plus concis, la mention "généré par ou avec l'aide de l'IA générative" suffit.
- ✦ L'usage de l'IA générative dans un contexte scientifique pose des problèmes de plagiat. Son usage doit toujours faire l'objet d'un accord explicite avec le responsable de recherche et doit pouvoir être décrit méthodologiquement. La rédaction des travaux ne doit pas être déléguée à l'outil.

Fiabilité et durabilité

L'IA générative est un outil adapté à certains usages et non à d'autres.

- ✦ La performance des outils d'IA générative dépend fortement des cas d'utilisation. Il est donc important de se questionner sur la pertinence de l'usage choisi (cf. partie 3). Un usage raisonné permet également de limiter l'impact environnemental.
- ✦ Quand l'objectif est de générer un contenu original (cf. Création et Recherche), toujours vérifier la véracité des faits évoqués ou, au mieux, demander des sources à l'outil.
- ✦ Rester vigilant-e au fait que les contenus produits sont sujets à la désinformation (i.e. Hallucination), à véhiculer des discriminations (i.e. Biais algorithmique) ou à l'omission (i.e. Mode collapse).

Responsabilité

Les individus sont responsables personnellement du contenu produit avec l'IA et de l'usage qui en est fait au même titre que des auteurs/trices.

- ✦ Après avoir utilisé un outil d'IA générative, relire le texte ou revoir le media puis valider manuellement le contenu généré afin de s'assurer qu'il est conforme à quelque chose que l'on aurait pu créer soi-même et dont on est prêt à endosser la responsabilité.
- ✦ L'usage des IA génératives doit être éthique et ne pas causer de tort à autrui.
- ✦ Être à jour concernant la politique et les directives de l'institution à jour et être au courant des lois et des clauses contractuelles en vigueur concernant l'IA générative.
- ✦ Toute faille ou erreur entraînant des risques pour l'institution ou les individus doit être signalée dès qu'elle est détectée.

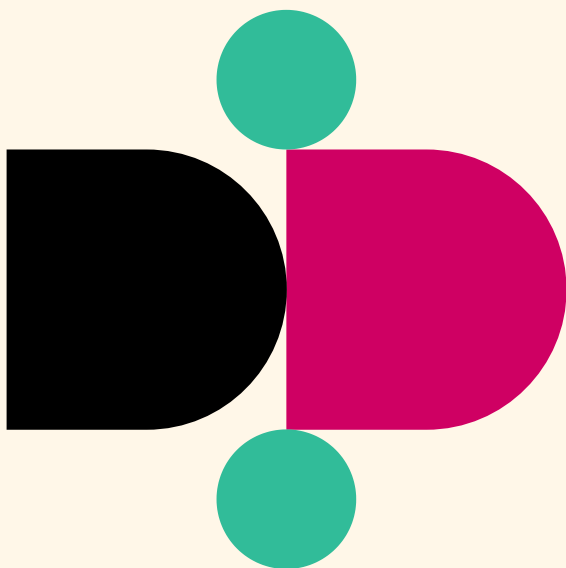
Propriété intellectuelle

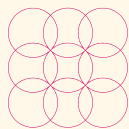
Toute donnée est soumise à la notion de propriété intellectuelle.

- ✦ Éviter, au mieux de sa connaissance, d'utiliser des outils basés sur des contenus ne respectant pas la propriété intellectuelle et vérifier les conditions d'utilisation pour déterminer ses droits de propriété intellectuelle pour créer un contenu.

PARTIE 2

GUIDES D'UTILISATION





Comment communiquer?

Loin des lignes de code et au-delà des mots-clés des moteurs de recherche, la communication avec une IA générative se fait via un agent conversationnel auquel il faut formuler une requête articulée (i.e.

prompt en anglais). L'expérience du *prompt* se veut simple et se passe de science. Les *prompts* sont pensés pour avoir une relation claire entre l'utilisateur/trice et l'IA.



Formuler des *prompts* est une **compétence** à part entière qui relève plus de la psychologie de l'IA que de l'expertise technique. Un-e ingénieur-e informatique ne sera pas plus apte qu'un-e novice pour générer un contenu de qualité s'il/elle ne sait pas expliquer à l'IA quelle est son attente. Il s'agit de quelque chose de comparable à décrire une scène à un-e aveugle. Tout est dans la **formulation**. Ce qui n'est pas dit explicitement n'est pas pris en compte et le sens commun ne suffit pas pour combler les lacunes entre attentes et informations disponibles.



La technique de *prompting* (i.e. *prompt engineering*) est une compétence comme une autre qui implique de prendre en compte un certain nombre d'éléments techniques et/ou créatifs. Surtout, cela implique d'avoir une idée claire de ce que l'on veut. Il existe des outils d'aide à la rédaction de *prompts*.

1 STRUCTURE

De la même manière qu'il n'est pas nécessaire d'être mécanicien-ne pour savoir conduire, on peut utiliser un outil d'IA générative sans maîtriser tous les aspects informatiques de son fonctionnement. Néanmoins, pour être pilote de course, il est utile de savoir ce qu'il y a sous le capot. La première partie de ce guide permet de comprendre la **structure d'une requête**.



2 TECHNIQUE

En deuxième partie, nous avons mis en place une série d'**étapes à suivre** et d'instructions à garder en tête avec des acronymes spécifiques comme moyen mnémotechnique pour appliquer les techniques de *prompting*.

STRUCTURE D'UNE REQUÊTE

Faire un prompt est similaire à faire un gâteau. Avec du beurre, de la farine, du lait et des œufs, on peut tout faire. Le reste est une question de quantités, d'ingrédients supplémentaires et de recette. Les ingrédients sont ici les exemples, les paramètres et le contexte. Il sont essentiels pour construire des consignes efficaces et évaluer les résultats de manière critique.

L'astuce, est de répondre à:
Qui? Quoi? Où?
Comment?
Pourquoi?

QUI?

Lorsque l'on crée un contenu, il est généralement adressé à un public. Comprendre son audience et préciser ce genre d'information est important pour donner un contexte pertinent. Mieux encore, c'est de préciser qui fait la requête.

Audience: À qui s'adresse le contenu?

Création: Qui s'adresse à l'audience?

👉 **Exemple:** Je suis chargée de projet en communication et je rédige un courriel à des étudiant-es universitaires.

QUOI?

Il s'agit de guider l'outil pour faire comprendre notre intention. Comme mentionné, il faut s'imaginer parler à une personne aveugle et être explicite sur ce que l'on souhaite obtenir.

Format: Quel type de contenu?

Sujet: À propos de quoi?

Paramètres: Quelle longueur? Quel ton?

👉 **Exemple:** Fais une liste de cinq idées de titres de conférence sur l'impact des écrans sur le développement des enfants.

OÙ?

Quel est l'objectif de diffusion ou de transmission? Préciser où se trouvera le résultat donne du contexte à l'outil. Lorsque cela est possible, il est utile de préciser la source d'information à utiliser, soit en ligne, soit dans la requête.

Canal: Où sera lu le contenu?

Source: À partir de quoi est-il formulé?

👉 **Exemple:** Dans le texte suivant, identifie trois mots-clés qui serviront à trouver ce texte dans un moteur de recherche: {{insérer texte}}.

POURQUOI?

L'intentionnalité peut être saisie par l'IA générative qui retrouvera un pattern identifié parmi de nombreux documents partageant le même objectif. Il ne faut donc pas hésiter à préciser pourquoi ce contenu est supposé avoir la forme attendue.

Objectifs: Quelle est l'action finale? Quel est le résultat idéal?

👉 **Exemple:** Rédige un paragraphe sur le numérique et l'environnement pour exprimer les enjeux économiques, sociaux et environnementaux tout en gardant une neutralité de point de vue.

COMMENT?

Les outils d'IA génératives peuvent produire et reproduire n'importe quel type de contenu et dans n'importe quel style. Particulièrement, ils sont doués pour reproduire des nuances émotionnelles, sociales ou hiérarchiques. Il est donc judicieux de préciser laquelle est attendu.

Émotion: Quelle note prendre?

Style: Comment habiller le contenu?

👉 **Exemple:** Répond positivement en soulignant l'originalité du projet tout en gardant un ton légèrement enthousiaste mais autoritaire et administratif.

POUR FORMULER UNE REQUÊTE, IL FAUT PROCÉDER PAR ÉTAPE

E Esprit

Il est utile de se mettre dans un état d'esprit de débutant-e ou d'allophone et éviter des formulations trop régionales ou personnelles et ne pas faire appel à des règles tacites ou habitudes.

✗ Ex: Explique la décision X du CODIR concernant le PAT.

✓ Ex: Explique la décisions X du Comité de Direction concernant le personnel administratif et techniques.

T Trame

La véritable puissance de ces outils est de construire une réponse. Posez des questions créatives et critiques pour obtenir des résultats intéressants et éviter les réponses évidentes (e.g. oui/non). Pour ne pas s'égarer, il est mieux d'avoir une idée de la trame voulue et de l'affiner avec des instructions précises et concises.

✗ Ex : La loi sur l'Université est-elle bien?

✓ Ex: Dans le contexte de la réforme de la loi sur l'Université en 2008, est-ce que l'enseignement supérieur et la recherche ont été harmonisés à l'échelle nationale?

A Action

Il faut indiquer l'action attendue. Cela passe par l'usage d'un verbe ou la mention d'une activité, par exemple des verbes de direction comme «discuter», «comparer», «concevoir» ou «évaluer». Le choix du verbe peut grandement influencer le résultat.

Ex: Pour reprendre l'exemple de la loi sur l'Université, préciser de «donner les principaux arguments» ou «d'agir comme un représentant de tel parti politique» permettra de clarifier la réponse et de lui donner une certaine couleur.

P Paramètres

Pour éviter les réponses vagues, non professionnelles, voire même chargées de contenu superflu, il faut affiner votre requête. Par exemple, précisez une méthodologie, le nombre de mots, la structure (e.g. courriel, mots-clés, rapport) ou le format (Python, HTML, CSV, Excel).

Ex: Rédige un paragraphe en 450 mots sur le sujet XY en format Markdown.

E Environnement

Le contexte crée un environnement pour l'outil. Utilisez des mots limitants (e.g. «en Suisse»), avec des exemples connus ou un public défini. Pour cibler le cadre, il est possible de donner le ton (formel, familier, etc.) ou le rôle attendu («une responsable assurance qualité»).

Ex: Élabore un atelier de première année dispensé en ligne pour les étudiantes et étudiants adultes.

Ex: Rédige une réponse compréhensible par un enfant de 8 ans.

Ex: Rédige dans un ton formel à la troisième personne pour un public universitaire.

LES BONS RÉSULTATS DEMANDENT DES EFFORTS APRÈS

A Ajouter

Il n'y a pas de réponse définitive, ce sont des possibilités et des idées. Il faut tester différentes variations de résultats, soit via le même outil, soit sur d'autres outils pour comparer. Cela permet d'explorer plusieurs points de vue ou d'attaquer un problème sous plusieurs angles.

Ex: Propose une autre réponse en prenant le point de vue d'une étudiante de bachelor

Conseil: Poser la même question sur ChatGPT, Bard et Llama.

P Préciser

Le processus est itératif. Cela demande d'élargir son champ d'intérêt à d'autres éléments puis de le recentrer ailleurs avec des consignes supplémentaires (e.g. un suivi sur un aspect de la réponse obtenue). Le but est de guider l'IA et d'identifier les éléments utiles pour les approfondir.

Ex : Précise-moi le troisième argument de la réponse avec un exemple concret.

R Reformuler

On a un plan et parfois il faut le repenser. Un piège courant est la formulation vague ou imprécise des consignes, ce qui peut perturber les algorithmes. Il faut alors simplifier les mots-clés et reprendre les verbes d'action et les paramètres pour obtenir plus de précision et avoir le contexte juste.

Ex: Pour reprendre l'exemple de la loi sur l'Université, on peut dire que plutôt que «la recherche» on s'intéresse «à la collaboration scientifique».

E Évaluer

Les informations sont parfois confuses. Il faut donc défier l'outil et lui demander d'expliquer ses réponses. Il ne faut pas hésiter à exprimer ses doutes et suggérer des approches alternatives.

Ex: J'ai l'impression que le troisième argument n'est pas juste. Peux-tu expliquer pourquoi tu proposes cela?

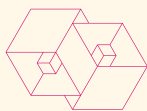
S Scruter

Il s'agit d'identifier les «hallucinations» et les biais potentiels que l'outil aurait produits. Cela passe par la mention des sources à l'appui des éléments générés ou par une vérification minutieuse des faits évoqués. De toutes les étapes, celle-ci est cruciale pour préserver sa propre crédibilité et celle de l'institution.

Ex: Si l'outil n'est pas utilisé comme moteur de recherche, avec comme résultat une liste de liens url, il est crucial de demander explicitement à l'agent de fournir des références à la réponse produite.



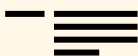
Pour rappel, les individus sont responsables.



Styles de travail

CONSTRUCTION

Il n'y a pas une manière unique de communiquer avec une IA générative pour faire des requêtes. Il est normal, et important, de trouver son propre style, ou celui qui convient le mieux à une tâche donnée.



Une fois pour toutes

L'approche consiste à écrire une seule requête précise. C'est un point de départ à éditer par la suite. Cette approche vise à éviter les efforts répétés ou les itérations successives. Elle s'applique dans divers contextes, tels que la résolution de problèmes, la prise de décisions ou la planification de projets, afin de maximiser l'efficacité et d'économiser du temps et des ressources.



Par couches

Cette approche consiste à faire une requête minimale, à l'évaluer, puis à l'affiner en ajoutant des exemples, du contexte et des paramètres (voir ETAPE) jusqu'à obtenir un ensemble satisfaisant. Cette méthode permet également de diviser un problème complexe en plusieurs niveaux plus simples, chaque couche étant responsable d'une partie du problème. À la fin, il suffit de demander à l'outil d'agrégier les réponses obtenues.



En collage

Cette manière de procéder exige de formuler une requête aussi complète que possible, chargée de contexte, de paramètres et d'exemples, puis d'itérer plusieurs versions de résultats. Puisque l'outil génère quelque chose de différent à chaque demande, il est possible de prendre les meilleures parties de chaque résultat pour assembler quelque chose de nouveau.

Il est impératif d'avoir un comportement engagé. Adopter aveuglement les résultats peut mener à des erreurs, en particulier si la tâche se situe au delà de la frontière «dentelée» où les capacités de l'IA deviennent imprévisibles. D'après Dell'Acqua et al. (2023), deux catégories de pratiques de collaboration émergent entre l'humain et la machine. Ces stratégies distinctes ont été spontanément adoptées dans des cadres professionnels intégrant l'IA générative. Les premiers délèguent les tâches à l'outil ou à eux-mêmes, les seconds intègrent complètement leur flux de travail avec la technologie.



Comportement Centaure

À l'instar de la créature mythique, humains et machines ont étroitement fusionné dans une division stratégique du travail. Les centaures discernent quelles tâches sont mieux adaptées à l'humain ou à l'IA générative en fonction des forces et des capacités de chaque entité et les attribuent en alternance.

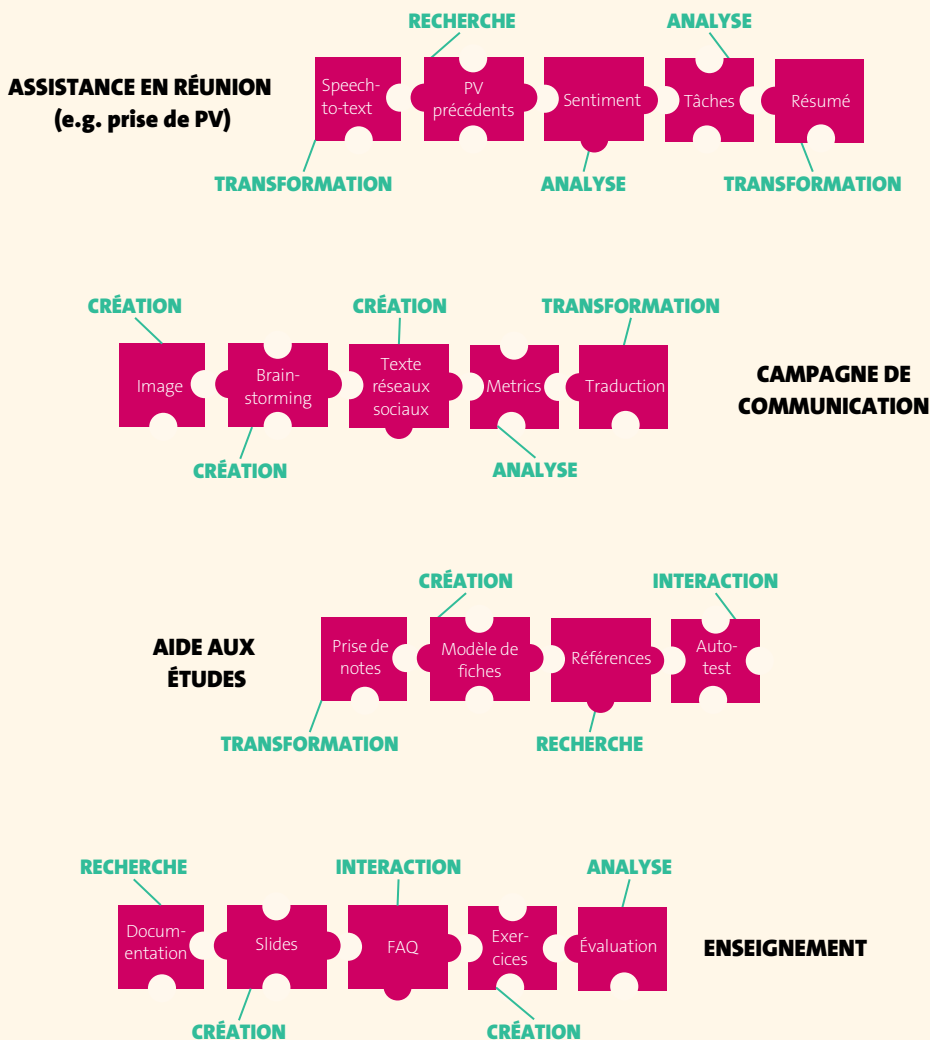


Comportement Cyborg

Les cyborgs intègrent les capacités de la machine et de l'humain au niveau des sous-tâches. On ne sait plus si le résultat a été produit par l'un ou l'autre. Avec cette approche, axée sur l'intégration complexe, il n'y a pas de division claire du travail. Les efforts se mêlent jusqu'à la frontière des capacités de l'IA générative.

COMBINAISONS DE FONCTIONS

Cinq types d'usages ont été identifiés (cf. pages suivantes). Si chacun nécessite une application des principes ETAPE et APRES et peut être travaillé en combinaison de requêtes, il ne faut pas perdre de vue que les usages peuvent également être combinés pour réaliser une tâche ou pour mener à bien un projet.





Choisir un usage

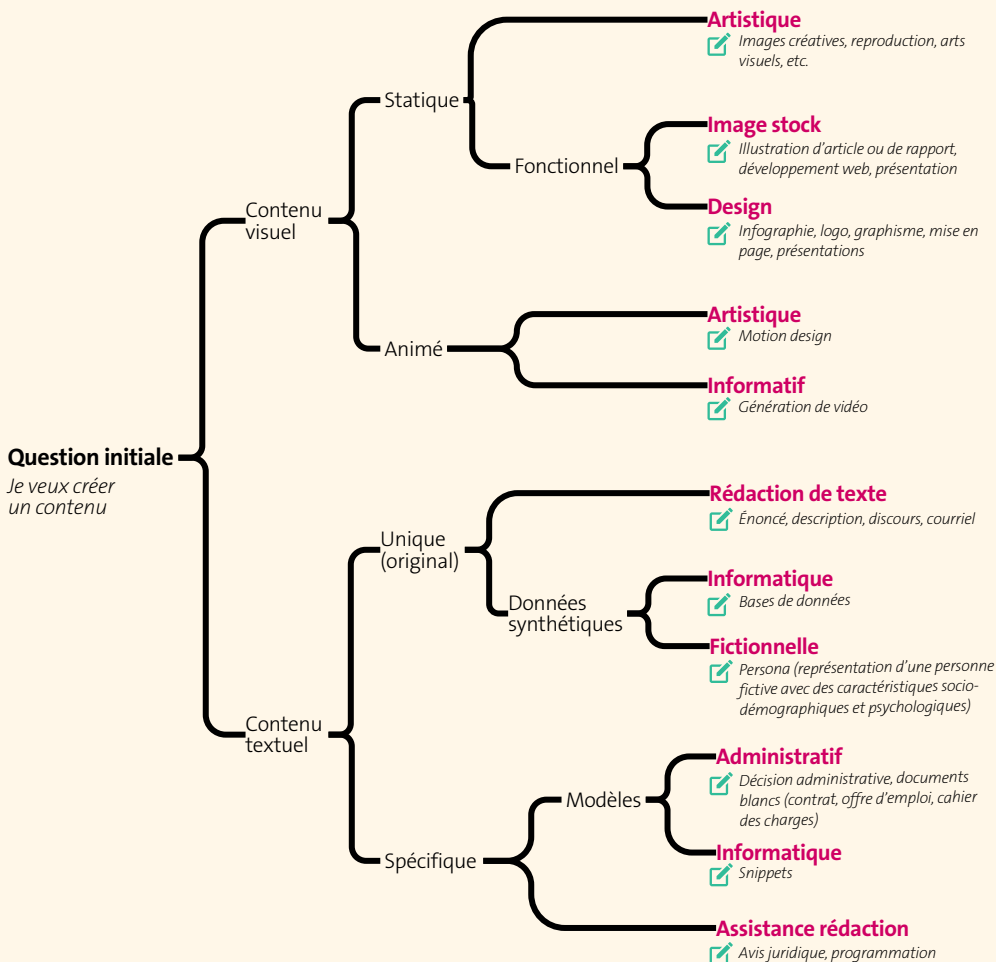
CRÉATION

La création de contenu se réfère au processus où, à partir d'instructions ou de paramètres initiaux, un système génère un contenu entièrement nouveau sans réutiliser ou modifier du contenu existant. **L'output est original.**



Input Description (paramètres, consignes, directives ou données initiales).

Output Nouveau contenu original généré à partir de zéro et basé sur l'input minimal.



TRANSFORMATION

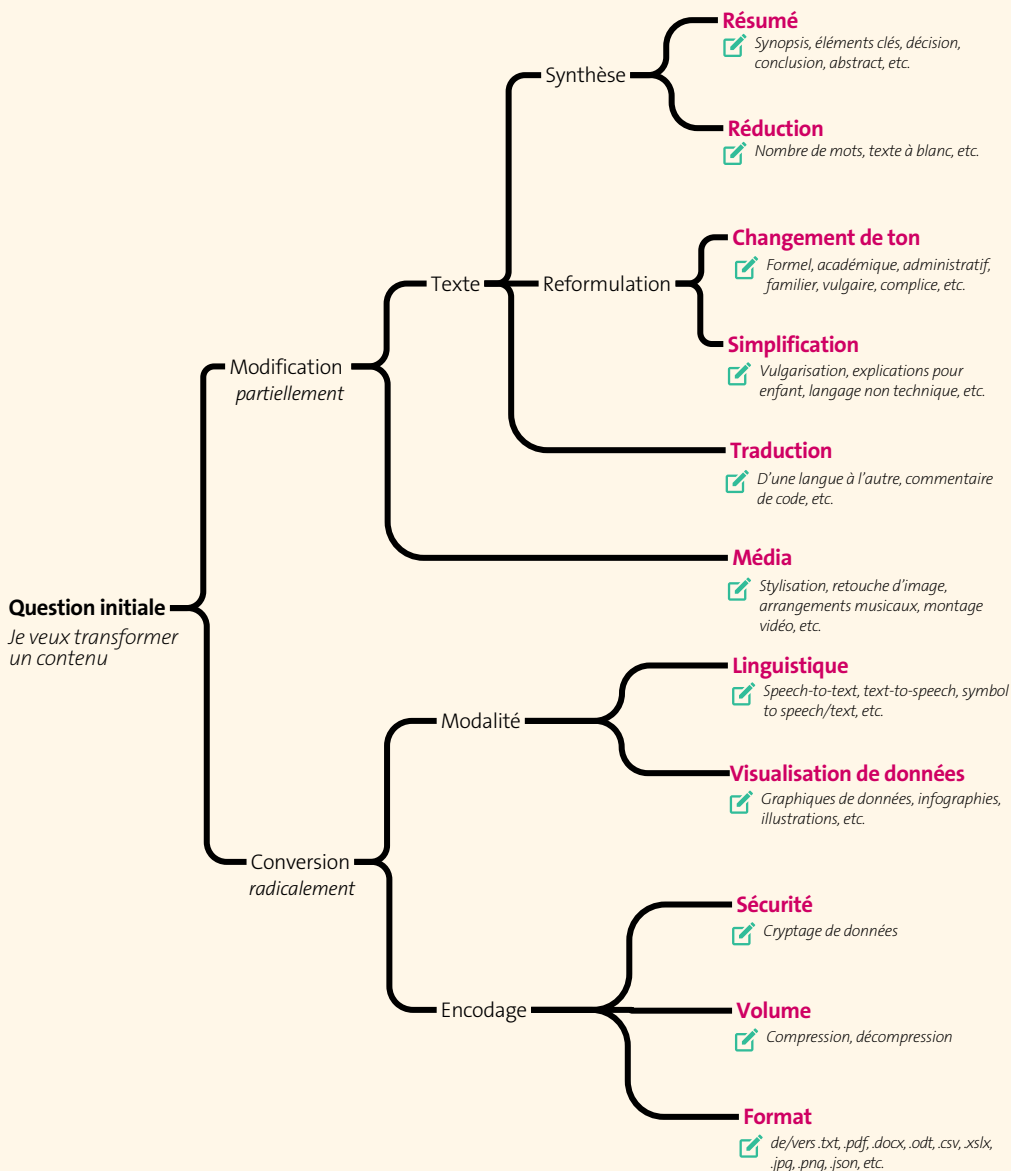
La transformation de contenu implique de soumettre un élément fini à une modification selon des directives spécifiques pour produire un nouveau rendu du contenu.

L'output est une altération de l'input original.



Input Contenu complet avec paramètres, consignes ou directives.

Output Le même contenu sous une forme différente.



ANALYSE

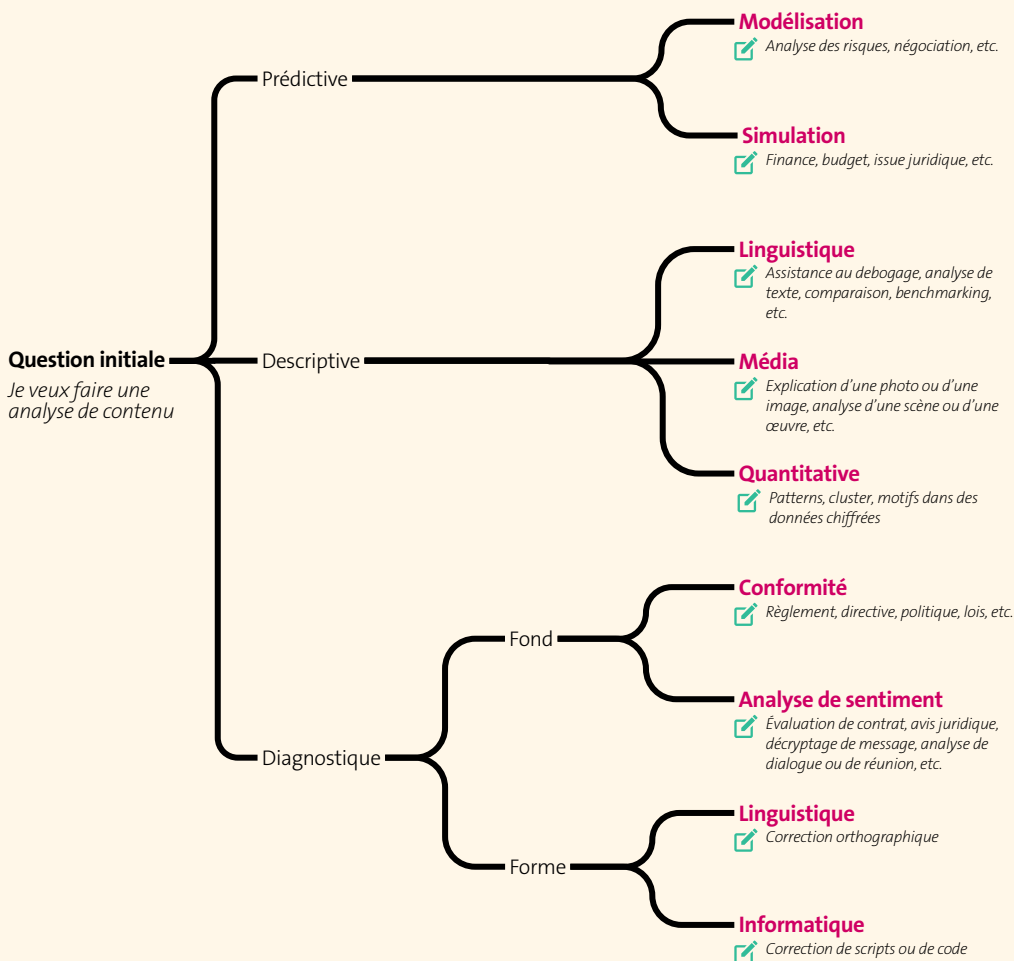
L'analyse de données consiste à examiner des contenus, à les modéliser ou à les nettoyer, dans le but d'en extraire des informations basées sur l'input initial.

L'output est nouveau, basé sur l'input original.



Input Des données brutes ou non avec des consignes ou directives.

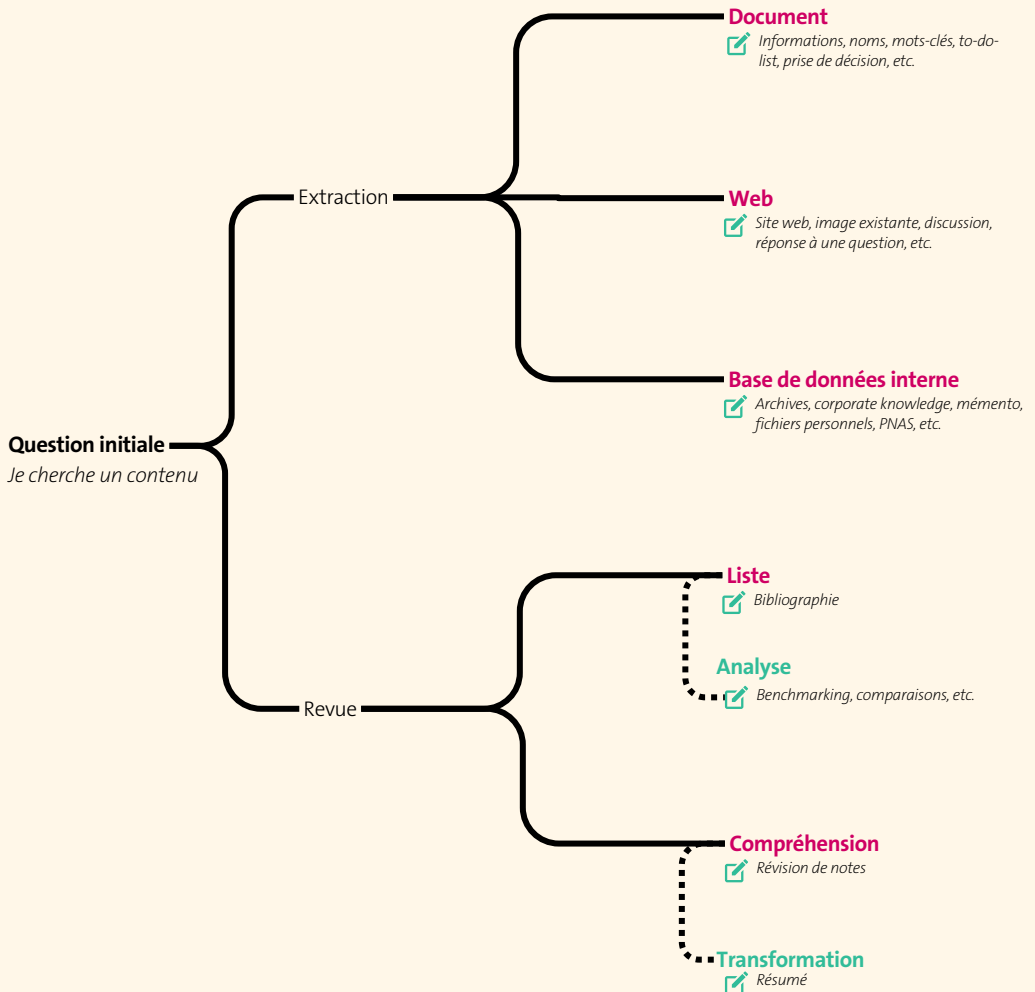
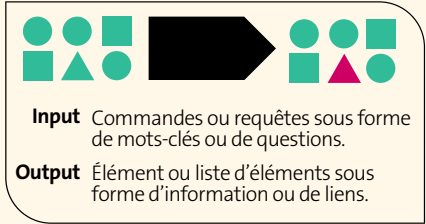
Output Insights, tendances, statistiques, pattern ou conclusion générale.



RECHERCHE

La recherche de contenu est un processus où, sur la base d'un système d'information, une requête permet d'obtenir des éléments pertinents à propos du contenu visé.

L'output est un élément désigné par l'input original dans l'ensemble existant.



INTERACTION

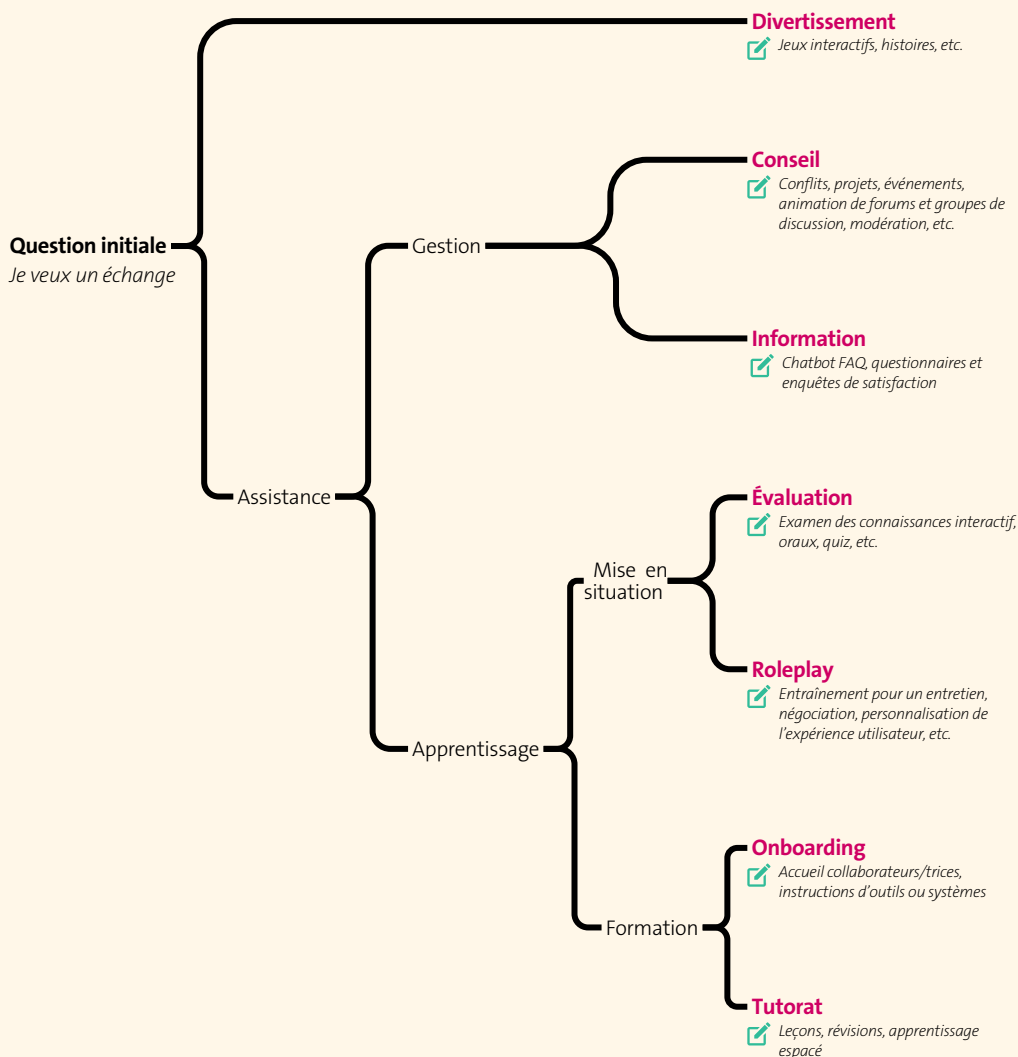
L'interaction se rapporte à la communication et à un échange bidirectionnel de contenu entre un utilisateur et un système.

Itérativement, le système et l'utilisateur/trice produisent des outputs originaux selon les inputs de l'autre.



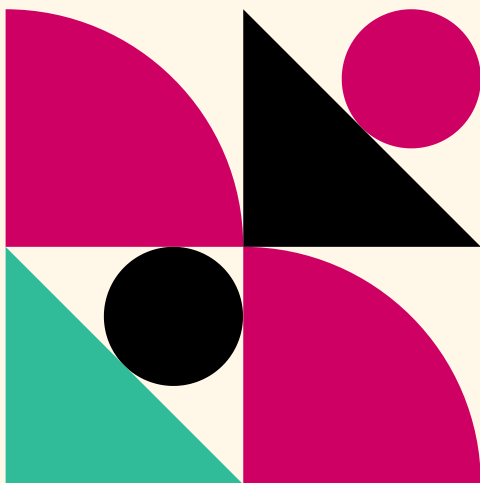
Input Commandes, requêtes, réactions.

Output Réponses, feed-backs, questions, réactions.



PARTIE 3

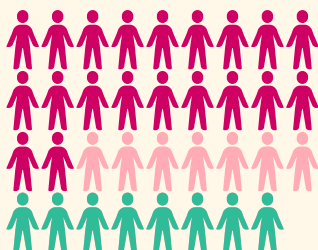
COMPTE RENDU DE L'USAGE DE
L'IA À L'UNIGE





Adoption à l'UNIGE

L'émergence d'une technologie aussi disruptive place l'UNIGE à l'intersection d'un besoin d'innovation et de préservation. Des entretiens menés en septembre 2023 auprès d'un échantillon de collaborateurs et collaboratrices du Personnel Administratif et Technique de l'Université concernant leur rapport à l'IA générative dévoilent l'existence de trois catégories d'individus. L'exploration de ces profils offre un aperçu des dynamiques pouvant influencer une adoption ou une réticence vis-à-vis de cette technologie.



■ **Aucun usage**
■ **Usage à titre personnel**
■ **Usage professionnel**

Il y a un **faible taux d'adoption spontanée** dans les pratiques professionnelles. Ce résultat souligne une prudence ou un désintérêt de la part du personnel.

Des variations sont clairement **marquées entre les différents métiers**. Le secteur de la communication émerge comme précurseur, tandis que juristes et secrétaires témoignent d'une utilisation actuellement limitée.

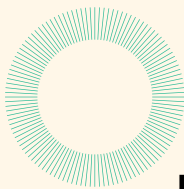
Les cas d'exploration hors de toute application professionnelle signalent un **potentiel d'adoption professionnelle** plus large, toutefois conditionnée par un soutien accru, tel que la formation et l'accès à des ressources adaptées.

De plus, la diversité sous-jacente des pratiques suggère l'importance d'une **approche personnalisée** d'appropriation de ces outils au sein de l'institution.

In fine, ces entretiens mettent en lumière certaines **attentes et appréhensions** face à l'IA générative, indépendamment du degré d'utilisation actuel.

ÉCOSYSTÈME

profil du personnel



Les corneilles

Intégration-innovation

Ouvert-es à l'inconnu et promptes à adopter de nouvelles méthodes, ils/elles ont déjà utilisé l'IA dans un cadre professionnel. Bien qu'ils/elles ne représentent de loin pas une majorité, leur présence reste significative.



Curiosité

La curiosité est le moteur principal de l'exploration initiale des outils disponibles. L'apprentissage est autodidacte, souvent via des vidéos ou des cours en ligne. En dépit d'une certaine familiarité, ou d'une aisance technologique, ils/elles font souvent cette démarche sans posséder d'expertise informatique, naviguant entre une utilisation intuitive et des connaissances de base. Ils/elles reconnaissent une utilité variable des outils, avec une valeur ajoutée dans certains domaines spécifiques seulement (i.e. création de contenu ou traduction).



Temps

Le principal obstacle à l'adoption est temporel: l'absence de méthodologie fiable et le temps requis pour maîtriser ces outils sont des contraintes limitant un usage plus généralisé dans leur travail quotidien. De plus, un manque de compréhension claire des usages possibles peut présenter un frein.



La diversité des applications concrètes explorées par ce groupe témoigne d'une recherche d'efficacité individuelle propre à chaque métier mais, surtout, à chaque besoin.

Création de contenu

Selon les cas, l'IA générative est un partenaire créatif, assistant à la rédaction et à l'idéation ou encore remède contre la page blanche. En communication, i.e. événementielle, elle est particulièrement utilisée pour la création de contenu visuel d'illustration.

Transformation

L'IA générative sert de traduction ou d'amélioration des textes rédigés dans une langue non maternelle. Elle joue également le rôle de synthétiseur efficace, condensant ou résumant courriels et autres documents.



Ces pratiques s'accompagnent toutefois de certaines réticences initiales et d'une comparaison inévitable avec des outils traditionnels établis, parmi lesquels l'IA se taille une place de choix, en imposant un nouveau standard d'accessibilité, en particulier pour les utilisateurs et utilisatrices non spécialisés.



Les pandas

Tradition-fidélité

Peu enclin-es au changement, n'y trouvant que peu d'intérêt, voire ouvertement critiques, ils/elles n'ont pas encore essayé ces nouveaux outils, préférant s'en tenir à ce qui est connu et maîtrisé.



Pertinence

Le manque d'opportunités concrètes, et d'applications évidentes, pour intégrer l'IA dans leurs activités se présente comme un facteur limitant. Plus particulièrement, des contraintes liées à des tâches professionnelles très spécifiques écartent, pour certain-es, l'IA de leur éventail d'outils. De plus, le manque de temps est souvent invoqué comme un frein majeur à l'investigation et à l'adoption de l'IA.



Technologie

Des appréhensions personnelles ont été exprimées, comme une certaine anxiété à l'idée de faire des erreurs ou de ne pas utiliser correctement la technologie. Ces inquiétudes mettent en relief les défis liés à la formation et à la perception de soi dans l'adoption de nouvelles solutions. Des préoccupations techniques, telles que les restrictions d'accès ou une familiarité insuffisante avec les outils, sont fréquemment mentionnées. Enfin, le scepticisme et la prudence quant à la fiabilité et à l'utilité de l'IA influencent également l'hésitation à l'adopter, ces sentiments étant particulièrement prégnants chez celles et ceux ayant une approche plus conservatrice vis-à-vis de la nouvelle technologie.



La saturation des outils gratuits, soulignée par certain-es, démontre des limitations d'accès pouvant entraver l'expérimentation et l'adoption de l'IA. La perception du risque associé à l'IA est également un frein significatif, témoignant d'une conscience aiguë des implications potentielles d'une mauvaise utilisation.

Les chats

Exploration-curiosité

Explorant de nouvelles solutions pour les comprendre et les évaluer, sans forcément s'y attacher, ils/elles ont voulu tester ou «voir» à quoi cela ressemblait, mais n'ont pas souhaité l'intégrer dans leur cadre professionnel.



Connaissance

L'intérêt initial et l'exploration de l'IA découlent d'une curiosité et d'un désir de connaître et comprendre l'outil afin d'évaluer son potentiel. Comme pour le cadre professionnel, ce premier contact et cette initiation personnelle apporte une satisfaction variable, un aspect qui se reflète dans la diversité des attentes et des applications.



Usage dans les tâches quotidiennes

Certain-es trouvent un écho favorable dans des applications domestiques et sociales, depuis la rédaction d'articles (i.e. syndrome de la page blanche) jusqu'à l'aide aux devoirs, en passant par des recherches juridiques complexes. L'aspect ludique n'est pas à négliger, les explorations sont souvent menées entre ami-es, soulignant l'aspect convivial et décontracté que peut prendre l'IA dans la sphère privée.



Scepticisme

Les doutes quant à l'utilité de ces outils dans le cadre professionnel subsistent. En outre certaines restent dubitatifs et dubitatives face à l'impact réel de l'IA générative sur leurs missions spécifiques à l'Université. D'autres se heurtent à des limitations temporelles ou financières. De plus, une critique plus philosophique émerge; si ces outils sont capables d'égaler la performance humaine dans certaines tâches, la valeur intrinsèque des ces dernières est remise en question. Cette remarque soulève notamment des interrogations fondamentales sur la nature du travail significatif à l'ère de l'automatisation croissante.



L'intérêt pour l'IA générative transcende les différents métiers au sein de l'Université de Genève, révélant une ouverture quant à son adoption.

Les assistant-es de direction interrogé-es sont unanimement favorables à l'intégration de ces outils dans leur travail, sous réserve d'un accompagnement. Les administrateurs/trices ayant participé à l'enquête voient également un atout pour leurs équipes, tout en reconnaissant qu'ils/elles ne peuvent pas répondre à l'ensemble des besoins spécifiques individuels. Ils/elles estiment que son acceptation peut varier au sein des différents groupes, tout en insistant sur l'importance de la protection des données. Le besoin de fiabilité est également mis en avant par le domaine juridique.

En communication, l'opinion est partagée parmi les personnes interrogées. Si une moitié est favorable, l'autre exprime des réserves, craignant un frein à la créativité ou des mauvaises pratiques. L'intérêt pour des tâches fastidieuses est mitigé par des limites techniques (orthographe des noms propres, images approximatives, etc.), questionnant la véritable valeur ajoutée. L'automatisation ne résout pas les défis fondamentaux, et le doute existe de savoir si cela entraînera une réduction ou une augmentation du travail.

Ces entretiens ont également mis en lumière un certain nombre d'attentes et d'appréhensions, indépendamment du degré d'utilisation actuel. Ce panorama constitue la base sur laquelle les guides d'utilisation (cf. partie 3) ont été élaborés afin de rester au plus près des besoins du personnel administratif et technique.

ATTENTES

Rapidité et efficacité

L'IA générative devrait contribuer à accélérer les processus, en particulier pour les tâches jugées sans valeur ajoutée. Elle pourrait prendre en charge les aspects répétitifs ou fastidieux du travail, tels que naviguer à travers de larges volumes d'informations ou mettre en relation des données. L'automatisation pourrait toucher, entre autres, la rédaction de procès-verbaux, la révision de contrats ou la synthèse des textes. Cette attente d'efficacité s'accompagne d'une certaine ambivalence. Si certain-es anticipent la délégation des travaux de rédaction, d'autres préféreraient la cantonner à des tâches rébarbatives, tout en préservant les fonctions plus intellectuelles et créatives.

Simplification et assistance

Il y a une attente claire pour que l'IA simplifie les processus. D'une part, en tant que première ligne ou en guise de tampon, comme support en répondant aux questions moins urgentes, habituelles ou pour des requêtes sur des sujets mouvants, notamment dans des domaines où la législation change fréquemment. D'autre part, l'IA pourrait agir comme soutien pour surmonter l'isolement ou encore pour combattre la page blanche. Dans ce contexte, l'IA serait une assistance dans le processus de création, telle une aide complémentaire, que ce soit pour la rédaction de courriels, la création d'images, ou l'adaptation visuelle dans divers formats et styles.

Autonomie et innovation

L'IA est perçue comme un moyen d'acquérir une plus grande agilité et autonomie dans l'accomplissement de missions originales ou dans la gestion de tâches complexes, qui apparaissent parfois en dehors du domaine d'expertise. Par exemple, si un graphiste professionnel a peu d'intérêt pour la génération d'images qu'il sait produire seul, il existe dans plusieurs métiers une multitude de tâches liées à la communication difficilement réalisables sans expertise. L'IA générative permettrait de générer des images originales pour le web en évitant les problèmes de droits d'auteur-e et l'homogénéité des banques d'images. Les outils de traduction constituent un autre exemple concret pour pallier le manque de moyens ou de disponibilité d'experts et gagner en autonomie.

Retour à l'essentiel

Ces outils pourraient libérer du temps afin de pouvoir se concentrer sur les aspects plus gratifiants et humains des métiers. Cette recherche de gain de temps n'est pas simplement une question d'efficacité, elle reflète un désir profond de se tourner vers des activités plus enrichissantes et de s'épanouir dans son activité professionnelle. L'IA serait une assistance potentielle capable d'aider à se débarrasser des charges annexes, souvent chronophages ou superflues, au cours d'une journée de travail, permettant ainsi de réduire la surcharge et de s'investir dans son développement personnel et professionnel.

CRAINTES ET FREINS

Perte de compétences

La perte de valeur intrinsèque personnelle dans les activités professionnelles fait peur. Il est primordial de préserver les aspects les plus gratifiants qui constituent le cœur du métier, i.e. la résolution autonome d'un problème apporte une satisfaction non substituable. L'anxiété de voir s'installer une addiction ou une paresse intellectuelle, menant à une atrophie des compétences, est également palpable, à l'instar de la correction orthographique automatique empêchant de s'améliorer. La standardisation de la pensée et la production d'informations sans personnalité suscitent des inquiétudes quant à la dévaluation de l'individualité dans le travail.

Données personnelles

La possibilité de manipulation d'informations confidentielles hébergées dans des systèmes externes à l'UNIGE fait craindre une insuffisance en termes de protection et de confidentialité. Le manque de maîtrise sur le traitement, le stockage et la sécurité des données personnelles, notamment la nécessité de les anonymiser, sont des préoccupations significatives.

Digital gap

Il existe le risque d'un fossé numérique entre celles et ceux qui adoptent rapidement les nouvelles technologies et celles et ceux qui sont réticent-es ou moins agiles. Ce clivage risque d'exacerber les peurs liées au remplacement ou à l'obsolescence professionnelle. La différence d'accès entre étudiant-es et enseignant-es, par exemple, aux versions plus avancées d'outils est également une source d'inquiétude.

Réticences technologiques

Une remise en question plus philosophique de la modernité et du progrès technique se manifeste par une résistance à ce qui est perçu comme un progrès technologique injustifié. L'éthique, l'utilité et la valeur ajoutée réelle de l'IA sont questionnées: la technologie est-elle adoptée pour répondre à un besoin authentique ou simplement parce qu'elle est disponible?

Fiabilité et qualité

L'IA générative pourrait ne pas être adaptée aux spécificités requises par certains rôles et manquer de la subtilité nécessaire pour gérer des communications délicates. Les nuances dans la rédaction de procès-verbaux, courriels sensibles ou documents politiques sont, par exemple, difficiles à capturer. Cette insuffisance perçue, compte tenu que le résultat n'est souvent qu'une ébauche nécessitant d'être retravaillée, soulève la question de l'utilité de l'IA si l'on doit doubler le travail pour s'assurer de sa véracité, en particulier lorsque des enjeux juridiques sont impliqués. De plus, les lettres de motivation et autres documents pourraient perdre de leur singularité.

Rapport à la machine

L'ambivalence de donner des ordres à une technologie pouvant reproduire une interaction humaine authentique, comme saluer ou remercier, suscite un malaise. De plus, il existe une méfiance généralisée à l'égard du contrôle exercé par des robots et une incompréhension des outils, accompagnée de craintes éthiques sur le rôle grandissant de la technologie dans la vie professionnelle.

Surcharge

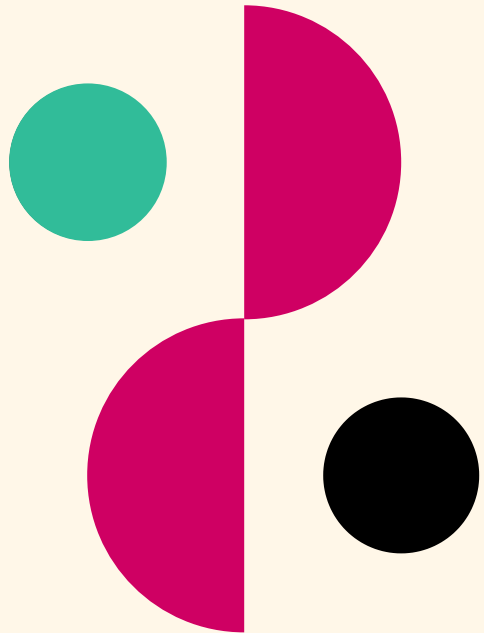
L'idée fausse que l'IA va libérer du temps en automatisant des tâches est contrecarrée par la crainte qu'elle entraîne une augmentation du volume de travail, en remplaçant les travaux simples par des responsabilités plus complexes.

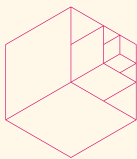
Manque d'intérêt ou de temps

Plus qu'une crainte, il ressort un manque d'intérêt manifeste, au mieux un attentisme pour évaluer l'engagement des autres. Les équipes se sentent surchargées et manquent de motivation pour se former à de nouveaux outils, d'autant plus que la technologie complexe demande un apprentissage constant. Le manque de temps et la volonté de laisser les autres «défricher» le terrain renforcent ce frein à l'adoption de l'IA.

PARTIE 4

CONCLUSION





Conclusion

La démocratisation des IA génératives marque un tournant significatif dans le paysage technologique contemporain. L'UNIGE embrasse cette révolution à travers de multiples initiatives, reconnaissant ces outils comme de puissants leviers d'optimisation. Une partie de son personnel a déjà spontanément expérimenté une transformation notable dans la gestion de ses tâches professionnelles. Ce document ouvre la voie au reste de la communauté universitaire tout en garantissant un usage informé et responsable.

L'intégration des IA génératives dans l'environnement professionnel n'est pas sans risques. Les questions éthiques, la fiabilité des informations générées et la sécurité des données sont des préoccupations majeures. Chaque utilisateur et chaque utilisatrice doit faire preuve d'une intégration responsable de ces technologies et comprendre que la responsabilité finale repose sur l'individu. Il est donc primordial de garder un esprit critique vis-à-vis d'un outil qui est complémentaire au processus décisionnel.

À l'issue de la rédaction de ce document, il apparaît que la mise à disposition de guides et de recommandations ne sera pas suffisante pour mener une transformation aussi profonde de l'écosystème universitaire. Ce document n'est qu'une étape vers une adaptation plus large et systémique de notre institution, portée par le plan stratégique de l'UNIGE, qui ne pourra se réaliser qu'en s'appuyant sur l'engagement, la curiosité et la collaboration de toutes et tous.



Intelligence artificielle générative

Guide à l'intention de la communauté universitaire

Le présent document a été créé par le Bureau de la transformation sous mandat du secrétaire général de l'Université de Genève et publié pour la première fois en juin 2024.



Nous exprimons notre sincère gratitude à tous les experts et toutes les expertes consultées, ainsi qu'aux représentant-es des différents corps de métier interrogé-es pour leur investissement et leur soutien affirmé.

ENQUÊTES & RÉDACTION

Dr KHAN Guive
Dr THÉZÉ Raphaël

GRAPHISMES

Dr THÉZÉ Raphaël

EXPERT-ES

Prof. BENHAMOU Yaniv
Prof. FLEURET François
Dre FRIHA Lamia
HUGENTOBLE Alain

MÉTHODOLOGIE

Un groupe de travail, formé en juillet 2023 et composé de membres du Bureau de la transformation numérique (BTN) et de la Division des systèmes et technologies de l'information et de la communication (DiSTIC), a été mandaté pour identifier les applications les plus pertinentes des outils en ligne d'IA génératives et pour concevoir des stratégies en vue de leur intégration par l'administration.

Lors des étapes de validation par diverses instances de gouvernance, la portée du document, initialement prévu pour le PAT, a été élargie à toute la communauté universitaire, à la suite du constat de son champ d'application plus global.

Pour élaborer ce document, le BTN a mené 35 entretiens avec des représentant-es de divers métiers au sein de l'UNIGE pour comprendre leurs besoins et cibler les applications concrètes de l'IA générative. Ces entretiens ont été complétés par des discussions avec des expert-es et une revue de la littérature sur les IA génératives. De plus, des outils d'IA générative (e.g. GPT-4) ont été utilisés comme aide à la rédaction des résultats et des guides.

SITES CONSULTÉS

Writer (start-up blog)

<https://writer.com/guides/prompt-crafting/>

Let's enhance (start-up blog)

<https://letsenhance.io/blog/article/ai-text-prompt-guide/>

Enrollify (Higher-ed marketing)

<https://www.enrollify.org/blog/chatgpt-cheat-sheet-for-higher-ed-your-guide-to-ai-powered-marketing>

RESSOURCES

Prise de position UNIGE

unige.ch > numérique > IA

Prise de position HES-SO

numerique.hes-so.ch > course > task force AI education

Guide Dall.E par Dall.ery Gall.ery

dallery.gallery > The prompt book

Guide IA pour les entreprises - DGDERI

ge.ch > dossier > entreprises et numerique > ressources et outils > guides

RÉFÉRENCES

-  Chui, Michael, Hazan, Eric, Roberts, Roger, Singla, Alex, Smaje, Kate, Sukharevsky, Alex, Yee, Lareina & Zimmel, Rodney (2023). The Economic Potential of Generative AI: The Next Productivity Frontier, McKinsey & Company.
-  Deloitte (2023). Generative AI's Fast and Furious Entry into Switzerland: Usages and Attitudes of the Swiss Workforce Towards Generative AI. Deloitte AG.
-  Dell'Acqua, Fabrizio, McFowland, Edward, Mollick, Ethan R., Lifshitz-Assaf, Hila, Kellogg, Katherine, Rajendran, Saran, Kraymer, Lisa, Candelon, François & Lakhani, Karim R. (2023). Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality. Harvard Business School Technology & Operations Mgt. Unit, Working Paper N° 24-013.
-  Flahaux, José Roger, Green, Brian Patrick & Skeet, Ann Greg (2023). Ethics in the Age of Disruptive Technology: An Operational Roadmap, The ITEC Handbook, Santa Clara: Santa Clara University
-  Fleuret, F. (2023). The Little Book of Deep Learning. A lovely concise introduction.
-  OpenAI. (2023). ChatGPT-4 (version du 15 octobre) [Large Language Model]. <https://chat.openai.com/chat>

Ce travail par le Bureau de la transformation numérique est sous licence CC BY-NC-SA 4.0 @2024 Université de Genève



**UNIVERSITÉ
DE GENÈVE**

Guide téléchargeable en ligne



www.unige.ch/numerique/IA



numerique@unige.ch



[@unigenumerique](https://twitter.com/unigenumerique)



Général Dufour 24, 1204 Genève, CH