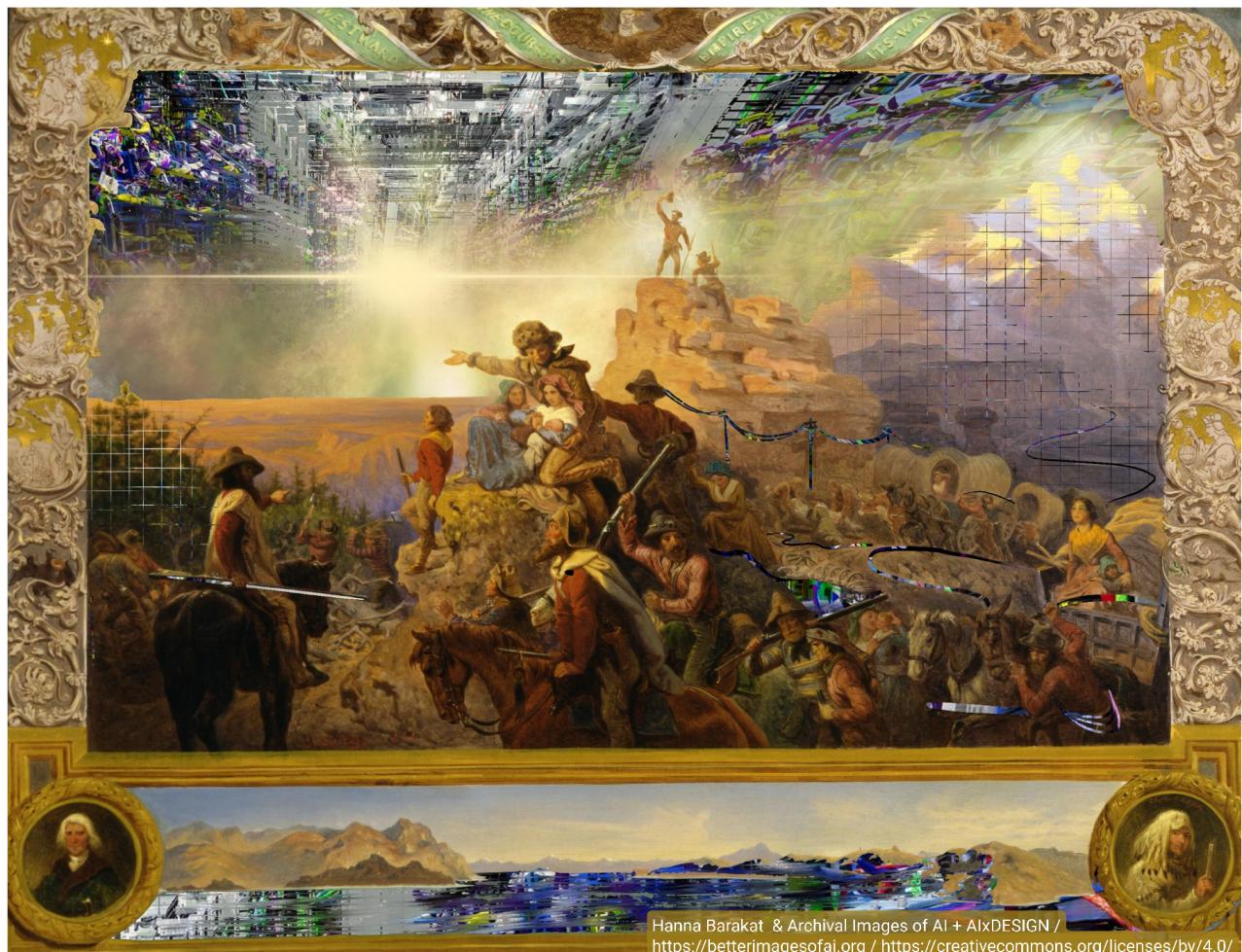


Questionner l'IA, son monde, son histoire et sa culture

Livret et wiki - version imprimable



Hanna Barakat & Archival Images of AI + AIxDESIGN /
<https://betterimagesofai.org> / <https://creativecommons.org/licenses/by/4.0/>

 **NUMETHIC**

conseil supérieur
de l'éducation
aux médias
csem

 **FÉDÉRATION
WALLONIE-BRUXELLES**

philocité

Avec le soutien de la Fédération Wallonie-Bruxelles et du Conseil Supérieur de l'Éducation aux Médias



Table des matières

Livret d'autoformation	1
Une collaboration entre NumEthic et PhiloCité	2
NumEthic	2
PhiloCité	2
Préambule	3
Introduction	4
Ce que l'on fait avec ce livret	4
Comment on le fait	5
À qui le livret s'adresse	7
Qu'est-ce que ça fait de travailler avec le livret	7
Présentation des chapitres	7
Chapitre 1 – Nos cultures et imaginaires de l'IA	9
Contexte du chapitre	9
Genèse du transhumanisme : le choix d'un mythe	10
La cybernétique ou le traitement identique de l'humain et de la machine	12
Un contexte idéologique	13
Atelier 1 : Sonder nos imaginaires sur l'IA	14
Enjeux	14
Présentation de l'atelier et consignes	14
Premier temps : Visionner un film ou un court-métrage	14
Deuxième temps : Discussion collective	15
Variations possibles	16
Debriefing	16
Chapitre 2 – Se repérer par rapport à l'IA dans nos pratiques médiatiques quotidiennes	17
Contexte du chapitre	17
L'IA comme médiation	17
Le milieu de l'IA	18
Atelier 1 - Petites histoires de nos usages	20
Enjeux	20
Présentation de l'atelier et consignes	20
Variations possibles	21
Debriefing	21
Atelier 2 - Enquête sur nos pratiques médiatiques	22
Enjeux	22
Présentation de l'atelier et consignes	22
Temps 1 : Enquête individuelle	22
Temps 2 : Par petits groupes de 3-4 personnes (premier temps)	23
Variations possibles	24
Debriefing	25

Table des matières

Chapitre 3 – Culture technique : entrer dans le monde des machines	26
Contexte du chapitre	26
La culture technique	26
La culture technique spécifique à l'IA	27
Atelier 1 : Représenter le fonctionnement d'une IA	28
Enjeux : attiser la curiosité et étoffer la relation	28
Présentation de l'atelier et consignes	29
Variations possibles	30
Debriefing	31
Atelier 2 – Langage et statistiques	31
Enjeux de l'atelier	31
Présentation de l'atelier et consignes	32
Debriefing	33
Chapitre 4 – Machines et conflits de valeurs	34
Contexte du chapitre	34
Éviter la moralisation	34
Discuter les conflits de valeurs	35
Atelier 1 : Esprit critique et analyse d'images	36
Enjeux de l'atelier	36
Présentation de l'atelier et consignes	37
Première séquence	37
Deuxième séquence	38
Variations possibles	38
Debriefing	38
Atelier 2 : Des valeurs dans la machine	39
Enjeux	39
Présentation de l'atelier et consignes	40
Variation possible	41
Debriefing	41
Conclusion	43
Annexe : quelques textes et citations	44
Chapitre 1	44
La méchanceté	44
La gouvernance par les nombres	45
Du mode d'existence des objets techniques	46
Chapitre 2	46
Fragment d'un village global	46
Chapitre 3	47
Du mode d'existence des objets techniques	47
Psychosociologie de la technicité	48
Persistance du merveilleux	48
La gouvernance par les nombres	49
Chapitre 4	50
De la bêtise numérique	50
Annexe : Ressources pour les ateliers	51
Esprit critique et analyse d'images	51
Des jeunes avec un ordinateur	51
Une classe d'électromécanique	52

Table des matières

Industrie textile au Ghana	53
Mary Jackson	57
Métier à tisser en 1935	60
Dessin du futur	60
Dessiner le fonctionnement de l'IA	62
Wiki	65
L'histoire de l'IA	66
Une histoire multiple de l'intelligence artificielle	66
L'idée de l'intelligence artificielle	66
La cybernétique	66
L'histoire technique et scientifique	67
Les automates	68
L'arrivée de l'informatique	68
Dans les années 2000	69
Les grands modèles de langage	70
Fonctionnement de l'IA	72
Algorithme	73
Les algorithmes	73
Histoire des algorithmes	73
Algorithme explicite et implicite	73
Apprentissage machine	75
Histoire de l'apprentissage machine	75
Neurone artificiel	76
Les réseaux de neurones artificiels	77
Dans les grands modèles de langage	79
Statistiques et grands modèles de langage	80
Le milieu des statistiques et les grands modèles de langage	80
Les grands modèles de langage	80
Alignement	81
Les probabilités	82
Le milieu des probabilités	82
Dans les grands modèles de langage	83
Dans les algorithmes	84

Livret d'autoformation

Une collaboration entre NumEthic et PhiloCité

NumEthic

Ni techno-rabat-joies, ni techno-enthousiastes, les membres de NumEthic souhaitent développer une posture techno-critique. Le terme « critique » ici est entendu dans le sens de « qui est capable de discernement, de jugement ». Afin d'acquérir cette aptitude de discernement, il paraît essentiel à NumEthic de cultiver les connaissances qui lui sont nécessaires et de créer des outils de réflexion qui permettent de penser les enjeux du numérique, en particulier dans le secteur de l'éducation. La démarche de questionnement nous paraît essentielle, afin de se garder de tenir un discours normatif ou moralisateur face aux publics que nous rencontrons.

Créée en 2021, NumEthic a pour vocation de : sensibiliser aux questions relatives à l'emploi d'un numérique éthique, mettre en évidence des outils alternatifs fonctionnels, congruents, respectueux et responsables ; créer une communauté / un collectif de personnes (de tous horizons) dont les préoccupations rejoignent celles mises à l'honneur par l'association ; mettre sur pied des actions concrètes (formation, journées de rencontre...).

PhiloCité

De la maternelle à l'université en passant par les maisons de jeunes, les centres culturels, ludothèques, les IPPJ et les prisons, PMS, écoles de devoirs, maisons de quartiers, centres de soins, etc., **PhiloCité** diffuse depuis 15 ans dans l'espace public les outils de la philosophie. Considérant que la philosophie – son histoire, ses méthodes et ses outils – est un vecteur d'émancipation, PhiloCité se donne pour objectif, notamment à travers les très nombreux ateliers philo qu'elle organise chaque année, de toucher et de mélanger tous les publics, en particulier ceux qui spontanément ne se dirigeaient pas vers la philosophie.

NumEthic et PhiloCité collaborent depuis 2022 autour du numérique et de l'éducation aux médias en développant une approche de ces questions à partir des outils de la philosophie pratique. Animations avec des jeunes et des adultes, conférences, journées d'échanges et de pratiques, formations d'enseignants, sont autant d'activités qui nourrissent cette collaboration. Ce livret est le fruit des expériences de terrain déjà menées. Nous tenons à remercier en particulier les personnes avec qui nous avons eu la joie et le plaisir de tester les différents modules de ce livret : les travailleurs du secteur associatif qui se sont réunis pour une journée de formation à Barricade à Liège, ainsi que les deux groupes constitués par l'Espace Public Numérique de la Ville de Liège / Fétinne et le Centre Informatique Sombreffois. Un grand merci de nous avoir chaleureusement accueillis.

Préambule

Info

Le présent livret a été publié le **23 novembre 2025**. L'histoire de l'IA ne s'arrêtant pas à cette date, des mises à jours et corrections sont peut-être disponibles. N'hésitez pas à consulter la dernière version sur notre wiki : <https://ia.numethic.education>.

Introduction

« *Si nous étudions le processus pendant qu'il nous frappe et ce que nous pouvons y faire, nous pourrons alors nous en sortir* » Marshall McLuhan, cité par Anne Alombert (De la bêtise artificielle, p.30)

Ce que l'on fait avec ce livret

Faire de l'IA un sujet de réflexion commun

« L'IA existe, il va falloir apprendre à bien l'utiliser ». Tel est le discours dominant depuis l'émergence des grands modèles de langage, des logiciels de génération d'images, de musiques et de vidéos.

Nous partons de l'hypothèse selon laquelle c'est le **déficit de culture numérique** qui produit à la fois le sentiment de fascination et d'écrasement face aux IA. Par « déficit de culture numérique », nous entendons le manque d'espace de discussion collective (que ce soit entre les travailleurs d'une équipe – qui a envie de prendre le temps de penser des problèmes informatiques dans le cadre professionnel ? – dans le cadre familial entre les jeunes et les adultes, etc.), le manque d'intérêt et d'attention que nous portons à nos usages quotidiens de ces technologies et à ce que celles-ci nous font, comment elles nous affectent, ce qu'elles nous font faire quand on y regarde attentivement (par exemple, des comportements compulsifs), ainsi que l'absence d'échange sur nos (non-)usages respectifs.

Ce déficit de culture n'est pas neuf, il s'inscrit dans une histoire où nous délaissions depuis longtemps la réflexion autour de nos choix médiatiques. Développer une culture autour de l'IA nécessitera de s'interroger aussi plus largement sur les technologies de l'information et de la communication dans laquelle l'émergence de l'IA s'inscrit.

Plusieurs conditions sont nécessaires pour faire de l'IA **un sujet de réflexion commun**. Il s'agira d'abord de créer des espaces de discussion collective pour sortir d'un rapport individuel à l'IA. Il est en effet important d'en parler ensemble, car l'IA (et de manière générale, les technologies de l'information et de la communication) structurent nos manières d'échanger des informations et de penser.

Transformer les technologies de l'information et de la communication comme c'est le cas aujourd'hui (et comme cela a été le cas maintes fois dans l'histoire de l'humanité avec l'émergence de l'écriture, de l'imprimerie, des médias de masse, des moteurs de recherche, des réseaux sociaux...), c'est toucher directement à nos manières de réfléchir, d'agir et d'être en relation les unes ¹ aux autres.

Ce livret est conçu pour mettre en commun les expériences individuelles et singulières des IA, en observant comment ces expériences s'ancrent dans une histoire (personnelle et collective), en se rendant attentifs aux différences entre elles afin de se donner des repères.

Une seconde condition est nécessaire pour entrer dans la réflexion : parvenir à penser ce qui nous arrive à l'aide de termes autres que ceux qui résonnent actuellement dans l'espace

1. Ce livret a été rédigé sans appliquer la règle du masculin qui l'emporte sur le féminin. Le féminin et le masculin sont utilisés l'un pour l'autre de manière aléatoire.

public et médiatique.

Alors que les technologies de l'information et de la communication prennent tant de place dans nos vies, nous nous y intéressons en général sous un angle très particulier, celui du « bon usage », de « ce qu'il faut faire » avec les machines. Cette perspective restreinte finit par occuper tout l'espace, au détriment d'autres dimensions de l'existence de ces machines. Elle est en outre un obstacle à l'examen minutieux de **ce que nous faisons réellement avec elles**.

Nous proposons plusieurs portes d'entrée dans le livret pour essayer à chaque fois d'opérer des décalages par rapport à la manière habituelle de s'interroger sur l'IA. Ces décalages passent parfois par l'introduction de *concepts* qui servent à nous mettre en action ou à nous placer dans une certaine perspective.

Note

Réfléchir à partir de nouveaux concepts peut provoquer un sentiment d'inattendu, avec tout ce que cela peut comporter de curieux mais aussi d'inconfortable. Ces concepts sont cependant nécessaires pour entrer dans les propositions d'ateliers et de questionnement.

Le livret vise à **acquérir collectivement de la culture numérique** afin de ne pas subir ce qui s'imposerait à nous sans discussion. Ceci suppose non seulement que son usage soit discutable, mais aussi que son existence et ses conditions d'existence le soient également.

La parole des « anti-IA » est tout aussi nécessaire que celle de celles qui l'utilisent quotidiennement. Nous partirons de là où chacun en est dans sa réflexion. Aucun prérequis, connaissance technique ou usage préalable n'est nécessaire pour aborder les réflexions proposées ici. Chacune part de son expérience de l'IA, quelle qu'elle soit.

Même si nous n'avons jamais utilisé l'IA, nous en avons entendu parler et celle-ci transforme déjà le monde dans lequel nous vivons. Le livret est conçu pour que des personnes de niveaux de connaissance technique différents puissent entrer en dialogue « à égalité ». Chacun pourra s'adresser les questions posées.

Il faudra néanmoins veiller à se prémunir contre les risques d'écrasement qu'entraîne toute connaissance répartie de manière inégale au sein d'un groupe : *ce qu'on sait techniquement de nos machines ne doit jamais dicter ce qu'il faut en faire.*

Comment on le fait

Mettre des mots

Le livret n'est pas une recette à appliquer pour obtenir un résultat prédéfini. Il ne s'agit donc pas d'une méthode toute faite ou d'un dispositif *clé en main* à appliquer à *la lettre* et qui automatiquement corrigera nos cultures numériques déficientes. Bien plutôt il mobilise des moyens pour plonger dans et partir des situations réelles des personnes qui se réunissent.

Plusieurs modalités sont proposées pour y parvenir, modalités dont le trait commun est de mettre des mots sur nos expériences respectives de l'IA :

- **Décrire et observer** : Alors que les spéculations sur l'avenir foisonnent dans l'espace public, il s'agira plutôt de se concentrer sur *ce qui est déjà là*, sur ce que nous faisons *vraiment* de l'IA, ainsi que sur ce que l'IA nous fait *vraiment* en retour. Le livret vise à soutenir un effort d'observation et de description de l'environnement médiatique et numérique qui est le nôtre au moment de l'atelier.

- **Raconter** : Ce qui se passe pour nous ici et maintenant avec l'IA est le fruit de notre histoire singulière et collective passée. Raconter des anecdotes (de nous ou de nos proches) de ce qui nous est arrivé avec l'IA, des rumeurs qui circulent à son sujet, etc, sont un moyen de la contextualiser dans une histoire (personnelle et collective).
 - **Imaginer** : Si les affects sont puissants au sujet de l'IA, c'est probablement parce que nous touchons à une corde sensible de notre imaginaire collectif. La figure de la machine qui prend vie apparaît dans des productions culturelles (films, romans, dessins animés, mythes, etc.) qui structurent nos imaginaires. C'est là une manière dont nos sociétés traitent d'un enjeu anthropologique « essentiel »² ou au moins « essentialisant » : savoir distinguer ce qui est humain et ce qui ne l'est pas.
 - **Connaître** : La connaissance technique pouvant en rebouter plus d'un, nous avons porté une attention particulière à *la façon d'exposer la connaissance technique*. En effet, nous voulons éviter autant que possible l'idée selon laquelle la seule connaissance technique suffirait à combler notre déficit de culture technique et à répondre à nos interrogations. Pour cela, en premier lieu, nous nous interrogerons collectivement sur ce que *nous* souhaitons savoir de ces machines. Pour répondre à notre curiosité légitime, on trouvera, en annexe, une série de textes à caractère informatif : l'histoire des techniques, de l'IA, leur fonctionnement, etc. Le but n'est pas de fournir une connaissance exhaustive du sujet en question, mais de le contextualiser historiquement, socialement et techniquement.
- Le livret entend donc stimuler la construction et l'acquisition d'une *connaissance située*, c'est-à-dire qui part de et intègre nos intérêts et affects. Dans la même logique que le reste du livret, cette annexe peut être abordée dans n'importe quel ordre et à n'importe quel moment. On retrouvera l'annexe en ligne sous la forme d'un wiki à l'adresse suivante : <https://ai.numethic.education>.

Chacune de ces modalités d'explicitation – de « mise en mots » – est expérimentée dans des ateliers grâce auxquels un groupe peut se mettre au travail. Chaque atelier, dont les enjeux sont présentés au préalable, propose une manière (parmi d'autres possibles) de se rassembler pour se poser des questions situées (à un moment donné, avec les personnes qui sont là). Les lectrices sont invitées à s'approprier les propositions, à les retravailler pour qu'elles soient les plus adaptées à leur contexte. Des variantes des ateliers sont proposées dans cette optique.

Les ateliers sont conçus pour aider (chacun et/ou au sein d'un collectif) à **développer une culture numérique** afin d'identifier ce dont il serait important de discuter alors que cela ne nous apparaissait pas jusqu'ici. Les questions posées dans les ateliers de discussions méritent qu'on y revienne, peut-être même qu'on se les pose de manière répétée, précisément parce que les réponses ne sont pas données une fois pour toutes. On pense généralement que, quand il faut parler de quelque chose, c'est pour « régler le problème », pour « trouver la solution », en espérant qu'une discussion puisse suffire (« ça, on en a déjà parlé », comme si on pouvait alors rayer la thématique pour de bon). Or certaines questions peuvent être plus intéressantes par les problèmes qu'elles soulèvent que par les réponses apportées ponctuellement.

2. Le caractère « essentiel » de la distinction de l'homme par rapport à d'autres êtres, notamment les êtres vivants, est remis en question aujourd'hui. Pourquoi tient-on à cette distinction ? À quoi sert-elle ? Dans le contexte des crises écologiques, cette distinction entre l'homme et le reste des vivants montre plutôt ses limites. Ce qui est certain, c'est que cette distinction est « essentialisante » puisqu'elle vise à déterminer quelle est l'essence de l'homme.

À qui le livret s'adresse

À toute citoyenne

Le livret a été conçu pour rassembler des personnes autour des transformations qu'opèrent les IA dans leur existence. Cette dimension *collective* est essentielle car c'est elle qui permet d'observer les écarts entre nos pratiques, nos affects, nos habitudes, afin de les interroger et de tirer parti de ces différences. Ce qui importe au moment de constituer les groupes, c'est le désir de s'intéresser ensemble aux transformations que les IA opèrent dans nos métiers et nos vies, quel que soit notre positionnement initial par rapport à elles. Nous pouvons détester l'IA, en avoir peur, mais également avoir envie d'en discuter avec d'autres.

Les groupes peuvent être déjà constitués (par exemple, une équipe de travailleurs d'une association qui souhaiterait traiter du sujet) ou se constituer au moment de travailler autour du livret (par exemple, à l'occasion d'une animation au sein d'un Espace Public Numérique).

Afin de toucher des publics le plus largement possible, des variations sont proposées au moment de présenter les ateliers afin de pouvoir s'adapter à différents contextes³.

Qu'est-ce que ça fait de travailler avec le livret

Si ce livret n'est pas une recette, cela signifie qu'il n'y a pas un résultat connu d'avance. Les effets que nous pouvons attendre d'une culture numérique plus solide sont multiples :

- nous sentir moins écrasés par l'IA
- avoir le sentiment d'être plus libres par rapport aux discours ambients
- augmenter la capacité et l'envie d'entrer en dialogue avec des personnes qui ont une position différente de la nôtre sur l'IA
- observer avec davantage de finesse la multiplicité et la complexité des questions qui se posent dans notre environnement médiatique actuel
- transformer nos (non-)usages de l'IA
- savoir donner des raisons qui nous poussent à agir de la sorte
- transformer collectivement nos pratiques pour qu'elles servent nos intérêts ou au moins ne nous portent pas préjudice
- retrouver de la puissance d'agir collectivement
- ...

L'attention aux effets produits par le livret est importante, c'est ainsi seulement que nous pourrons mesurer ce que pourrait signifier « **éduquer au numérique** ».

Présentation des chapitres

Le livret est structuré en quatre chapitres. Bien qu'absolument **non contraignant**, l'ordre des chapitres a été pensé pour faciliter une prise en main progressive des ateliers et questionnements. En effet, les sentiments d'écrasement et de fascination nous poussent à l'urgence d'une réponse à laquelle nous essayons de résister. Ce temps de suspens est indispensable

3. Avant d'être publié, le livret a été testé avec des travailleurs de l'associatif et des collectifs constitués dans des Espaces Publics Numériques de la Fédération Wallonie Bruxelles. Il est donc en partie conçu à partir des expériences menées avec ces publics spécifiques. Qu'ils en soient remerciés.

Introduction

pour éviter de passer à côté de quelque chose d'intéressant qui nous implique et qui pourrait nous donner un levier d'action.

Le premier chapitre précise **notre imaginaire vis-à-vis des IA**, dans la culture en général et dans la culture audiovisuelle en particulier. Quelles histoires nous racontons-nous vis-à-vis des machines et autres robots ? Quelles tensions ces histoires mettent-elles en scène entre les êtres humains et les machines ? Qu'est-ce qui inquiète dans ces mises en scène ? Qu'est-ce que les robots, androïdes et autres cyborgs viennent ébranler dans nos imaginaires ?

Dans le deuxième chapitre, nous menons **une enquête sur nos pratiques médiatiques actuelles** : chercher des informations, les synthétiser, en garder une mémoire, etc. Nous prenons le temps de les observer et de les décrire, d'examiner les différences entre nos pratiques respectives. Alors que nous avons en général tendance à nous demander quelles pratiques nous « devrions » adopter, nous revenons à une dimension souvent négligée : ce que nous faisons effectivement déjà maintenant, et ce que cela nous fait. Nous discutons également de comment l'émergence des logiciels d'intelligence artificielle vient bousculer et modifier nos pratiques actuelles.

Dans le troisième chapitre, nous entrons dans **le monde des objets techniques** (les logiciels d'intelligence artificielle mais également le matériel et les infrastructures nécessaires à leur fonctionnement). De quoi ces objets techniques ont-ils besoin pour exister, pour fonctionner ? En termes de ressources naturelles ? De travail ? À quelles contraintes (chimiques, physiques, mathématiques, techniques) les concepteurs font-ils face au moment d'inventer une machine ? Ces questions servent également à observer notre rapport habituel aux objets techniques : prenons-nous le temps de nous intéresser à leur fonctionnement ? Ou au contraire, les considérons-nous en général comme des purs moyens nécessaires à la poursuite des fins qui sont les nôtres ? Et qu'est-ce que cela change de prendre le temps de s'arrêter et de nous y intéresser ? Qu'est-ce qui attise notre curiosité ? Qu'est-ce que nous gagnons à nous intéresser aux machines ?

L'annexe informative qui accompagne le livret donne certains éléments informatifs et historiques qui soutiendront notre curiosité aiguisée.

Après avoir observé le fonctionnement des machines, le dernier chapitre aborde **les valeurs véhiculées par les IA**. Ces valeurs dépendent des contraintes techniques avec lesquelles les concepteurs doivent composer, du contexte social, des choix économiques, de la politique culturelle, des systèmes juridiques, etc. Chacune réagit de manière singulière à l'émergence des logiciels d'intelligence artificielle en fonction de ce qui importe dans son existence. Ou, plus précisément, dans les différentes dimensions de son existence : vie familiale, professionnelle, loisirs, vie affective et intime, etc. Nous explorons dans ce chapitre d'une part les endroits où les logiciels d'intelligence artificielle tendent à nous emmener. D'autre part, nous examinons la manière dont l'émergence de ces logiciels nous mettent en tension les uns avec les autres, individuellement et collectivement, au sein des communautés auxquelles nous appartenons.

Chapitre 1 – Nos cultures et imaginaires de l'IA

« *Je veux être un vrai petit garçon.* » – Conte de Pinocchio

« *Tu n'es pas un vrai petit garçon.* » – Extrait du film *Artificial intelligence : AI* de Steven Spielberg sorti en 2001. La mère s'exprime à son « fils » artificiel¹.

Contexte du chapitre

« Intelligence artificielle », telle est l'expression qui résonne depuis quelque temps dans l'espace public et médiatique pour désigner de façon générique (et peut-être floue) des nouvelles technologiques qui ont émergé : grands modèles de langage, logiciels de génération d'images, de musiques et de contenus audio-visuels. Pour un technicien ou un ingénieur, ce vocable, « intelligence artificielle », est utilisé à tort et à travers et ne renvoie à aucun objet technique précis. Il a une histoire – scientifique d'abord, puis marketing – et recouvre en fait aujourd'hui un vocabulaire bien plus touffu servant à désigner un certain genre d'algorithmes : réseaux de neurones, apprentissage machine (supervisé ou pas), apprentissage profond, apprentissage par renforcement, etc.

L'expression « intelligence artificielle » focalise notre attention sur certaines questions particulières : En quoi l'intelligence humaine est-elle différente de celle de la machine ? Que serait le « propre de l'homme » que la machine ne pourrait pas remplacer ? L'intelligence artificielle va-t-elle remplacer nos emplois ? Les artistes qui ont recours à l'IA sont-ils des imposteurs (sous-entendu : parce qu'ils ne recourent pas à leurs propres ressources humaines uniquement) ? Etc. De nombreux commentaires s'emploient ainsi à insister sur *la différence de nature entre l'être humain et la machine*. Ainsi met-on en évidence soit l'impossibilité de remplacer l'être humain par la machine, soit les risques pour l'être humain d'être dépassé par la machine. Cette perspective d'une obsolescence de l'humanité est mise en scène notamment par les entreprises propriétaires de logiciels d'intelligence artificielle. Brandissant en effet des scénarios proches des récits de science-fiction, voire de mythes, elles revisitent un imaginaire de fin du monde : le risque d'un retournement de la créature (une « super-intelligence ») sur son créateur qui serait incapable de la contrôler, la prédiction de la fin de l'ère humaine (certains se risquant même à préciser l'année où celle-ci aura lieu)². Ces idées focalisent notre attention sur la distinction entre l'être humain et la machine, ainsi que sur les potentiels risques encourus. Notre réflexion tend dès lors à se porter exclusivement sur ces questions au détriment de bien d'autres.

1. Au XXIIe siècle, alors que des robots androïdes répondent aux besoins des êtres humains, des parents adoptent un « enfant artificiel » dans le cadre d'une expérimentation d'ingénieur en intelligence artificielle : est-il possible de créer des robots capables d'aimer ? Leur fils biologique est plongé dans un coma artificiel au moment de l'adoption. Les parents ne savent pas s'il pourra s'en sortir un jour. Après la convalescence (inespérée) du fils biologique, des conflits éclatent entre les deux « frères ». La mère s'exprime ainsi à son fils artificiel peu de temps avant de s'en séparer.

2. Par exemple Sam Altman, patron d'OpenAI, en mai 2023, qui co-signe une lettre parlant de la possibilité de l'extinction humaine causée par l'IA.

Avant de porter notre attention sur d'autres questions dans les chapitres suivants, peut-être un bon point de départ est-il d'observer les discours ambiants sur l'IA de l'intérieur, par une plongée attentive dans cet imaginaire collectif où ils vont puiser. Le mythe, la fiction peuvent ici enrichir les idées présentées par les discours : les récits s'adressent à nous d'une autre manière qu'un discours qui nous expose des idées. Ils parlent à notre existence plutôt qu'à notre connaissance. Ils provoquent des émotions, ont la capacité de résonner avec nos expériences et nos histoires singulières et peut-être avec des éléments profondément partagés de ce que c'est que d'exister pour un être humain.

Cette perspective permet d'aborder de nouvelles questions :

Pourquoi l'imaginaire de ces récits de fin du monde (ou du moins de l'humanité) ont un tel succès aujourd'hui ?

Qu'est-ce qu'ils éveillent en nous ?

Pourquoi aimons-nous (ou n'aimons-nous pas) ces histoires ?

Qu'est-ce qui fait qu'on se sent menacés en particulier par des machines qui visent à nous remplacer dans ces récits ?

Et finalement, qu'est-ce que nous pouvons apprendre de nous-mêmes et de notre époque à partir de ces observations ?

Pour ce faire, nous repartons de l'imaginaire, nous prenons le temps de nous y attarder, de ne pas le quitter trop vite.³

Genèse du transhumanisme : le choix d'un mythe

Pour mieux saisir d'où viennent ces discours sur l'IA, de quelles histoires elles sont les fruits, nous allons donc voir dans quels imaginaires ils vont puiser et de quoi sont faits ces imaginaires. Des histoires de création de créatures artificielles existent depuis bien longtemps, bien avant de voir apparaître des robots, des androïdes ou des grands modèles de langage.

Ainsi, le mythe du Golem⁴ apparaît dans les *Psaumes* dans des versets qui présentent la création comme le fait de Dieu. Ce mythe est saisissant car il mobilise déjà des caractéristiques similaires aux discours actuels sur l'IA, notamment dans les discours transhumanistes.

Le Golem est une créature d'argile qui prend vie à partir des lettres dessinées sur son front ou en introduisant un parchemin dans sa bouche⁵ qui reproduisent le nom de Dieu (dans le judaïsme, la combinaison des lettres hébraïques est à l'origine de la création du monde, le monde entier est fait par les lettres). Ce sont en général des rabbins et des savants qui connaissent les rituels et les protocoles qui permettent de créer des golems. La création est un acte de vénération de Dieu, c'est un geste de prière. Réitérer l'acte de création permet de se rapprocher de Dieu, c'est expérimenter le mystère de la création divine. C'est aussi un acte de magie de donner vie à une créature surnaturelle. Créer un golem demande une pureté d'âme, une pratique mystique qui n'est pas accessible au commun des mortels. Les rabbins et les savants s'essayent à créer une créature la plus parfaite possible (à l'image de Dieu qui crée l'homme). Le Golem est dès lors un être perfectible, jamais inachevé, toujours en perfectionnement.

Mais le Golem est un être ambivalent. En tant que créature, il n'est pas doué de raison, ni de conscience, il n'est ni tout à fait vivant ni tout à fait mort. Il n'a pas le même statut que

3. Voir François Flahaut, *La pensée des contes*, Éd. Economica (2001). Disponible en pdf en ligne sur le site de l'auteur.

4. Voir *Golem, Avatars d'une légende d'argile*, Catalogue de l'exposition publié par le Musée d'art et d'histoire du Judaïsme de Paris en 2017.

5. Un épisode des Simpson met en scène ce personnage et une des manières dont il prend vie. Voir l'émission BiTS sur ARTE, disponible sur Youtube, *Le Golem : une créature entre la religion et Terminator*

l'homme. C'est un être dont on fait ce qu'on veut. On peut le détruire, le jeter voire le tuer de la même manière qu'on jette une machine. L'analogie avec la création d'une machine s'accentue au cours de l'histoire. « À partir du XV^e siècle, le Golem devient [...] un serviteur, un « aide-technique », qui libère son créateur des travaux serviles jusqu'à pouvoir échapper à tout contrôle⁶. »

La différence de statut entre l'homme et le Golem est rejouée à la Renaissance. Alors que les indiens acquièrent le statut de l'homme et les droits qui y sont associés, la question de savoir si le Golem a une âme est posée. Ces créatures peuvent-elles être aussi dotées d'une âme ?

Avec le siècle des Lumières, l'homme s'émancipe de la toute-puissance de son créateur. Ce n'est plus Dieu qui fixe le destin de l'homme, il exerce son libre-arbitre et sa volonté. Une nouvelle question se pose alors avec plus d'intensité : Quel est le rapport de la créature à son créateur ? Celui-ci en est-il responsable ? Le mythe pourrait bien également mettre en scène l'hubris, la déraison et les excès de l'homme dans ses tentatives d'emprise absolue sur la nature. Selon Charles Mopsik (1956-2003), spécialiste de la Kabbale et de la mystique juive, « le Golem, et c'est là son principal danger, ne fait que refléter et exacerber la dimension mécanique, déterminée de l'être humain, qui se perçoit comme un être pensant, libre, oubliant à quel point il est en réalité lui-même un être limité et conditionné⁷. » « Ce n'est pas tant le Golem qui est créé à l'image de l'homme que l'homme qui se révèle être un golem. »

Au XIX^e siècle, apparaît la version du mythe la plus connue aujourd'hui, celle de la création d'un golem par le Maharal de Prague (un personnage légendaire inspiré d'un rabbin ayant réellement existé au XVI^e siècle). Comme une machine, ce golem l'assiste dans des travaux pénibles, mais il protège également la communauté juive de la ville. Il la défend contre les fausses accusations dont elle est victime et des violences (réelles) qu'elle subit. Dans la plupart des versions de la légende du Maharal de Prague, le Golem est un être colossal et surnaturel. Un jour où le Maharal avait oublié de le désactiver, alors que celui-ci est en maîtrise de sa créature, le Golem lui échappe et commence à tout dévaster⁸.

Le Golem connaîtra d'innombrables descendants – ou du moins des parents imaginaires plus ou moins proches – qui rompent avec la tradition juive. Il entre en résonance notamment avec le roman de Frankenstein écrit par Mary Shelley au XIX^e siècle. Le roman reprend la trame du mythe avec la création d'une créature (à partir de morceaux de cadavres) qui se retourne contre son créateur, Victor Frankenstein. Comme le Golem, la créature de Frankenstein n'est ni morte ni vivante. Par contre, la question de la transgression d'un interdit (la créature est créée de nuit avec des cadavres) et de l'incapacité du créateur à endosser la responsabilité de son acte n'apparaissent pas dans le mythe du Golem. La magie est licite dans le judaïsme tant qu'elle est pratiquée pour accomplir l'œuvre de Dieu⁹, il n'est pas question de transgression dans le mythe. Ainsi, contrairement à sa fonction dans l'imaginaire juif, la reprise du mythe dans l'univers chrétien signifie un *acte de défi lancé à Dieu*¹⁰, c'est un acte sacrilège. À cet égard, le récit de Mary Shelley se rapproche davantage du mythe de Prométhée (c'est d'ailleurs le sous-titre du récit, *Le Prométhée moderne*). Alors que les dieux de la Grèce Antique ont créé un homme vulnérable, Prométhée transgresse les lois des dieux de l'Olympe et dérobe le feu aux dieux. En ce sens, il est un sauveur pour la communauté des hommes. Il s'agit bien d'un Prométhée moderne dans l'histoire de Mary

6. Ibid, p. 115

7. Voir *Golem, Avatars d'une légende d'argile*, Catalogue de l'exposition publié par le Musée d'art et d'histoire du Judaïsme de Paris en 2017, p.24

8. Émission BiTS sur ARTE, disponible sur Youtube, *Le Golem : une créature entre la religion et Terminator*.

9. Le judaïsme considère que chacun a besoin de magie, de ce « petit plus » surnaturel dans nos existences humaines, mais seules certaines personnes peuvent la pratiquer. Voir l'émission « *Frankenstein ! Bienvenue dans le monde des créatures artificielles* » (Épisode 3 / 5 : *Le Golem*) diffusée sur France Culture en 2016.

10. Émission BiTS sur ARTE, disponible sur Youtube, *Le Golem : une créature entre la religion et Terminator*.

Shelley. Victor Frankenstein se désintéresse des pratiques mystiques telles que l'alchimie sur les conseils de ses professeurs à l'université. Il se consacre à la science et c'est celle-ci qui lui permet de parvenir à créer une créature.

Parmi les innombrables résonances de l'histoire du Golem avec d'autres récits, nous pouvons citer également les super-héros qui apparaissent dans les bandes-dessinées du début du XX^e siècle aux États-Unis tels que Batman et Superman¹¹. Ces histoires de super-héros empruntent au mythe certains traits du Golem. Ils ont des pouvoirs surnaturels ou hors-normes, défendent une communauté et sont mus par un idéal de justice. Les auteurs, des immigrés d'origine juive, emportent dans leur migration le mythe du Golem.

À l'instar des hypothétiques super-intelligences artificielles, les variantes du mythe mettent en récit des *apparitions de créatures sans recourir à la procréation*, en outrepassant les déterminations biologiques des êtres humains. Comme les IA, le Golem arbore un double visage, tantôt une figure de l'espoir (Golem sauveur, protecteur et promesses de nouvelles prouesses médicales pour l'IA), tantôt un monstre¹², tantôt un idiot (le Golem n'est pas capable de parole et l'IA est, après tout, une « bête » machine), tantôt un être surpuissant. Tous ces récits font ainsi écho à l'apparition d'une « entité surpuissante, artificiellement créée par l'homme pour faciliter son existence et assurer sa protection, mais capable parfois d'échapper à son contrôle, voire de se retourner contre lui¹³ ».

La mise en perspective historique du mythe du Golem entre en résonance avec les discours actuels autour de la créature de l'intelligence artificielle : Quel est le statut de l'intelligence artificielle ? Peut-elle acquérir le même statut que celui de l'être humain ? A-t-elle une conscience d'elle-même ? Une « âme » ? Quel est le rapport du créateur à sa créature ? En est-il responsable ? Elle met aussi en exergue l'ambivalence du Golem, sauveur et protecteur de l'homme, mais aux pouvoirs dévastateurs dont les transhumanistes ne cessent de nous mettre en garde. L'intelligence artificielle va-t-elle se retourner contre son créateur ? Est-elle capable de le dépasser ?

La cybernétique ou le traitement identique de l'humain et de la machine

Les sciences et les techniques ne sont pas hermétiques à ces imaginaires collectifs ; au contraire, ils se nourrissent réciproquement. Cette proximité est particulièrement visible avec la cybernétique. L'expression « intelligence artificielle » apparaît avec la cybernétique dans les années 1950 (voir le wiki). Cette science interdisciplinaire étudie la communication et la circulation de l'information *dans les systèmes naturels comme dans les systèmes artificiels*. Les uns comme les autres réagissent aux informations qui leur parviennent de l'environnement et sont capables de s'adapter en fonction des informations qu'elles ont traité antérieurement.

Cette nouvelle science est placée sous le signe du Golem par Norbert Wiener lui-même, un de ses pionniers, dans son ouvrage *God & Golem, inc*¹⁴. Ce qui deviendra plus tard l'ordinateur (en particulier les logiciels) est pour Wiener « l'homologue moderne du Golem du rabbin de Prague »¹⁵ : alors que le Golem est animé par des lettres hébraïques, l'ordinateur l'est par

11. Ibid.

12. Nicolas Nova reprend également l'image du monstre pour qualifier l'IA dans son livre *Persistance du merveilleux*, Éd. Premier Parrallèle (2025).

13. *Golem, Avatars d'une légende d'argile*, Catalogue de l'exposition publié par le Musée d'art et d'histoire du Judaïsme de Paris en 2017, p.23

14. Voir *Golem, Avatars d'une légende d'argile*, Catalogue de l'exposition publié par le Musée d'art et d'histoire du Judaïsme de Paris en 2017, p.114

15. Ibid, p.114

un système de 0 et de 1 ; alors que le Golem est fait d'argile, les puces électroniques sont faites de silicium.

L'analogie de l'homme et de la machine permet aux cybernéticiens de modéliser l'apprentissage. Celui-ci consiste pour l'homme comme pour la machine à s'adapter à partir des signaux (des « informations ») qui leur parviennent de leurs environnements respectifs. La plasticité de la machine (comme du cerveau humain) lui permet d'évoluer à partir des réponses qu'elle apporte aux signaux antérieurs. Cette analogie n'est pas seulement théorique. Les cybernéticiens vont essayer de créer des machines capables de pouvoir le faire effectivement. Ce sont les prémisses des algorithmes dits « d'apprentissage machine ». Contrairement aux machines mécaniques dont le fonctionnement est uniquement déterminé par son créateur, ceux-ci sont capables de s'adapter à partir des perceptions qu'il reçoivent de leur environnement.

Pour les cybernéticiens, les ordinateurs fournissent une image opérante de ce qu'est l'être humain : même si celui-ci n'est pas « qu'un robot », ce que les cybernéticiens ne manquent pas de souligner, il est comme l'ordinateur *un être d'information*.

Le Golem ordinateur n'est plus l'œuvre mystique d'un rabbin au cœur pur, c'est la figure du scientifique qui s'impose. Les scientifiques et les ingénieurs connaissent les protocoles pour faire passer le Golem à l'existence. Celui-ci bénéficie dès lors de l'autorité de la science.

Un contexte idéologique

Il est aussi intéressant de noter que cette nouvelle version du mythe du Golem naît dans *l'après seconde guerre mondiale*. La confiance dans l'espèce humaine est ébranlée, la faillite de l'humanité est éclatante, les dangers de la science sont apparus de manière évidente. C'est la représentation de l'homme comme un être responsable, qui exerce son libre arbitre qui est ébranlée. Il ne suffit plus d'invoquer un golem pour protéger l'humanité contre les dérives sanguinaires de ses dirigeants (comme le Golem de Prague qui protégeait les Juifs contre les violences et les accusations dont ils étaient victimes), il s'agira désormais de remodeler l'être humain¹⁶ en un golem qui sera plus fiable que lui.

Alors que l'homme du XIX^e siècle défie Dieu, l'homme du XXI^e siècle a le sentiment de ne pas être à la hauteur par rapport aux machines qu'il conçoit. Il est dépassé par la puissance et la qualité des productions des machines, ainsi que par la vitesse à laquelle les technologies se renouvellent. Les technologies ne nous simplifient plus la vie, mais nous dépossèdent de nos savoir-faire (pensons aux grands modèles de langage et à notre capacité à écrire). Dans ce contexte, l'homme inachevé est en recherche perpétuelle de perfectionnement, d'amélioration pour atteindre son plein potentiel.

Ce long parcours par les variantes du mythe du Golem met en évidence la parenté entre ce mythe et notre imaginaire collectif actuel qui s'incarne dans d'innombrables discours (récit, littérature, films, discours techniques et scientifiques, voire politiques¹⁷). Les histoires que nous nous racontons sur les machines et nos affects sont aussi pris dans ces imaginaires mythologiques. Si le prisme du Golem nous permet d'observer une manière de raconter le monde et ce qui se passe aujourd'hui avec l'IA, c'est que ce mythe véhicule des idées, des valeurs et des croyances, qui peuvent servir différents objectifs, intérêts et horizons : maintenir un ordre social, poser le problème de notre rapport à la machine, renforcer une idéologie dominante, la critiquer, servir Dieu, défier Dieu, etc.

16. Ibid, p.117

17. Pour des exemples, voir *Golem, Avatars d'une légende d'argile*, Catalogue de l'exposition publié par le Musée d'art et d'histoire du Judaïsme de Paris en 2017, pp.22-23.

Le mythe du Golem n'épuise pas la dimension mythique des récits du transhumanisme actuel. Nous pourrions nous intéresser à d'autres mythes. Pygmalion¹⁸ permet d'observer le thème de la machine qui prend les traits d'une femme fatale¹⁹. Le rêve d'Icare quant à lui met en scène l'hubris de l'homme grisé par ce que la technique lui permet de réaliser.

Au contraire du mythe du Golem où l'homme est face à une machine extérieure à lui, les transhumanistes imaginent également des hybridations entre l'homme et la machine. Il est alors davantage question de l'homme augmenté et du cyborg. Mais ce qui compte ici, c'est de mieux saisir d'où viennent les discours des personnes qui vantent les mérites de l'intelligence artificielle, de quelle mythologie ils s'inspirent (quelle qu'elle soit), et surtout, ce que cela nous fait d'être exposées à ces discours.

Atelier 1 : Sonder nos imaginaires sur l'IA

Enjeux

Afin d'observer nos imaginaires concernant les machines, en particulier des robots, androïdes et autres cyborgs, nous nous concentrerons dans ce premier chapitre sur des productions culturelles populaires qui les mettent en scène. Nous essayons de mettre en évidence les tensions imaginaires qui structurent ces récits et dans lesquelles nous sommes pris. Qu'est-ce que ces récits nous font ? Comment est-ce qu'ils nous affectent ?

Ce passage par la fiction permet de faire un pas de côté vis-à-vis des discours ambients sur l'intelligence artificielle. Il s'agit d'aller au-delà des idées, du message de leurs auteurs et de ce qu'ils suscitent, de sortir du registre de la connaissance. Si ces récits résonnent avec nos existences aujourd'hui, c'est qu'ils évoquent quelque chose qui nous travaille, qui nous fait réagir, qui nous affecte. Nous cherchons dans cet atelier à mettre collectivement des mots sur ce que ces récits nous renvoient de nous-mêmes et de la société dans laquelle ils s'inscrivent.

Présentation de l'atelier et consignes

Premier temps : Visionner un film ou un court-métrage

Quelques propositions de films et court-métrages qui mettent en scène des « intelligences artificielles » (robots, androïdes, logiciels...) :

- [Blade Runner](#)
- [Matrix](#)
- [A.I. Intelligence artificielle](#)
- [Ex machina](#)
- [Her](#)
- [2001, L'odyssée de l'espace](#)
- [Wall-E](#)

18. Mythe de la Grèce antique où un homme, Pygmalion, tombe amoureux d'une statue. Celle-ci finit par prendre vie et à s'unir à lui.

19. Dans la plupart des versions du Golem, les créatures sont de sexe indéterminé. Alors que la statue de Pygmalion a un corps parfait et idéal, celui du Golem est inachevé et il n'est pas l'objet de désir. Il existe néanmoins des versions du mythe où un golem de sexe féminin au physique parfait est créé. Pour en savoir plus, voir *Golem, Avatars d'une légende d'argile*, Catalogue de l'exposition publié par le Musée d'art et d'histoire du Judaïsme de Paris en 2017, p.25.

- Terminator
- Black Mirror
- Zima Blue
- L'apprenti sorcier
- Frankenstein (il existe beaucoup d'adaptations)
- Sayônara

Deuxième temps : Discussion collective

À partir de ce que les participantes ont vu, des questions permettent d'engager une discussion collective :

- Aimons-nous ce récit ? Quels moments, scènes avons-nous particulièrement aimé ?
- Au contraire, quels moments du récit nous ont déplu ?
- Quels sentiments le récit éveille-t-il en nous (peur, dégoût, fascination, vertige...) ?
- Ces sentiments sont-ils partagés au sein du groupe ?
- Qu'est-ce que nous n'avons pas compris ?

Plus les mots qui qualifient les sentiments (éventuellement contrastés à différents moments de l'histoire) sont précis pour chacune, plus des différences peuvent apparaître. Est-ce que ce qui horrifie l'une ravit l'autre ? Discuter de nos sentiments permet d'observer comment ces récits entrent en écho avec nos existences singulières. Ce sont les différences entre eux qui viennent enrichir l'image de ce que le film produit dans notre imaginaire collectif. Cette première étape est essentielle, c'est elle qui permet d'essayer de mettre en discussion ce que nous ressentons depuis nos expériences propres.

Les participants pourront ensuite s'interroger sur l'histoire en elle-même :

- Qui est le créature ? Apparaît-il dans le récit ?
- Qu'est-ce qui est créé ?
- Dans quel contexte la créature apparaît-elle ?
- **Quelle(s) tension(s) le créateur cherche-t-il à résoudre au moment de sa création ?**
- **Quelle(s) tension(s) l'usagère de l'IA cherche à résoudre ?**
- Y parvient-elle ?
- **Quelle(s) nouvelles tensions apparaissent après la création (pour la créatrice, avec l'usager) ? Avec quels effets sur les personnages ?**
- Quelle relation le créateur entretient-il avec sa créature ?
- Comment la création est-elle mise en scène ?
- Quelles relations la créature entretient-elle avec d'autres êtres humains que son créateur ? Avec d'autres machines ?
- Y a-t-il des ressemblances avec d'autres récits de fiction que nous connaissons ? Les-quel(s) ? Quels différences voyons-nous entre les choix des autrices ?

Pour répondre à ces questions collectivement, il est toujours intéressant de revenir au film pour percevoir ensemble ce qui permet à quelqu'un de dire ce qu'il dit. Lors du visionnage d'un film, nous ne sommes pas attentives aux mêmes éléments, de sorte que l'interprétation que nous construisons est différente. Revenir au film permet d'observer les raisons qui ont conduit à cette interprétation de l'œuvre.

Variations possibles

1. Afin d'interroger nos sentiments par rapport à une histoire, les participantes peuvent également visionner le début d'un film dont elles ne connaissent pas la fin. Le visionnage s'arrête avant la fin et les participants s'interrogent sur la fin qu'ils souhaiteraient voir, celle qui leur ferait plaisir. Est-elle partagée au sein du sous-groupe ? Fait-elle débat ? Quels sont les éléments du début du film qui permettent aux participantes de l'imaginer ? Elles pourront se demander de la même manière quelle fin elles ne souhaiteraient surtout pas voir.
2. Il est possible de discuter à partir d'un film que les participants connaissent déjà. Ils font alors appel à leur mémoire pour en discuter collectivement.

Debriefing

1. Les participantes reviennent aux discours sur l'intelligence artificielle (par exemple, les prises de parole de transhumanistes). De la même manière que pour les films, elles s'interrogent sur les sentiments qu'ils suscitent : Aiment-elles ces discours ? Ou contraire, ceux-ci leur déplaisent-ils ? Quels sentiments ces discours éveillent-t-ils en elles (peur, dégoût, fascination, vertige...) ? Sont-ils partagés au sein du groupe ? Qu'est-ce qui est incompréhensible dans ces discours ?
Comme pour les films, il s'agit de demeurer un moment sur cette question des affects. Ceux-ci peuvent être puissants, les participants gagnent à en faire l'objet d'une discussion collective afin d'observer ce qui se passe pour chacun d'eux.
2. À partir de l'analyse du film et des discours transhumanistes, il est peut-être possible de comprendre un peu mieux ce qui nous arrive aujourd'hui avec l'émergence de l'intelligence artificielle. Qu'est-ce que ces récits peuvent nous apprendre de nous-mêmes ?
Les participantes y répondront d'abord pour elles-mêmes, à partir de leurs expériences singulières au sein du groupe qui se réunit. Elles peuvent ensuite essayer de se poser la question de manière plus générale. Qu'est-ce que ces récits peuvent nous apprendre de notre époque ? Pourquoi ont-ils un tel succès aujourd'hui ? Avons-nous des hypothèses ?

Chapitre 2 – Se repérer par rapport à l'IA dans nos pratiques médiatiques quotidiennes

Contexte du chapitre

L'IA comme médiation

Depuis l'émergence des technologies numériques, le mot « outil » s'est imposé pour désigner les logiciels et le matériel informatique que nous utilisons. Ce mot entraîne une certaine manière d'envisager notre rapport aux technologies qui nous entourent :

- à la manière d'un marteau ou d'une scie, l'outil est *extérieur* à nous et nous entretenons rapport *individuel* avec lui ;
- il est déjà là dans une *forme figée* ;
- il nous faut apprendre à l'utiliser tel qu'il l'exige, nous y faire, transformer nos gestes pour nous l'approprier.

L'outil produit donc avant tout un *rapport d'usager* devant apprendre les bonnes pratiques (pratiques qui sont d'ailleurs souvent dictées de l'extérieur voire transmises par des personnes « expertes »). D'autres rapports possibles comme celui du *bricoleur* ou du « *hacker* » sont marginalisés, tandis que ceux de la *conceptrice* et encore plus de la *technicienne* deviennent peu visibles. En outre, un outil, ça se perfectionne. Aussi le discours ambiant est-il souvent accompagné de l'idéal du *progrès* et de l'*innovation*. Ce sont ces valeurs qui dynamisent la marche en avant des nouvelles technologies et cherchent à légitimer le fait qu'elles s'imposent à tous et à chacune.

L'intelligence artificielle ne déroge pas à cette conception. Depuis son émergence, les formations à l'IA se focalisent sur l'apprentissage du « bon usage » de cette nouvelle technologie, sans interroger son existence, sans chercher à la transformer d'une manière qui pourrait peut-être nous convenir, sans questionner l'idéal de progrès que l'intelligence artificielle véhicule.

Nous pouvons pointer un premier problème de ce discours : il nous focalise toujours sur « ce qu'il faudrait faire », ce qui nous empêche d'observer ce que nous faisons réellement, ce que nous faisons déjà. Or nous faisons énormément puisque les technologies de la communication et de l'information sont devenues omniprésentes dans notre quotidien. Elles nous accompagnent désormais dans des activités fondamentales de notre existence : communiquer, produire des textes et des images, nous informer sur les actualités, acquérir de nouvelles connaissances, etc. Dans ce chapitre, nous prendrons le temps d'observer et de décrire ces activités : Comment est-ce que nous nous y prenons au fond ? Quelles pratiques diverses sont les nôtres actuellement ? Pourquoi opérons-nous ces choix-là plutôt que d'autres ? Et, surtout, comment l'intelligence artificielle vient-elle bouleverser – ou non – nos pratiques existantes ?

Nous pourrions nous interroger dans la solitude de l'usager individuel de l'outil « intelligence artificielle ». Mais pour pouvoir comprendre ce qui se passe au moment de l'émergence des

logiciels d'intelligence artificielle (et de manière plus générale des technologies de l'information et de la communication), il faudra prêter attention à deux dimensions supplémentaires : 1) ce qui se passe pour les autres (non-)usagères, 2) les impacts sur les relations entre les personnes. En effet, les technologies de l'information et de la communication opèrent une *médiation*. En ce sens, elles mettent en relation, de différentes façons, des êtres humains parfois proches, parfois lointains, parfois conscients de cette opération, parfois à leur insu. Ainsi, par exemple, les grands modèles de langage sont des médiateurs entre les auteurs des textes qui servent d'entraînement et l'usager du logiciel. Mais le logiciel *Whatsapp* sert aussi de médiation dans les relations entretenues par les usagères. Ces dimensions doivent être pensées conjointement pour avoir une image la plus juste possible des modifications qu'entraîne l'émergence des logiciels d'intelligence artificielle.

Le milieu de l'IA

Pour pouvoir observer de plus près ces médiations et les conditions matérielles dans lesquelles elles s'opèrent, il est intéressant de penser au *milieu* de l'intelligence artificielle, de la situer quelque part. Ce milieu est double : il y a un milieu *physique* et un milieu *symbolique*.

a) Un milieu physique : les logiciels d'intelligence artificielle ont besoin d'un milieu physique favorable pour se déployer. De la même manière qu'il nous serait impossible de discuter en l'absence d'air, nous ne pouvons pas utiliser un logiciel de génération d'images sans la présence de serveurs, d'une capacité de calcul suffisamment puissante, de l'extraction minière des composants nécessaires aux machines, des infrastructures électriques, du réseau internet, etc. La liste est longue. Comme la crise environnementale nous le rappelle, ces technologies ne pourraient pas exister sans ce milieu physique, ni sans une organisation du travail, de conception et d'entretien pour répondre à leurs besoins.

b) Un milieu symbolique : les technologies de l'information et de la communication organisent notre manière de nous informer et de communiquer depuis toujours. Que ce soit le langage parlé, l'écriture, l'imprimerie, les chiffres ou les images, chacune de ses manières d'« informer » et de « communiquer » impose ses propres contraintes. Alors que l'écriture permet de garder une trace et favorise un travail architectural de la pensée par l'organisation grâce aux titres et sous-titres, le langage parlé fait travailler la mémoire et permet la répétition ou l'insistance par des inflexions de la voix (il est plus musical qu'architectural), etc. Les transformations des milieux symboliques ont des répercussions fortes sur nos existences, c'est-à-dire qu'elles ont un impact structurel sur notre société : une justice basée sur des textes de loi favorise les personnes qui comprennent ces textes et ces modalités (écrites), mais défavorise les analphabètes en les condamnant à se fier à autrui. Les logiciels d'intelligence artificielle transforment donc également notre manière d'échanger des signes et des symboles (textes, images, sons et musiques, nombres). Les contenus qu'ils produisent sont porteurs de significations pour les êtres humains et agissent donc aussi sur les milieux symboliques.

Tentons de voir ce que cela change d'observer les transformations, voire les bouleversements, liés à l'émergence de l'intelligence artificielle en se rendant attentifs aux *médiations* qu'elle opère et aux *milieux* dans lesquelles elle se déploie plutôt qu'en se plaçant dans la perspective d'un outil que nous prenons en main :

- Si nous restons dans la perspective de l'outil, nous pourrions croire que l'usager seul est impacté par l'émergence de l'intelligence artificielle. Celui-ci doit apprendre de nouvelles pratiques et faire avec les nouvelles contraintes que son outil lui impose. Or, si c'est le milieu médiatique qui est transformé, les bouleversements touchent tout le monde, usagère ou pas, en maîtrise ou non de ce nouvel outil. Par exemple, si la majorité des travailleurs de l'associatif recourent aux grands modèles de langage pour

rédiger des rapports aux pouvoirs subsidants, la demande future pourrait peut-être se faire plus exigeante et les rapports plus intéressants ; si la majorité des personnes recourent au même logiciel pour choisir le cadeau à offrir à une amie à l'occasion de son anniversaire, la diversité des cadeaux à acheter et à offrir, voire ce que recevoir un cadeau signifie, risquent également d'être impactés. Prenons un exemple d'une transformation passée, celle opérée par le smartphone – adopté aujourd'hui par la grande majorité de la population. L'émergence de cette technologie a touché des dimensions de la vie de chacun, usager ou non, que nous le voulions ou pas et a créé une nouvelle norme sociale : acheter des tickets de train par application plutôt qu'au guichet ou à la machine, accéder à son compte en banque via une application, communiquer avec ses proches en visioconférence plutôt qu'au téléphone, etc. Ce sont finalement nos habitudes et nos normes sociales communes qui se sont transformées. Et pas toujours de manière homogène, que ce soit entre classes d'âges (enfants, adolescents, adultes et personnes âgées) ou classes sociales (personnes dans la précarité ou porteur d'un handicap visuel, technicienne, etc). Cela peut même créer de nouvelles communautés, par exemples celles qui refusent la transformation de leur milieu, qui rejettent le smartphone dans ce cas-ci. C'est donc la question des fractures sociales potentielles qui échappe à la perspective de l'outil. D'autant plus dans un contexte où les tensions entre groupes sociaux ne font qu'augmenter.

- La maîtrise de l'usage de l'outil qui reviendrait à la volonté de chacun des usagers (une fois en maîtrise d'un générateur d'images, l'usager obtiendrait ce qu'il veut du moment qu'il sait ce qu'il veut faire et ce que la machine peut réaliser) s'oppose ici à l'imprévisibilité des transformations d'un milieu, pour les usagers comme pour les conceptrices. Celles-ci ne peuvent pas prévoir quelles transformations des rapports sociaux les nouvelles technologies vont opérer. Un concepteur ne peut pas facilement prévoir quelle nouvelle technologie va s'imposer et l'histoire recèle nombre d'exemples où des technologies ont échoué. Par exemple, le Metavers de Meta (anciennement Facebook) n'a jamais pu percer chez le grand public et ce malgré une campagne marketing longue et très coûteuse. Dans la perspective de l'outil, la question qui se pose est de savoir ce que nous voulons faire ou pas (parce que c'est possible de le prédire). La perspective du milieu déplace la question : qu'est-ce qui nous échappe et qu'est-ce qui a échappé à la conceptrice ? Quelles seront et ne seront pas ou plus les nouvelles pratiques dans le nouveau milieu transformé ? Ces questions permettent de se rendre attentifs de manière plus complète aux effets réels des nouvelles technologies sur nos existences.
- Alors que l'outil s'impose d'un coup au moment de son apparition, comme s'il était déjà là, sans histoire (nous avons un marteau dans la main), la transformation du milieu est progressive, elle peut être tantôt brutale (elle nous affecte alors fortement), tantôt discrète. L'émergence des grands modèles de langage est progressive et nous pourrions même supposer qu'elle n'est pas terminée car elle n'est pas stabilisée (par exemple, ces technologies ne sont pas rentables pour les entreprises qui les déploient, le modèle économique n'est pas encore établi...). Si nous considérons l'émergence de ces nouvelles technologies comme une transformation de notre milieu, nous pourrons être attentifs aux changements qui ne sont pas toujours faciles à observer : ils peuvent nous échapper, parce que nous ne parvenons pas à les cerner, parce que des petites transformations permanentes sont réalisées par les entreprises qui les déploient, parce qu'ils sont complexes et que leurs causes sont multiples, parce qu'ils ne nous touchent pas tous de la même manière.

Nous pouvons prendre une analogie avec la transformation des textes de loi¹ : même si

1. L'analogie ne tient pas si nous nous intéressons à qui décide des transformations du milieu. Laisser les décisions aux mains d'un état démocratique ou d'une entreprise marchande qui ne dépend pas de la souveraineté

de nouvelles applications de lois dans l'enseignement touchent seulement une frange de la population à court terme (les enseignantes et les travailleurs du milieu scolaire ainsi que les élèves), les conséquences concernent le corps social dans son ensemble. L'éducation d'aujourd'hui n'est plus celle d'hier, parce que les règles que nous nous donnons collectivement ont changé. Les changements sont tantôt opérés de manière brutale, tantôt discrets au point qu'ils pourraient passer inaperçus. Il faudra du temps et du recul pour les observer, avec des effets qui peuvent s'avérer bien plus profonds que ce que nous voyons apparaître au départ (d'autant plus si nous nous donnons initialement des indicateurs qui passent à côté de certains effets que personne n'avait prévus). Les transformations opérées ne sont pas totalement prévisibles ni pour le gouvernement, ni pour les citoyennes.

Transformer les médias, comme les textes de loi, c'est transformer nos conditions d'existence. Cela revient à transformer notre cadre de vie, physique et symbolique. Ce chapitre vise à nous situer collectivement dans ces milieux. Nous l'écrivons au pluriel, car ils ne sont pas toujours partagés par tout le monde ; certains milieux sont partagés par certaines communautés et nous appartenons à plusieurs d'entre elles avec des pratiques potentiellement différentes pour chacune d'elles (nous ne communiquons pas de la même manière avec notre famille, nos amis, notre compagnon ou nos collègues). Nous gagnons à les observer de près, dans le détail, pour pouvoir observer ces différences de pratiques entre nous. L'émergence des logiciels d'« intelligence artificielle » n'aura pas les mêmes effets en fonction du milieu dont nous sommes issus.

Atelier 1 - Petites histoires de nos usages

Enjeux

L'intelligence artificielle (et les technologies de l'information et de la communication en général) peut être mobilisée dans différentes dimensions de nos existences que nous dessinons ici à gros traits : le travail (synthétiser des informations nécessaires à la rédaction d'un appel à projets, produire un résumé pour des participants à une formation, créer une affiche ou un site internet) - les désirs, la vie affective et intime (demander conseil pour offrir un cadeau à une amie, pour langer un nouveau-né, explorer les causes possibles d'un symptôme) - la vie publique (écrire une lettre pour des pouvoirs publics, obtenir un résumé d'un texte de loi...). Cet atelier d'écriture vise à articuler une histoire, une anecdote personnelle à une question que celle-ci nous pose. Si nous nous intéressons à l'intelligence artificielle, c'est qu'elle nous touche dans une – ou plusieurs – dimensions de notre existence. Se poser une question, c'est mettre en évidence ce sur quoi nous désirons nous attarder, ce sur quoi nous souhaiterions réfléchir collectivement. C'est aussi une manière de concevoir collectivement un mini-programme de recherche à partir des questions qui émergent.

Présentation de l'atelier et consignes

Chaque participante écrit un petit texte pendant une quinzaine de minute à partir des consignes suivantes :

1. Identifier une pratique précise de l'IA que vous avez déjà expérimentée, que ce soit dans votre travail, dans votre vie affective, dans la vie publique ou dans vos loisirs.

de l'Etat qui gouverne pose d'autres questions que nous aborderons au chapitre 4.

2. Décrire le contexte dans lequel elle intervient : à quelles occasions l'avez-vous utilisée ? Était-ce seul ou à plusieurs ? Avec qui le logiciel vous a-t-il mis en lien ?
3. Interrogez-vous sur les motifs de votre choix de l'utiliser à ce moment-là (quelle est votre intention ?), de cette façon-là (quel est le dispositif technique utilisé ? Avez-vous utilisé un smartphone, une tablette, un ordinateur ? Comment avez-vous choisi le logiciel ?). Une autre manière de vous adresser cette question serait de vous demander si vous auriez fait quelque chose dans le même contexte si l'IA n'existe pas (l'IA vous permet-elle de faire de nouvelles choses ou de le faire d'une façon que vous jugez plus riche ?).
4. Terminer le texte par une question.

Au moment d'écrire, les participantes essaient de raconter une histoire qui met en évidence le milieu dans lequel l'expérience a lieu (quels sont les acteurs de l'histoire ? Quelles sont leurs habitudes ?...).

Variations possibles

1. Les consignes de cet atelier d'écriture suppose que les personnes qui y répondent ont déjà eu une pratique des logiciels d'intelligence artificielle. Ce n'est pas toujours le cas et l'expérience peut être trop pauvre pour mettre en évidence quelque chose qui compte pour la personne qui y répond. Si une participante n'a pas de pratique d'IA, il est possible d'écrire une histoire qu'on se raconte et qui circule (comme une rumeur) à son sujet et qui a marqué, quelque chose qu'on a entendu ou qu'on a lu et qui nous a troublés. Cela peut également être une anecdote racontée par une proche. Les participants terminent leur texte par une question.
2. Cet atelier peut être proposé comme un temps particulier du deuxième atelier de ce chapitre, au moment d'entamer le Temps 3 en sous-groupe (voir ci-dessous).

Debriefing

Les textes sont mis en commun en favorisant les prises de parole qui étoffent les questions posées par l'émergence des logiciels d'intelligence artificielle. Au moment de la lecture, deux rôles sont confiés à des participantes volontaires :

1. Produire une synthèse des problématiques soulevées par les textes lus : quelles sont les différentes facettes des questions posées par l'intelligence artificielle qui apparaissent dans les textes ? La thématique de l'intelligence artificielle est riche, il s'agira ici de préciser les dimensions qui *importent* pour le groupe qui travaille autour du livret.
2. Relever quelques éléments qui permettent de mettre en évidence des différences entre les milieux médiatiques qui apparaissent dans les récits. Avec quoi et avec qui les logiciels d'intelligence artificielle mettent-ils en relation leurs usagers ? Quelles sont les médiations qu'ils opèrent ?

Atelier 2 - Enquête sur nos pratiques médiatiques

Enjeux

Alors que nous nous concentrons souvent sur les pratiques médiatiques que nous devrions adopter, nous nous attardons en général très peu sur l'observation et la description fine de nos pratiques actuelles. Cet atelier vise à mener une enquête sur notre milieu médiatique à partir d'une liste de questions (non-exhaustive) qui servent avant tout à préciser les multiples dimensions possibles lors de ce travail d'observation. Dans ce livret, nous avons choisi de nous concentrer sur nos manières de nous *informer* : elles ne doivent pas dicter ce qu'il convient d'observer, mais plutôt indiquer l'esprit du travail d'observation proposé. L'atelier est individuel dans un premier temps afin de se concentrer chacune sur la description de ses pratiques singulières. Une mise en commun a lieu dans un second temps de sorte à mettre en évidence les ressemblances et différences de pratiques : partageons-nous le même milieu ? À quelles occasions ne le partageons-nous pas ?

Présentation de l'atelier et consignes

Dans un premier temps, nous menons une enquête personnelle sur la manière dont nous nous informons. Nous indiquons ci-dessous quelques sous-questions pour essayer de caractériser le(s) milieu(x) médiatique(s) que nous arpentons pour rechercher des informations. Ces questions servent à nourrir votre enquête, elles ne sont pas exhaustives et peuvent être reformulées pour coller davantage à votre contexte.

Temps 1 : Enquête individuelle

- Dans nos métiers, quels sont les médias que nous consultons ? Et pourquoi en choisissons-nous certains et pas d'autres ? (radio, TV, réseaux sociaux, livres, e-books, podcast, plateformes numériques, via des moteurs de recherche, flux RSS, blogs...). Nous pouvons détailler les médias choisis (quelle radio et quelles émissions ? Dans quel contexte l'écoutons-nous ? Sur quelles machines ?...) et les raisons qui nous poussent à opérer ces choix.
- À quels types d'informations (textes, images, audio/vidéo, nombres) sommes-nous exposés ? Quelle est la longueur des articles / livres / films /... que nous consultons ? Y a-t-il des formes d'expressions que nous privilégions (par exemple le texte ou l'audio-visuel) ? Est-ce la même chose pour notre travail et nos loisirs ? Certains types d'informations sont-ils réservés au travail ou aux loisirs ?
- Le média est-il consulté d'une traite ? Devons-nous y revenir à plusieurs reprises ? Quand les consultons-nous ? À quelle fréquence revenons-nous au même média ? Avons-nous des pratiques de veille (consulter des médias régulièrement afin d'être au courant de ce que ces médias publient, éventuellement sur des thématiques particulières) ? Lesquelles ? Pourquoi ?
- Quels types d'information retenons-nous des médias que nous consultons ? Les cherchons-nous volontairement (par exemple, choisir un podcast pour travailler sur un sujet précis) ? Quelles stratégies avons-nous développer pour nous en souvenir ?
- Quand nous recherchons une information, préférons-nous regarder une vidéo, lire un texte, écouter un podcast pour trouver une information ? Est-ce variable en fonction de l'information recherchée ?

- La présence d'un comité éditorial permet de savoir quel genre d'informations nous trouverons dans un journal (le comité éditorial de *Médor* ne suit pas les mêmes lignes éditorial que *Le Figaro*, ce que les lecteurs savent en achetant leur journal). Avec la transformation des médias, les pratiques éditoriales se transforment. Sur les réseaux sociaux, il n'y a pas de ligne éditoriale officielle, mais il existe cependant une modération des contenus (certains pouvant être supprimés), et une hiérarchisation des informations sur base de critères qui ne sont pas toujours connus des usagères. Ces nouvelles pratiques médiatiques posent de nouvelles questions : Y a-t-il une modération des contenus (par un algorithme, par un comité de rédaction ou par la communauté) dans les médias que nous consultons ? Faisons-nous attention à la manière dont les contenus sont modérés ? Choisissons-nous les médias en fonction de la manière dont les contenus sont modérés ? À partir de quels critères avons-nous confiance dans les contenus que nous consultons (choix du média, de l'auteur, etc) ? Est-il aisément vérifiable la fiabilité des informations ?
- Des lois encadrent ce qui peut être fait d'une œuvre ou de la production de données. En particulier le droit d'auteur et le RGPD. Quand et comment ses contraintes affectent notre travail et nos loisirs ? Par exemple, choisissons-nous une banque d'image libre de droits ou préférons-nous payer aux ayants droit pour utiliser ou réutiliser une œuvre ? Quand nous voulons organiser une activité dans un cadre professionnel ou privé avec d'autres participantes, et que nous avons besoin de certaines informations personnelles, comment cela affecte-t-il la mise en place de notre activité ?
- Qu'est-ce que nous faisons des informations que nous collectons ? Comment gardons-nous une mémoire des médias consultés ? Pouvons-nous facilement retrouver les informations que nous citons ? Essayons-nous d'être fidèle ? La transformons-nous (volontairement) ? Sommes-nous en mesure de citer nos sources ? Opérons-nous des synthèses entre différentes sources ?
- Quel est le rapport au temps que nous cultivons au moment de nous informer ? Par exemple : une concentration profonde où nous ne voyons pas le temps passé ; lire pour tuer le temps, parce que nous nous ennuyons. Comment qualifier le régime d'attention dans lequel nous nous trouvons à ce moment-là ?
- ...
- **À quelles tensions, à quelles difficultés faisons-nous face dans vos pratiques actuelles ? Qu'est-ce qui est insatisfaisant ?**

Temps 2 : Par petits groupes de 3-4 personnes (premier temps)

Les réponses à ces questions seront mises en commun, en particulier la dernière, afin d'observer comment ces milieux informationnels sont partagés (ou non).

Temps 3 : Discussion collective² Chaque petit groupe relève quelques éléments de la discussion qui a précédé. En fonction de ce qui est mis en commun, la discussion en grand groupe permettra d'observer comment l'émergence des algorithmes d'intelligence artificielle transforme notre milieu. Les participants (usagers ou non³) sont-ils affectés par cette transformation de leur milieu ?

2. Les pratiques de discussion philosophique collective que PhiloCité a l'habitude d'animer sont présentées dans le livre de PhiloCité, *Philosopher par le dialogue*, Éd. Vrin (2020). Ces méthodes sont également présentées sur le blog Dans la grotte de PhiloCité (www.philocite.eu/blog).

3. L'atelier 1 de ce chapitre peut également servir d'introduction à ce moment de l'atelier 2.

Nous vivons un moment particulier de l'histoire de ces algorithmes : ils sont nouveaux et nous pouvons nous rendre davantage sensibles aux transformations qu'ils opèrent alors que celles-ci seront difficiles à détecter une fois que leur présence dans notre milieu sera normalisée. Alors que les discours marketing nous avertissent que les intelligences artificielles sont partout, nous tâcherons d'observer précisément ce qu'il en est depuis notre perspective singulière. Voici quelques exemples de questions qui pourront servir de point de départ à la discussion :

- Où se trouvent les intelligences artificielles dans nos pratiques médiatiques ?
- Leur présence est-elle manifeste ? Visible ? Uniquement accessible à des techniciennes qui connaissent les algorithmes ? Quand les avons-nous repérées pour la première fois ?
- À partir d'un exemple concret (par exemple : la première entrée d'un moteur de recherche qui est un texte généré par un grand modèle de langage) : Comment cela nous affecte-t-il ? Qu'est-ce que cela nous a fait la première fois que nous l'avons vu apparaître ? Est-ce que cela nous a attiré ?
- Plus largement : Si nous assistons à une transformation de notre milieu, quelles sont nos réactions à ces transformations : qu'est-ce que cela nous fait ?
- Avons-nous déjà utilisé volontairement une intelligence artificielle ? Dans quel but ? À quelles occasions ? Aurions-nous cherché la même information par un autre moyen auparavant ? Si oui, comment nous y serions-nous pris ?
- Comment les intelligences artificielles transforment-elles actuellement nos pratiques médiatiques ? Quelles sont les nouvelles pratiques qui apparaissent ou disparaissent (pour nous-mêmes ou pour d'autres) ? Voyons-nous plus ou moins d'informations sur tel ou tel sujet ? Est-ce qu'il est plus difficile ou plus facile d'accéder à certaines informations (par exemple des sujets de pointe ou des textes didactiques) ?
- Les IA apportent-elles des solutions à des problèmes que nous avions déjà observés ? De nouvelles difficultés apparaissent-elles ?
- Comment réagissons-nous à l'absence d'autrice identifiable derrière les contenus générés ?

Variations possibles

1. Nos manières de communiquer de nous informer sont imbriquées, il peut être artificiel de vouloir les séparer. Il est inutile voire contre-productif de vouloir les séparer à tout prix. Si nous avons choisi ici de nous concentrer sur nos manières de nous informer, c'est surtout pour avoir un point de départ circonscrit pour entamer la réflexion.
2. L'enquête sur nos pratiques médiatiques gagne à être menée sur un temps un peu plus long (quelques jours voire quelques semaines), par exemple à la manière d'un journal à tenir quotidiennement, afin d'être au plus proche des pratiques réelles. Ces observations peuvent être rapportées plus tard dans le collectif. Les effets produits par les médias sur nous ne sont pas toujours faciles à percevoir à moins de se donner des moyens d'y être attentives. Ce journal est à mi-chemin entre un journal intime qui augmente notre conscience des événements de la journée et de leur retentissement affectif sur nous (en particulier notre attention se portera aux difficultés et aux tensions que nous rencontrons afin d'observer ce qui nous pose problème déjà actuellement) et un relevé plus technique et plus objectif de nos usages, comme un médecin nous demanderait de relever ce qu'on mange précisément et à quelle heure pour identifier le régime dans lequel on se trouve.
3. Si ce sont des problèmes particuliers qui nous occupent (par exemple : la désinformation, l'esprit critique, l'économie de l'attention...), les questions pourront être

précisées pour les traiter davantage. Par exemple, savoir comment nous choisissons les sources et les autrices que nous lisons, comment nous pouvons avoir confiance en ces ressources, permet d'amorcer l'observation de la façon dont nous nous y prenons pour exercer notre esprit critique, mais aussi de préciser peut-être ce qui nous pose problème pour l'exercer déjà aujourd'hui.

Debriefing

Les effets que nous pouvons attendre de ce travail d'observation et de description sont de différentes natures. Ils pourront être précisés au moment du debriefing par chacune des participantes. Parmi les effets possibles, en voici quelques uns :

- Peut-être avons-nous déjà l'habitude de discuter collectivement de nos pratiques médiatiques ; peut-être est-ce nouveau. Dans les deux cas, notre manière de mettre des mots sur nos expériences, de les raconter, pourra être transformée. La façon dont nous parlons de l'intelligence artificielle avec nos proches pourrait également être altérée.
- En général, l'émergence des logiciels d'intelligence artificielle ne laisse pas indifférent. La description collective de nos pratiques médiatiques actuelles pourrait transformer nos affects. Par exemple, nous pourrions observer que d'autres personnes que nous éprouvent des difficultés similaires dans leurs pratiques médiatiques et réfléchir ensemble à la manière de les transformer.
- Même si les médias contraignent nos modes de communication et d'information (nous ne pouvons pas faire ce que nous voulons), les usagers (individuellement et collectivement) ont toujours développé des stratégies d'appropriation voire de détournement pour s'en emparer d'une manière qui leur convient davantage⁴. La mutualisation de la description des pratiques médiatiques permet de s'inspirer les unes les autres des pratiques actuelles voire d'inventer ensemble de nouvelles manières de faire.

4. Cette idée est dépliée pour les usages du smartphone dans la thèse de Nicolas Nova, *Smartphones, Une enquête anthropologique*, Métispresses (2020).

Chapitre 3 – Culture technique : entrer dans le monde des machines

Contexte du chapitre

Pourquoi devrions-nous nous intéresser au fonctionnement des logiciels d'intelligence artificielle ? Ne suffit-il pas d'apprendre à les utiliser ? Qu'aurions-nous à gagner en prenant le temps de savoir mettre des mots sur ce qui se passe dans nos machines ? D'entrer dans leur monde ?

La culture technique

D'après le philosophe Gilbert Simondon, notre *déficit de culture technique*, c'est-à-dire le fait de ne pas nous intéresser aux machines, de ne pas en faire un objet de réflexion (que ce soit en famille, à l'école, au sein d'une équipe de travail, dans notre vie affective, etc.), mène à *deux types de relations* aux objets techniques qui sont toutes les deux aliénantes. D'une part, nous sommes amenés à traiter les machines en *esclaves*. Nous nous concentrons alors sur l'usage de l'objet technique et nous ignorons délibérément son fonctionnement, son histoire, ses contraintes... et donc là où il nous entraîne. Notre ignorance nous pousse en effet à devenir les esclaves de nos esclaves, les esclaves de nos machines. Dépendance dont on se rend parfaitement compte lorsque la machine ne fonctionne plus. D'autre part et en même temps, une *sacralisation* des machines est à l'œuvre. Nous les mettons sur un piédestal et plaçons une confiance aveugle en elles, aveugles aux contraintes techniques et à la réalité physique. Nous pensons qu'il suffit d'y investir temps et argent pour qu'à la fin nous soyons récompensés par et grâce au « progrès technique ».

Il est important de considérer que les objets techniques ont un *mode d'existence propre*, différent de celui des êtres vivants. Ils ont leurs propres contraintes, leurs propres temporalités, un développement historique qui se distinguent de la réalité humaine. Les contraintes internes des objets techniques définissent non seulement les capacités de l'objet en lui-même, mais également la manière dont celui-ci va interagir avec son milieu¹ (humain ou non). Le *fonctionnement* de l'objet technique permet de mettre en évidence ces contraintes internes. Par exemple, un moteur thermique fonctionne par la succession d'explosions contrôlées provoquant une expansion de gaz et actionnant grâce à cela un mouvement de piston. Son fonctionnement spécifique impose donc des contraintes elles aussi spécifiques. Il faut maîtriser :

- les explosions du gaz en expansion et donc connaître la thermodynamique sous-jacente,
- la durée et le rythme des explosions successives,
- la résistance des matériaux pour faire face à la violence des explosions,
- ...

1. Voir [Chapitre 2 – Se repérer par rapport à l'IA dans nos pratiques médiatiques quotidiennes](#) pour une compréhension plus précise de la notion de « milieu ».

Or il existe des *tensions* entre ces différents éléments. En effet, a) augmenter la puissance des explosions demande une plus grande maîtrise des matériaux, b) augmenter la fréquence des explosions demande une plus grande maîtrise des explosions, c) augmenter la résistance des matériaux, en les épaisissant, demande plus de puissance pour mettre le tout en mouvement. Pour un moteur, ces tensions sont à la source de la création de plusieurs **architectures de pistons**. En fonction du contexte, du besoin (technique, économique,...) de nombreuses possibilités techniques ont été inventées, du simple monocylindre, facile à mettre en place, à produire et à réparer, à une disposition complexe en étoile pour les avions ou encore aux classiques dispositions en V.

Ce sont les choix effectués pour gérer ces tensions *objectives* qui sont au cœur du développement dynamique de la technique en question. Ce n'est donc pas, comme on pourrait le croire de prime abord, l'idée d'une inventrice qui se matérialise dans la matière, mais bien plutôt un *compromis* entre la réalité physique, mathématique et la volonté d'un ou plusieurs techniciens évoluant dans un contexte social particulier. *C'est de ce compromis qu'émerge une culture technique*. La culture technique n'est donc pas simplement une connaissance de l'objet technique ou des sciences, c'est une pratique dans un contexte social et historique situé, un regard sur notre réalité par le prisme de l'objet technique. Avoir une culture technique, c'est être capable de construire une relation réciproque avec les objets techniques, c'est-à-dire de tenir compte des réalités humaines et techniques et d'en construire une synthèse. L'on se donne alors la possibilité de comprendre ensemble deux choses : 1) ce qui est mobilisé socialement dans la conception des machines, 2) les contraintes internes à la machine. En d'autres termes, on voit de quel compromis est issue la machine, quelle synthèse elle incarne. C'est également cette capacité qui permet d'imaginer des transformations de la machine qui soient possibles compte tenu des réalités humaines et techniques avec lesquelles la machine est forcée de composer.

La culture technique spécifique à l'IA

Afin de remédier au manque de culture technique relative aux objets qui nous concernent ici, penchons-nous sur le fonctionnement interne d'un logiciel d'intelligence artificielle. Pour comprendre leur fonctionnement interne, il faut d'abord désigner d'un point de vue technique quel logiciel nous voulons étudier (grands modèles de langage, reconnaissance faciale, chemin le plus rapide avec GPS, etc). Nous allons nous concentrer dans un premier temps sur les **grands modèles de langage** (Large Language Models ou LLM) qui permettent, entre autres, une discussion comme agents conversationnels à l'instar de ChatGPT. Ces programmes informatiques ont pour but de produire la suite **probable** d'un texte, ce qui leur permet de répondre à une question qui leur est posée. Le fonctionnement repose sur l'établissement d'un modèle dit « statistique », c'est-à-dire qu'à partir de textes qui ont été collectés en très grande quantité, une relation statistique va être établie entre les mots qui se suivent dans ces textes. Sont ainsi repérés les mots qui ont le plus de chance de suivre. Par exemple, il est très probable qu'un verbe suive un sujet ou encore que, dans un texte parlant d'un personnage situé dans un verger et devant manger, le personnage mange un fruit du verger. Pour établir ces relations **statistiques**, les objets mathématiques utilisés sont des **réseaux de neurones profonds** (dans le cas des grands modèles de langage). Ce sont ceux qui donnent actuellement le résultat le plus rapide et le meilleur. La **rapidité** et la **qualité** des résultats sont justement deux contraintes majeures dans l'établissement du modèle, et ce n'est que depuis 2017 qu'une astuce technique, mathématique, a permis de concilier les deux (en tout cas suffisamment pour être acceptable à l'usage). Nous pouvons déduire de ce fonctionnement quelques contraintes internes aux grands modèles de langage :

- Une grande quantité de textes de qualité pour établir un modèle qui puisse mettre en évidence les relations statistiques qui se trouvent majoritairement dans notre langage afin d'avoir une réponse adéquate,
- Un algorithme statistique qui permet de faire la synthèse de la multiplicité des relations,
- Un algorithme statistique qui permet à partir du modèle (de la synthèse) de produire facilement et rapidement la réponse,
- ...

Toutes ces contraintes entrent en tension les unes avec les autres. Et, à nouveau, ce sera à la technicienne, à l'inventeur, de gérer ces tensions en opérant une série de choix en fonction de ce qui est physiquement possible et culturellement valorisé pour essayer d'arriver à un objet technique satisfaisant.

Ce chapitre vise à entrer dans le monde des machines, à s'y attarder pour connaître un peu plus ces réalités qui composent avec nous au quotidien, tout autant que nous composons avec elles.

Atelier 1 : Représenter le fonctionnement d'une IA²

Enjeux : attiser la curiosité et étoffer la relation

Que nous soyons usagères des logiciels d'intelligence artificielle ou pas, il est utile de se demander quelles *relations* nous entretenons avec les objets techniques et ce qui est nécessaire à leur fonctionnement. Cet atelier vise à prendre le temps de s'arrêter sur le fonctionnement des machines afin d'observer la distance que nous entretenons avec leur monde, leur fonctionnement, leurs besoins en termes de travail humain ou de ressources naturelles. Ce moment d'arrêt est nécessaire pour commencer à développer notre culture technique. A-t-on envie de plonger dans ce monde ? A-t-on des résistances à le faire ? Ce monde nous paraît-il étranger ? Familiar ?

Travailler collectivement est primordial. Rappelons en effet qu'il ne s'agit pas de corriger individuellement sa connaissance des machines, mais plutôt de renforcer notre curiosité, d'augmenter notre envie d'en savoir plus, sachant que celle-ci n'existe pas forcément pour chacun au départ.

Plonger dans le monde des objets techniques permet également d'observer la *complexité des relations* que l'objet technique tisse avec celles (et ce) qui composent son milieu : les êtres humains qui le maintiennent ou qui le développent, les usagères, les employés des sociétés commerciales, celles qui les fabriquent et qui travaillent sur les matières premières, mais aussi les autres objets techniques avec lesquelles il interagit, le monde physique qui l'entoure, etc. Quand ces relations sont visibles, nous pouvons nous demander comment les objets techniques agissent dans notre société : « acteurs culturels » car ils transforment nos imaginaires ; « acteurs économiques » puisqu'ils transforment les marchés et les chaînes de production et le travail ; « acteurs sociaux » car ils transforment nos relations sociales, etc.

À la manière de la rencontre avec un étranger dont nous ne sommes familiers ni du langage ni des coutumes et des modes de vie, il s'agira pour entrer en relation avec l'objet technique de s'intéresser à *ce qui compte pour lui*. Travailler sur le fonctionnement nous permettra de préciser ce qui fait que nous sommes fascinées ou pas par la machine, d'observer comment

2. Cet atelier est inspiré du livre *Enfantillages outillés, Un atelier sur la machine*, de Fanny Béguery et Adrien Malcor, Éd. L'Arachnéen (2016).

l’existence d’une machine particulière nous affecte. Si nous sommes plus familiers avec la machine, nous serons capables de la traiter (un peu) moins comme une étrangère, nous pourrons mieux comprendre ce qui se joue entre elles et nous. Le monde des machines n’en sera pas désenchanté pour autant ; au contraire, nous pourrons observer avec plus d’acuité comment le merveilleux est à l’œuvre encore aujourd’hui dans nos machines³.

Présentation de l’atelier et consignes

Redisons-le : *travailler collectivement est primordial*. Le regard de chacune enrichit le regard qu’initialement chacune portait. Et chacun est une ressource pour les autres.

Les participantes à l’atelier travaillent en sous-groupes de 3 ou 4 personnes (avec des niveaux de connaissance différents). Chaque sous-groupe a du matériel de dessin (marqueurs et crayons) ainsi qu’une grande feuille de dessin (A3 ou A2).

1. Chaque sous-groupe choisit une situation vécue par l’un d’entre eux en relation avec un logiciel d’intelligence artificielle dans un contexte précis (par exemple, écrire un prompt pour générer une image à partir du logiciel Midjourney disponible en ligne). Les participantes représentent par un dessin ou un schéma le fonctionnement de l’objet technique choisi dans ce contexte précis ainsi que son milieu technique. Chaque sous-groupe peut ainsi examiner :
 - les infrastructures et les machines nécessaires à son fonctionnement (réseau internet, borne wifi, réseau électrique, etc.),
 - la dimension matérielle de l’utilisation (ordinateur, smartphone, objets connectés, etc.),
 - les logiciels (observer différents logiciels pour effectuer une même tâche),
 - les matériaux nécessaires à la fabrication et au fonctionnement des intelligences artificielles.

Puisqu’il s’agit entre autres d’un *atelier d’observation*, les participants peuvent observer le logiciel, les infrastructures physiques de la pièce (voire du bâtiment) dans laquelle ils se trouvent.

Si les participantes ne savent pas comment le logiciel fonctionne (à un moment donné de l’atelier, ce sera le cas, y compris pour les techniciens puisque les entreprises qui détiennent ces logiciels ne dévoilent pas précisément comment ils s’y prennent), elles imaginent comment cela pourrait fonctionner.

2. Une fois le fonctionnement de l’objet technique représenté, les participantes indiquent sur le dessin (par exemple à l’aide de pictogrammes) où se trouve le *travail humain*. Le travail humain peut prendre des formes multiples que les participantes peuvent détailler : création et conception de logiciels, modération des contenus générés, entretien du matériel, des logiciels et des infrastructures, marketing, etc. Quelle part du travail est visible ? Qu’est-ce qu’on ne parvient pas à se représenter ?
3. Au moment de représenter le fonctionnement (ou à la suite), les participants consignent les questions qu’ils se posent. Ces questions viennent appuyer ce qui les rend curieux, ce qu’ils voudraient savoir qu’ils ne connaissent pas encore. Il n’est pas question à ce stade de l’atelier de chercher à y répondre, ceci viendra plus tard, au moment du debriefing.
4. À l’issue de l’atelier, les participantes se demandent s’il reste du merveilleux dans la machine, quelque chose qui leur paraît curieux, étonnant, admirable. Ce merveilleux

3. Voir Nicolas Nova, *Persistance du merveilleux*, Éd. Premier Parallèle (2025).

est indiqué sur le dessin à l'aide d'un pictogramme (par exemple des étoiles ou des licornes).

Variations possibles

1. Ne pas être usager de l'intelligence artificielle ne constitue pas un frein à l'atelier. En effet, celui-ci vise à cultiver une familiarité avec le monde des machines. Il est possible de réaliser le même atelier avec un objet technique nécessaire au fonctionnement d'une IA (par exemple, l'usage d'un navigateur internet pour consulter un site). Les sous-groupes peuvent également se diviser le travail pour examiner différents objets techniques nécessaires au fonctionnement des logiciels d'IA. Il est ainsi possible de pousser l'atelier d'observation du matériel encore plus loin en démontant des (vieilles) machines et en allant examiner ce qui se passe à l'intérieur (les nouvelles machines de par leur miniaturisation tendent à rendre de plus en plus opaque leur fonctionnement).
2. Il peut arriver que le déficit de connaissances du fonctionnement des machines – et de la capacité à les observer – soit tellement important qu'il soit difficile pour certains sous-groupes de se lancer. Dépasser l'inhibition que peut susciter notre ignorance des machines fait partie intégrante de l'exercice. L'imagination peut parfois être la seule piste d'exploration possible. Le travail sur une description de ce que fait l'IA à l'aide d'une métaphore est une autre manière de répondre à la question. On gagnera alors à filer la métaphore pour qu'elle produise des questions sur le fonctionnement de l'IA.
3. Le travail de questionnement sur l'objet technique peut venir à la suite du travail d'observation et de manière individuelle.⁴ Il s'agit alors, à la suite de l'observation, de faire une liste de questions (par exemple 25, en tout cas un nombre important), sans chercher à y répondre directement et en se demandant ce qui nous rend curieuse et ce que nous aurions envie de savoir. Le grand nombre de questions permet d'épuiser celles qui sont évidentes et voir apparaître des questions singulières, originales.
4. Il est possible de se concentrer sur des questions *écologiques* à partir de cet atelier. Le dessin peut alors être réalisé en essayant de cartographier où se situent les besoins en ressources naturelles (ressources minières, ressources en électricité, en eau, etc.). Ces besoins peuvent être directement liés au fonctionnement de l'objet technique étudié (par exemple : le métal des ordinateurs) ou des infrastructures nécessaires (les câbles électriques pour acheminer l'électricité, elle-même générée par des centrales nucléaires, à gaz, à charbon, etc.). L'atelier peut devenir vertigineux quand l'observation porte sur les besoins en matière des objets techniques nécessaires indirectement (par exemple : quelles sont les ressources nécessaires pour les machines des usines qui fabriquent des machines extrayant les minéraux qui sont utilisés dans la fabrication des ordinateurs). Cette modalité de l'atelier permet de mettre en évidence la complexité des connaissances, pratiques et des moyens techniques nécessaires à l'existence des logiciels d'intelligence artificielle.
5. Une autre manière de travailler le merveilleux dans nos machines consiste à observer certains mots utilisés dans le jargon. Si nous suivons l'hypothèse de l'anthropologue Nicolas Nova, le vocabulaire de l'informatique indique précisément que quelque chose de l'ordre du merveilleux persiste dans nos machines (avatar, cheval de Troie, virus et anti-virus, ver, bug, hallucinations, trolls...). Les participants peuvent s'attarder aux choix des mots : les connaissent-ils ? Existe-t-il un synonyme plus technique ? Quel mot pourrait être inventé pour mettre davantage en évidence le merveilleux de la machine tel qu'il est perçu par les participantes ?

4. Voir Nicolas Nova, *Exercices d'observation*, Éd. Premier Parallèle, 2022 – Exercice 2 : Série de questions

Debriefing

1. Il est intéressant de revenir sur les *affects* de départ des participantes et sur leur évolution au cours de l'atelier : quelle a été la dynamique collective au sein de chaque sous-groupe ? Y a-t-il eu des résistances de certains à entrer dans l'atelier ? Celles-ci se sont-elles transformées au cours de l'atelier ? À travers l'imaginaire des autres participantes, la curiosité est-elle apparue ? S'est-elle amplifiée ? Est-ce que les participants se sentent moins (ou plus !) écrasés face à la complexité des objets techniques étudiés (qui sont précisément des techniques de pointe) ?
2. Les différents dessins sont exposés et les participantes les présentent en mettant en évidence ce que l'atelier leur a donné envie de connaître, ainsi que les questions qu'elles se sont posées. Quelles sont les questions qu'ils se posent ? Sont-elles originales par rapport à celles qu'elles s'étaient posées jusqu'ici ?
3. Un temps est consacré à la recherche de réponses – ensemble avec un petit programme de recherche collectif ou individuellement en fonction des intérêts singuliers. Il ne s'agit pas de tout savoir sur les logiciels d'IA, comme si c'était une condition nécessaire à leur usage, mais plutôt d'avoir une clarté sur ce chacun des participants désire connaître au moment de l'atelier. Le wiki qui accompagne le livret donne déjà quelques contenus à consulter. Au moment du debriefing, il peut être important de rappeler que ce qui est visé, c'est 1) de comprendre les contraintes (techniques et humaines) avec lesquelles l'objet technique doit composer et 2) d'observer la complexité des relations qui tissent son milieu. Le merveilleux de la machine ne devrait pas disparaître à l'issue de l'atelier, mais il se sera peut-être un peu transformé avec la culture technique (la connaissance, mais aussi la capacité de mettre ensemble des mots sur le fonctionnement, de s'attarder à l'observation des machines, etc.). Il n'y a pas de « bon » chemin pour mieux connaître le monde des objets techniques, c'est aux participantes de faire ensemble leur chemin pour mieux comprendre ce qui les intéresse. Une spécialiste peut également intervenir à ce stade de l'atelier pour donner des réponses aux questions. Sa curiosité à l'égard des objets techniques peut encore nourrir celle des participants.
4. Le temps est limité, il s'agira dès lors de faire des choix sur ce que les participantes veulent observer. Au terme de l'atelier, un temps peut être pris pour se demander collectivement ce qu'il reste à observer et ce que chacun aurait envie d'observer de plus près.

Atelier 2 – Langage et statistiques

Enjeux de l'atelier

Les statistiques sont aujourd'hui omniprésentes dans nos existences : prévisions météorologiques, études médicales sur les vaccins, GPS et gestion du trafic, algorithmes de recommandation pour le ciblage publicitaire, création d'indicateurs pour la prise de décision des États tels la « croissance », le « chômage », les « déficits budgétaires », les « sondages électoraux », les « taux de natalité », etc. Elles sont devenues indispensables à l'administration et la gestion des États modernes⁵, les décisions étant prises sur la base des statistiques disponibles.

Les fonctionnements des logiciels d'intelligence artificielle reposent eux aussi sur des calculs statistiques de grande ampleur. La puissance de ces techniques mathématiques utilisées

5. Voir Olivier Martin, *Chiffre*. Éd. anamosa (2023) Voir également Alain Supiot, *La gouvernance par les nombres*, Éd. Fayard (2015).

aujourd’hui massivement ne doit pas nous faire oublier que nous avons toutes une certaine familiarité avec ce que manipulent les statistiques : l’aléatoire et le *hasard*. Quand nous jouons au dé, par exemple au poker, nous avons bien une petite idée des risques que nous prenons en faisant une enchère. Quand nous observons le ciel, même si nous ne sommes pas certains de la météo des prochaines heures, nous connaissons les tendances possibles. Tout n’est pas possible... ou du moins certaines choses sont largement improbables.

Nous nous appuyons sur cette familiarité avec l’aléatoire pour présenter une *approche intuitive* des *Les statistiques* dans le *wiki* qui accompagne le livret. Celle-ci permet de comprendre les bases des *contraintes mathématiques* avec lesquelles les grands modèles de langage doivent composer. Les problèmes inhérents à ces modèles (par exemple, « hallucinations », textes qui paraissent plats, manque de fiabilité des résultats, etc.) n’apparaissent dès lors plus comme des faiblesses d’un modèle qui vient d’être conçu, mais plutôt comme le produit de contraintes mathématiques.

Les statistiques sont des objets techniques particuliers. Si nous poursuivons l’hypothèse qui parcourt ce livret, même si nous avons une intuition de l’aléatoire, nous pourrions tout de même bien souffrir d’un déficit de culture statistique. Une fois de plus, nous ne cherchons pas ici à apprendre à utiliser les mathématiques et à réaliser des calculs statistiques, mais plutôt à les considérer comme des actrices de notre société, qui ont une histoire et qui nous emmène quelque part.

Présentation de l’atelier et consignes

Une présentation intuitive des statistiques peut être introduite comme préalable à l’atelier (voir le *wiki* qui accompagne le livret).

1. Les joueuses s’installent par table de 3 idéalement (pour jouer avec des dés à 6 faces). Un dé de 6 faces se trouve sur chaque table. Le début d’un texte est exposé au regard de chacune. Ce texte est suffisamment long pour donner quelques éléments de contexte⁶. Le principe du jeu consiste à générer la suite du texte à partir d’un lancer de dé.
2. Avant de commencer la partie, les numéros du dé sont attribués (par exemple : 1 et 2 pour la première joueuse, 3 et 4 pour la deuxième, 5 et 6 pour le troisième). Les joueurs s’organisent également pour les lancers de dé et l’écriture du texte généré qui est consigné au fur et à mesure du jeu par l’une d’eux sur une feuille de papier.
3. Chaque joueuse pense à un mot qu’elle ne révèle pas aux autres ; celui-ci doit avoir du sens par rapport au texte qui précède. Le texte généré doit donc être crédible. Le dé est lancé et le joueur qui est désigné par le dé révèle son mot et celui-ci est noté (les autres joueuses peuvent révéler leur mot également). Les joueurs réfléchissent au mot suivant qui pourrait suivre et un nouveau lancer de dé a lieu, et ainsi de suite jusqu’à la fin du temps imparti (de 10 à 15 minutes en fonction du rythme du jeu). Notons que la ponctuation compte comme un mot. Afin de coller au fonctionnement des grands modèles de langage, des classes de mots peuvent être considérées comme équivalentes : « le/la/l’ », « ce/c’ », « ne/n’ », « il/elle », « ils/elles », « ton/ta »...
4. À l’issue du jeu, l’ensemble des textes sont lus à voix haute. Les participantes s’interrogent alors si ces textes ont du sens, si elles détectent des éléments qui paraissent improbables voire impossibles.

6. Voici un exemple de texte qui peut servir de point de départ. Il s’agit du début d’un article paru sur le site de la RTBF qui relate un match des Diables Rouges : « *On s’attendait à un début de match compliqué, mais c'est un peu dommage d'encaisser ce premier but, a révélé Kevin De Bruyne au micro de la RTBF. Mais je pense que nous avons ...* »

Variations possibles

1. Si le nombre de joueurs n'est pas un multiple de 3, il est possible d'utiliser des dés à 4 faces (4 joueuses par table) ou à 10 faces (5 joueurs par table). Le déroulement de l'atelier reste le même.
2. Il est possible d'ajouter des contraintes au jeu : par exemple, demander à chacune des joueuses de prendre des points de vue ou des partis pris sur la situation décrite dans le texte. Les effets produits sur le texte produit seront alors observés au moment du debriefing.

Debriefing

1. Les participants reviennent sur le processus de génération de texte qui recourt à l'aléatoire. Les choix des autres joueuses les ont peut-être surpris (par exemple, le texte mentionne le château d'Aigremont et un joueur le situe en Catalogne, le nom du village ne correspondant pas à cette région du monde). Au moment du debriefing, les participantes peuvent revenir sur ces moments particuliers et s'interroger :

- Comment réagissons-nous lorsque le texte perd son sens ?
- Comment réagissons-nous lorsque le texte possède deux idées qui sont en tension ? Choisissons-nous une des deux idées à suivre ? Essayons-nous de faire une synthèse des deux idées ?
- Y a-t-il des personnes qui partagent un imaginaire proche du nôtre ? Qui choisissent des mots proches (voire les mêmes que nous) ? Des personnes qui choisissent des mots dont le sens nous échappe ?
- Que se passe-t-il quand nous ne connaissons pas (ou peu) la thématique traitée (par exemple, si le texte parle des Diables Rouges et que nous ne suivons pas les résultat de l'équipe nationale de football belge) ?

2. Des éléments de connaissance sur les moyens techniques mis en œuvre pour générer des textes avec des grands modèles de langage et les problèmes inhérents aux contraintes mathématiques⁷ qui s'exercent sur eux pourront être relus à la lumière de cet atelier : le manque de fiabilité et les erreurs qui apparaissent dans les textes générés dans le grand modèle de langage, les hallucinations, les textes qui paraissent plats... Est-ce que les textes générés avec l'usage des dés possèdent des caractéristiques semblables ? Lesquelles ? À quoi sont-elles dues dans le processus de jeu ?

7. Voir aussi la page du wiki consacrée aux probabilités et aux statistiques.

Chapitre 4 – Machines et conflits de valeurs

Contexte du chapitre

Éviter la moralisation

Lorsqu'on considère les logiciels d'intelligence artificielle comme des outils, la question des valeurs est souvent posée de manière *morale*. En effet, l'usager est face à un outil qu'il devra apprendre à *bien* utiliser. C'est un rapport individuel de l'usagère à la machine et des injonctions morales lui dictent les « bonnes pratiques » à adopter et les « mauvaises pratiques » à éviter. En d'autres termes, il n'appartiendrait qu'à l'usager d'être un bon ou un mauvais usager, et ce, sans que celui-ci se questionne sur l'idéal qui permet de définir le bon ou le mauvais usage.. Cette conception implique au passage la supposition d'une *maîtrise* de l'outil. On peut en effet le maîtriser puisqu'il est prévisible. On est donc *responsable* de son usage. Pour prendre un exemple provocateur, si une arme existe et qu'elle est entre nos mains, il nous suffira d'apprendre à l'utiliser à bon escient. Et c'est notre responsabilité qui est engagée si ce n'est pas le cas. Plus proche de nous : c'est notre responsabilité de ne pas passer trop de temps sur le smartphone ou les réseaux sociaux. Et si nous sommes sensibles aux enjeux écologiques posés par l'IA par exemple, il reviendra à l'usagère d'apprendre à écrire de « bons prompts », ceux qui consomment le moins de ressources énergétiques possibles.

Cette conception occulte une première question : même si les moyens techniques pour inventer des grands modèles de langage sont disponibles, même si ceux-ci sont déjà là, est-ce que nous *souhaitons* les utiliser? Plus encore : *Méritent-ils d'être inventés?* *Doivent-ils exister?* C'est précisément l'absence de ce questionnement qui divise le champ social en « technophiles » et « technophobes ». Car l'*histoire des techniques* est parcourue de mouvements de résistance aux machines. Des mouvements luddites du XIX^e siècle qui se sont opposés à l'industrialisation en Angleterre en brisant des machines jusqu'aux mouvements néo-luddites actuels, il s'agit de supprimer purement et simplement la machine. Cette résistance par la seule désobéissance violente indique aussi le rapport de force entre des personnes puissantes et des personnes impuissantes : les secondes subissent des transformations de leur milieu de vie décidées par les premières. C'était le cas entre les patrons des manufactures textiles et les ouvrières et artisans au XIX^e siècle ; c'est le cas aujourd'hui entre les GAFAM et des citoyennes dans le contexte du capitalisme numérique.

Nous assistons de plus aujourd'hui à un enfllement des technologies¹ de l'information et de la communication qui nous accompagnent dans (quasiment?) toutes les dimensions de notre existence. Ces technologies « s'imposent à nous » : parce que les entreprises nous les offrent « gratuitement », et que des administrations nous incitent à les utiliser, nous pourrions dire qu'« on nous les impose ». D'un autre côté, « ça s'impose » dans des normes sociales et des habitudes partagées, parce que ces technologies sont pratiques, simplifient la vie, font gagner du temps... La question de leur existence et de leur usage devient une *nécessité* dès

1. Voir Jacques Ellul, *La Technique ou l'enjeu du siècle*, Éd. Economica (2008)

lors que ces technologies sont omniprésentes. Elles recouvrent tous les pans de nos milieux de vie, nous ne pouvons pas y échapper.

Toute résistance n'a pas à prendre une forme violente pour mériter que l'on se penche dessus. Il importe d'observer où cela résiste *pour chacun d'entre nous*, là où des tensions apparaissent quand une nouvelle machine émerge. C'est l'attachement à nos métiers, à nos manières de faire, d'agir, d'aimer, d'être en relation, etc. que les nouvelles technologies mettent en question. Celles-ci ne sont toutefois pas seulement entre les mains des géants du numérique, le numérique ne forme pas un tout homogène. D'autres logiciels cohabitent avec ceux délivrés par les géants du numérique, notamment les logiciels libres qui sont d'autant d'exemples d'une tentative de résistance², mais aussi de la création de nouvelles manières de faire exister le numérique.

Discuter les conflits de valeurs

Réfléchir en termes d'outil pose encore un autre problème : nous ne pouvons pas nous contenter de nous adresser la question des attachements de manière individuelle, nous devons nous la poser *collectivement*. Car la machine ne vient pas seulement toucher à nos attachements singuliers. Et puisque son émergence ne touche pas chacune de la même manière, il est nécessaire d'observer à la fois les résistances et réactions *communes* et celles que nous ne partageons pas au sein d'un groupe social, d'une communauté. C'est ici que la conception d'un logiciel d'intelligence artificielle comme « acteur culturel » (voir [Chapitre 3 – Culture technique : entrer dans le monde des machines](#)) qui transforme son milieu de manière imprévisible (voir [Chapitre 2 – Se repérer par rapport à l'IA dans nos pratiques médiatiques quotidiennes](#)) montre sa richesse et son importance.

Lorsque le milieu est transformé par l'apparition de la machine, ce sont les conflits de valeurs entre nous qui se rejouent autrement. La nouvelle machine peut aussi faire apparaître des conflits latents. Dans les luttes luddites, par exemple, les valeurs capitalistes de rendement entrent en contradiction avec le désir du travail bien fait de l'artisan. Des conflits de valeurs peuvent aussi s'exprimer entre usagères. Toutefois, il ne faudrait tomber à nouveau dans l'opposition grossière et non réfléchie entre « technophiles » et « technophobes ». Par exemple, on évitera de se dire que les réseaux sociaux et *Whatsapp* nourrissent les personnes éprouvant le désir d'être en contact permanent avec leurs proches alors qu'elles terrifient quelqu'un qui a besoin de solitude et de déconnexion. En effet, une telle position du problème fait obstacle à la discussion. On peut être à la fois technophile et technophobe, de façon variable, selon les nécessités, selon les techniques. Mais de telles nuances ne sont possibles qu'à condition d'observer, d'expliciter et de questionner les tensions que les techniques suscitent inévitablement entre nous et en nous.

Par ailleurs, de tels conflits de valeurs sont imprévisibles. Ils ne sont pas issus d'un plan dans la tête de la conceptrice. Qui aurait pu imaginer par exemple que l'application *Tinder* émerge, transformant *par son existence même* nos manières d'entamer une relation amoureuse ? On n'aurait pas pu la prévoir, certes, mais cela n'empêche pas qu'on la comprenne rétrospectivement. Cette application qui propose une modalité de rencontre amoureuse sur base d'une photo et d'un court texte s'inscrit en effet dans une histoire sociale et technique (les bals de village, les agences matrimoniales, les sites de rencontre, etc.). Les concepteurs de l'application ne pouvaient néanmoins prédire avec certitude, ni qu'elle survivrait, ni les effets qu'elle produirait sur nos manières d'entrer en relation. Si nous considérons *Tinder* comme un acteur culturel de notre société, avec un mode d'existence particulier (voir [Chapitre 3](#)), nous pouvons observer les effets du logiciel sur notre société, en particulier quelle

2. Voir Sébastien Broca, *Pris dans la toile, De l'utopie d'internet au capitalisme numérique*, Éd. Seuil (2025).

nouvelle configuration du milieu social et technique il fait émerger. Nous pouvons donc être attentives à de nouvelles questions : Quelles sont les transformations à l'œuvre qui nous échappent, malgré nous et malgré la volonté du concepteur ?

Nous pourrions encore croire à ce stade que les conflits de valeurs se jouent uniquement entre les êtres humains (usagères, concepteurs, entreprises, techniciennes). Mais si nous considérons que les machines sont des acteurs culturels, cela nous oblige à observer les tensions qui se jouent également à l'intérieur même de la machine. Celle-ci est soumise à des contraintes techniques, d'ordre physique, biologique, mathématique, avec lesquelles la machine doit composer (elle n'est pas le pur produit de l'imagination des concepteurs). Face à ces contraintes, les conceptrices opèrent des choix en fonction des critères qui sont à la fois techniques et humains, elles font des compromis. La machine est ainsi elle-même aux prises avec des conflits de valeurs : Est-ce l'accessibilité partout dans le monde qui sera privilégié ? Ou la limitation des ressources énergétiques ? Le fait de parvenir à produire des textes les plus cohérents possibles ? Ou ceux qui ressemblent le plus à ce que des êtres humains pourraient produire ? Ou ceux qui sont créatifs ?

Ce dernier chapitre vise d'abord à observer et à expliciter ces conflits de valeurs. Les tensions sont au moins de trois ordres : a) à l'intérieur de la machine, b) entre la machine et les êtres humains, c) entre les êtres humains. Nous nous demanderons ensuite ce que nous faisons face à ces conflits de valeurs.

Atelier 1 : Esprit critique et analyse d'images

Enjeux de l'atelier

Travailler l'« esprit critique » apparaît comme essentiel alors que les logiciels d'intelligence artificielle génèrent des textes, des images, des musiques, des vidéos... sans qu'une autrice soit identifiable, ni ses intentions. Encore faut-il préciser ce que nous entendons par « esprit critique ». Il a tendance aujourd'hui à se réduire à un certain type de questionnement : savoir si c'est une machine ou un être humain qui est à l'origine du contenu produit ; essayer de le détecter dans un contexte où cela devient de plus en plus difficile ; et finalement, quand il s'agit d'un contenu de savoir, observer si le contenu produit est « vrai » ou « faux ». Les efforts se concentrent alors sur la lutte contre la désinformation ou sur le « décryptage » des informations, un terme aussi omniprésent que vague et réducteur.

Cette version de l'esprit critique suppose qu'il existe des faits, des réalités objectives et sans biais qu'on peut énoncer. Le travail de l'esprit critique s'arrête au moment où on s'est assuré de l'objectivité des faits. Il existe des risques à cette approche de l'esprit critique : 1) installer un rapport de méfiance par rapport aux informations qui nous parviennent (qui pourrons-nous encore croire ?), 2) se concentrer uniquement sur la question du vrai et du faux sans nous donner la peine d'observer plus finement quelles sont les raisons de croire, d'adhérer aux informations.

Nous partons ici de l'hypothèse que les informations passent toujours par des filtres (autant de l'auteur que de l'auditrice) qu'il est utile d'observer (le journal dans lequel l'information est publiée, le point de vue de l'autrice, son intention, sa formation, son milieu socio-économique, le contexte dans lequel l'information est mise en forme...). Ces filtres sont observables et objectivables par l'auditrice à partir du moment où on se donne la peine de s'y attarder. Dans la même optique, il ne s'agit pas tant de corriger tous les biais des informations (ce travail serait infini), mais plutôt de rendre visible les biais, de les reconnaître. C'est cette

version de l'esprit critique que nous poursuivrons ici, en l'adaptant au cas particulier des logiciels d'intelligence artificielle.

Le fonctionnement des logiciels d'intelligence artificielle brouille les catégories avec lesquelles nous avions l'habitude d'observer la fiabilité, la solidité des contenus auxquels nous avons accès et les filtres de l'« auteur » : il n'y a pas d'autrice de texte, elles ne sont issues d'une perspective identifiable (historique, sociale, culturelle, politique...), il n'y a pas de comité de rédaction affiché, les critères de valorisation des données d'entraînement des logiciels ne sont pas connus (les concepteurs de logiciels ne les révèlent pas)... Quels repères pouvons-nous retrouver dans ce nouveau paysage médiatique ?

Pour ce faire, nous tâcherons de déplacer la focale, en nous concentrant sur la génération d'images. Alors que le réflexe dominant est de se demander si une image a été générée par un être humain ou une machine, nous déplaçons la question : quel que soit le concepteur de l'image (une personne seule ou assistée d'une IA), nous nous demandons plutôt quel genre de message sur le monde se trouve dans l'image. L'esprit critique se situe aussi là : quel que soit le contenu (vrai ou faux), d'être attentive au message que celui-ci délivre. Qu'est-ce que l'image dit du monde dans lequel nous vivons ? Quel message fait-il passer ? Quel genre de monde les générateurs d'images sont-ils occupés à façonner ? Qu'est-ce que ces images valorisent ? Être attentif de manière systématique à ces enjeux face aux images générées par les modèles d'intelligence artificielle permettra peut-être de nous rendre un peu plus sensibles à l'esthétique promise par les entreprises qui détiennent ces logiciels. Ceci nous permettra d'objectiver un peu plus les « filtres » qui sont les leurs. Quelles sont les valeurs inscrites dans les machines par leurs conceptrices (peut-être même à leur insu) ? Et nous demander également comment ces valeurs nous affectent personnellement : comme nous choisissons les magazines que nous allons consulter, avons-nous envie d'être exposés aux images générées par des logiciels d'intelligence artificielle ? À quelles occasions ?

Présentation de l'atelier et consignes

Première séquence

L'animateur de l'atelier dispose d'un stock d'images qui fonctionne par deux : une image de départ prise par une photographe (amateur ou professionnelle) et une image générée en deux temps par des logiciels d'intelligence artificielle : dans un premier temps, un grand modèle de langage génère une description textuelle de l'image initiale ; celle-ci est soumise ensuite à un logiciel de génération d'images qui génère la seconde image(voir annexe ³).

Les participants travaillent en sous-groupes de 2 ou 3 personnes. Chaque sous-groupe reçoit une première image sans savoir si celle-ci a été produite par un logiciel ou non. On suspend cette question, en considérant a priori que les deux images sont également dignes d'intérêt. Chaque sous-groupe constitue collectivement une lecture de l'image en répondant à une série de questions pour travailler leur interprétation commune :

- Que se passe-t-il dans l'image ?
- Quelles sont les couleurs qui apparaissent ? Les formes (leur position, leur orientation, leur dimension respective) ? Y a-t-il des motifs récurrents ? Des lignes dominantes ? Des orientations qui structurent l'image ?⁴

3. Un stock d'images est disponible en [Annexe : Ressources pour les ateliers](#). Elles sont accompagnées des descriptions générées par le grand modèle de langage. La qualité des images de départ est importante, en particulier si les images sont agrandies, la pixellisation rendant l'exercice plus difficile.

4. Pour aller plus loin dans la lecture d'images, voir par exemple le syllabus *Apprendre à lire* des images disponible sur le blog [Dans la grotte de PhiloCité](#).

- Y a-t-il des dissonances dans l'image ? Des éléments qui paraissent incongrus ?
- De quelle esthétique l'image se rapproche-t-elle ? Où pourrait-elle avoir été publiée ?
- Si cette image avait une autrice, quelle pourrait être son intention ? Qu'est-ce qui l'intéresse ?

Dans un second temps, les deux sous-groupes qui ont travaillé sur les images qui fonctionnent ensemble sont rassemblés. Ils reçoivent la description de l'image générée par le grand modèle de langage sur base de l'image initiale. Des questions structurent ce deuxième temps en sous-groupes (de 4 à 6 personnes) :

- Chaque sous-groupe présente sa lecture de l'image sur laquelle il a travaillé.
- Comparaison des deux images avec la description textuelle : quelles sont les différences entre les deux images ? Quels sont les éléments de la description qui apparaissent dans les deux images ? Ceux qui ont été omis ? Quelles sont les ambiguïtés de la description textuelle qui sont visibles dans le passage d'une image à l'autre ?
- Est-il possible de se mettre d'accord sur l'image qui a été générée par les logiciels d'intelligence artificielle ? Qu'est-ce qui permet de le savoir ?

Deuxième séquence

Les participantes travaillent en sous-groupes de 2 ou 3 personnes. Chaque sous-groupe travaille sur une photographie prise par un photographe (amateur ou professionnel). De la même manière que lors de la première séquence, une description textuelle de l'image initiale est produite, cette fois-ci par les participantes. La description est dactylographiée de sorte à la soumettre à un générateur d'images.

Chaque sous-groupe pourra alors observer l'image produite et se demander ce qui a manqué dans la description, ce qui pourrait être précisé, les ambiguïtés du texte qui ont pu mener à l'image produite... de la même manière que dans la première séquence.

Variations possibles

1. La deuxième séquence est l'occasion de comparer des logiciels de génération d'images, en soumettant une description identique à des logiciels différents afin de comparer les résultats. Il est également possible de tester ainsi l'évolution des logiciels en comparant des versions plus ou moins récentes d'un même logiciel. Alors que les « hallucinations » et les éléments incongrus dans les images générées étaient fréquents il y a plusieurs années, ils se font de plus en plus rares.
2. Dans la deuxième séquence, il est également possible de soumettre plusieurs fois la même description d'images au même logiciel. Celui-ci ne générera pas exactement la même image puisque son fonctionnement est basé sur un algorithme statistique. Les images générées pourront également être comparées.

Debriefing

1. Il est utile de se demander quel genre d'images les générateurs d'images valorisent par défaut (sans ajustement de la part de l'usagère). Qu'est ce que ces images donnent à voir et, éventuellement, qu'est-ce qu'elles ne donnent pas à voir ? Comment pourrait-on qualifier les représentations du monde que ces logiciels génèrent ? Et si on tient compte du fait qu'une image n'est pas qu'une représentation du monde, mais qu'elle

en fait partie et participe donc de son façonnement, quel genre du monde ces images contribuent-elles à façonner ? Par ailleurs, les images générées ont-elles des caractéristiques communes ? Lesquelles ? Correspondent-elles aux *a priori* que nous avions par rapport aux logiciels de génération d’images ?

2. Si ces images construisent un monde, nous pouvons également nous demander ce que cela nous fait d’être exposés à ces images. Où pourrions-nous les retrouver avec plaisir ? Où pourrions-nous nous attendre à les voir apparaître (parce qu’elles renvoient à une représentation du monde partagée) ? Avons-nous envie d’utiliser ces logiciels ? À quelles fins ?
3. Si nous suivons l’hypothèse de la philosophe Marie-José Mondzain, « *nous voyons infiniment plus de choses que ce que nous pouvons en dire et nous sommes capables de dire infiniment plus de choses que ce que nous voyons* »⁵. La description textuelle de l’image reste dès lors toujours partielle, l’image étant inépuisable par le texte et inversement. Les générateurs d’images proposent une certaine modalité de mise en relation entre le texte et l’image. Parvenons-nous à produire l’image que nous décrivons ? Avec quel niveau d’exigence sur l’image produite ? Mettre plus de mots permet-il d’avoir une image plus à propos ? Est-ce toujours le cas ?
4. Cet atelier peut également servir à s’interroger sur le fonctionnement des générateurs d’images, notamment les contraintes statistiques avec lesquelles ils doivent composer. Il peut alors être mis en relation avec le deuxième atelier du Chapitre 3.

Atelier 2 : Des valeurs dans la machine

Enjeux

Au Chapitre 3, nous avons plongé dans le monde des objets techniques en étudiant le fonctionnement d’un logiciel d’intelligence artificielle et les conditions techniques de son existence. À l’issue de ce livret, il est temps de revenir sur les grandes entreprises qui conçoivent et mettent à disposition ces logiciels. En pariant sur le fait que notre perspective est enrichie par une plus grande familiarité avec les machines, c’est-à-dire par une culture technique plus importante.

Malgré leur puissance financière démesurée, les grandes entreprises d’intelligence artificielle sont aussi aux prises avec le réel, elles doivent composer avec les contraintes imposées par le bon fonctionnement des logiciels qu’elles produisent. Face à ces contraintes, elles opèrent des choix à la fois techniques et humains (un *design* de l’interface avec l’usager, un déploiement du logiciel, un modèle économique, une organisation du travail, etc.). Il s’agit dès lors d’observer les valeurs que ces entreprises colportent en observant de plus près les choix qu’elles opèrent parmi toutes les possibilités techniques et humaines possibles. À partir de l’observation des valeurs qui se trouvent dans nos machines, nous pouvons mettre en évidence indirectement ce qui importe pour les entreprises, quel monde construisent les technologies qu’elles développent.

L’émergence des logiciels d’intelligence artificielle s’ancre dans un contexte socio-historique et politique particulier⁶. Depuis l’arrivée d’internet (au départ, un réseau décentralisé), le marché du numérique s’est progressivement concentré entre les mains de quelques entreprises surnommées GAFAM. Ces entreprises imposent en effet une certaine culture

5. Citation de Marie-José Mondzain dans l’émission *Éducation à l’image, pour quoi faire ?* (épisode 4/7) en 2014. Vidéo disponible sur Dailymotion.

6. Pour aller plus loin, voir Sébastien Broca, *Pris dans la toile, De l’utopie d’internet au capitalisme numérique*, Éd. Seuil (2025).

du numérique (facilité d'usage, opacité du fonctionnement, etc.) ainsi que des modèles économiques particuliers (« service gratuit », captologie, économie de l'attention, collecte de données et ciblage publicitaire, etc.).

Puisque ces valeurs sont inscrites dans les machines, nous pouvons les observer. Nous nous demanderons ce que les logiciels d'intelligence artificielle valorisent et dévalorisent. Comment celles-ci rentrent en tensions avec nos valeurs ? Mettre des mots sur les valeurs (les nôtres et celles des machines) est un travail collectif. Il vise à objectiver les valeurs des grandes entreprises à partir de nos perceptions singulières. Et, finalement, mieux nous positionner collectivement par rapport aux machines.

Présentation de l'atelier et consignes

Les participantes dressent collectivement un tableau reprenant les valeurs *actuelles* qui sont comprises à l'intérieur d'un objet technique spécifique (par exemple un générateur de texte comme ChatGPT). Ce tableau comporte deux colonnes :

- 1) les valeurs techniques internes à la machine, celles qui proviennent
 - des *contraintes physiques* (par exemple, débit d'informations disponibles sur le réseau internet, refroidissement des serveurs)
 - des *contraintes mathématiques* (par exemple, les théorèmes mathématiques qui décrivent les lois statistiques) ;
- 2) les *valeurs humaines*, qui proviennent des contextes *économique* (qu'est-ce qui permet aux entreprises de tirer leur épingle du jeu ?), *culturel* (quelles sont les pratiques médiatiques dominantes ?), *juridique* (quelles sont les dispositions légales en vigueur ?) dans lequel l'objet technique évolue.

Les deux catégories ne sont pas étanches ; elles sont là pour nous aider à observer les différentes valeurs qui sont à l'œuvre au sein de la machine.

En particulier, elles visent à ne pas passer à côté des valeurs inscrites dans les contraintes liées au fonctionnement de la machine. La méconnaissance du fonctionnement de la machine ne doit néanmoins pas avoir d'effet inhibant pour les participants.

Tip

L'enjeu de l'exercice ne consiste pas à distinguer les deux catégories (valeurs techniques et valeurs humaines), mais bien de mettre des mots collectivement sur les valeurs du logiciel.

Valeurs techniques (qui découlent des contraintes internes à la machine)

ÊTRE DANS LA MOYENNE :
Utiliser les mots les plus fréquents
...

Valeurs humaines

ÊTRE ACCESSIBLE : Être utilisable quel que soit le niveau de littératie numérique, quel que soit le public
...

Une manière d'entrer dans les valeurs humaines de la machine est de se concentrer sur les entreprises qui les conçoivent : que cherchent-elles à produire ? Comment parviennent-elles à leurs fins malgré les contraintes techniques auxquelles la machine est soumise ?

Il est important de s'attarder sur chacune des valeurs et de prendre le temps de mettre des mots car nous n'avons pas la même sensibilité, nous ne percevons pas la machine exactement de la même manière. Par exemple, est-ce que la valeur technique de la machine, c'est d'être dans la moyenne (c'est-à-dire de véhiculer des discours dominants, avec les mots les plus fréquents) ou est-ce plutôt d'être plausible (c'est-à-dire d'être le plus cohérent, de ressembler le plus à un discours écrit par un être humain)? Elles seront travaillées collectivement en faisant droit aux sensibilités de tous les participants.

Une fois que le tableau de valeurs paraît suffisamment riche, l'atelier peut se poursuivre en observant les tensions entre les valeurs. Dans un premier temps, les participantes examinent les tensions entre deux (ou plusieurs) valeurs humaines. Par exemple, si dans les valeurs humaines apparaissent à la fois le fait d'être accessible à tout le monde et le fait d'être utile : le texte généré sur une thématique particulière ne sera pas satisfaisant pour un spécialiste de la thématique. Le jargon utilisé ne correspondra probablement pas à ce qui est attendu, des erreurs apparaîtront probablement dans le texte pour un lecteur averti, *etc.* Dans un second temps, les participants explorent les tensions entre des valeurs techniques et des valeurs humaines. Par exemple, si le grand modèle de langage génère des textes « dans la moyenne » et qu'il poursuit l'objectif d'être utile pour tout le monde, les usagers qui cherchent à générer des contenus originaux, ne parviendront pas à leurs fins, sauf à fournir eux-mêmes cette originalité. Nous pourrions même craindre que la signification du mot « original » soit transformée avec l'émergence de l'intelligence artificielle. Des tensions peuvent également exister entre plusieurs valeurs techniques, auquel cas les conceptrices des logiciels opèrent des choix (voir Chapitre 3). Par exemple, une tension existe entre produire une réponse rapide et produire un texte fidèle à la requête de l'usager.

Variation possible

Si nous souhaitons aller un peu plus loin dans l'atelier, il peut être intéressant de faire l'atelier avec des logiciels d'intelligence artificielle différents. Les valeurs sont-elles les mêmes ? Sommes-nous capables d'observer des différences de valeurs entre les logiciels ? Évoluent-elles au cours du temps ? Les choix de modération des contenus sont éclairants : ChatGPT interdisait jusqu'en 2025 des contenus érotiques ; DeepSeek ne peut pas générer de contenus à propos de certains événements historiques comme les manifestations de la place Tian'anmen de 1989.

Certains logiciels permettent aux usagères de changer des paramètres (par exemple : vitesse d'exécution, qualité du résultat, résultat plus ou moins original...). Ces paramètres peuvent également être explorés. Cette expérimentation met en évidence le degré d'appropriation potentiel pour l'usager.

Debriefing

1. Est-ce facile de mettre des mots collectivement sur les valeurs ? Est-ce aisément de se mettre d'accord ? Les valeurs techniques sont-elles perceptibles pour l'ensemble du groupe ? Qu'est-ce qui pourrait nous permettre d'entrer davantage dans les contraintes techniques du logiciel (ou de l'objet technique choisi) ?
2. Parmi les valeurs qui apparaissent dans le tableau, y en a-t-il qui dérangent, qui entrent en tension avec les valeurs des participantes (personnellement et collectivement) ? Celles-ci sont-elles prêtes à composer avec ces valeurs (éventuellement, pour certains usages, dans certaines dimensions de leur existence) ? Des valeurs paraissent-elles problématiques pour les participants ? Sont-elles rédhibitoires (pour

certains usages, dans certaines dimensions de votre existence)? Est-ce partagé au sein du groupe?

3. Y a-t-il une valeur qui n'est pas suffisamment prise en compte par les concepteurs de la machine et qui serait souhaitable? Par exemple : être économe en ressources (eau, électricité, matériaux); augmenter la capacité de l'usager à quitter le logiciel en ne l'incitant pas à poser une autre question; etc. Comment pourrions-nous transformer la machine pour qu'elle soit plus conforme à nos valeurs? Est-ce imaginable?
4. Lorsque nous faisons face à un problème concret, le réflexe dominant est de trouver une solution technique (c'est la sacralisation de la machine qui est abordée au Chapitre 3; elle peut solutionner tous les problèmes). Comme l'exercice le montre, un logiciel d'intelligence artificielle mobilise une série de valeurs qui peut créer d'autres tensions par ailleurs. Il est donc judicieux de les observer. Peut-être est-ce effectivement une solution qui convient, peut-être pas. Quelles autres solutions pourraient être envisagées (avec ou sans autres objets techniques)? Si nous faisons l'impasse sur cette discussion, les logiciels d'intelligence artificielle risquent de nous emmener dans une direction imprévue sans même que nous nous en rendions compte.
5. Les logiciels d'intelligence artificielle évoluent et les valeurs qu'ils véhiculent aussi. Pouvons-nous observer ces transformations? Comment qualifier les anciennes et les nouvelles valeurs? Est-ce que cela transforme également nos affects par rapport au logiciel?

Conclusion

Nous ne savons pas comment les logiciels d'IA transformeront notre monde. Le livret est conçu pour pouvoir y revenir au moment où de nouvelles questions se posent, de nouvelles tensions se manifestent ; il ne s'agit pas de faire les ateliers « une fois pour toutes ». Y revenir permet de continuer à développer notre culture technique, d'observer comment notre manière de percevoir ce qui se passe s'est un peu transformée, mais aussi d'inventer et de cultiver d'autres manières d'observer ce qui nous arrive que celles qui dominent l'espace public et médiatique.

Dans un contexte où l'IA est déjà utilisée, où son usage se normalise et où de nouvelles habitudes se prennent, il s'agit d'abord de dépasser les sentiments de fascination et d'écrasement pour sortir de la sidération, personnellement et collectivement. Le livret est un outil de réflexion afin de se donner les moyens de se positionner. Souhaitons-nous utiliser l'IA ou non ? Pour quelles tâches ?... Ces questions nous renvoient forcément à des discussions démocratiques qui dépassent largement la simple question de l'usage de l'IA : Quelle éducation voulons-nous ? Quelles activités souhaitons-nous réaliser dans le cadre professionnel ? Quels liens affectifs souhaitons-nous tisser et avec qui ?...

Il s'agit de se réapproprier ces questions à la fois quotidiennes et fondamentales, en observant les transformations que l'IA opère. Nous ne pouvons pas laisser le soin à quelques entreprises de façonnner nos milieux de vie. Pour sortir de la sidération, développer collectivement notre culture technique est une condition : comprendre les contraintes qui s'imposent à nous, mais aussi, celles que l'IA imposent aux autres êtres, humains (par le travail) et non-humains (quel est le sort des machines et de l'environnement). C'est l'organisation de notre société dans son ensemble que l'IA vient bousculer.

Ce questionnement vise à nous redonner du pouvoir d'agir sur ce qui nous arrive. Et *in fine*, sur la société dans laquelle nous vivrons. Ce livret s'inscrit dans une démarche qui est déjà à l'œuvre, chez toutes les personnes qui choisissent d'en faire un sujet de discussion collective, chez celles qui se réapproprient les logiciels pour répondre à leurs besoins et non à ceux des entreprises¹, chez celles qui inventent et réinventent de logiciels... la liste est longue – heureusement.

1. Voir par exemple Anne Alombert, *De la recommandation algorithmique privée aux pratiques citoyennes et contributives : assurer nos libertés à l'ère de l'intelligence artificielle*, article publié en 2024, disponible en ligne.

Annexe : quelques textes et citations

Chapitre 1

La méchanceté

C'est là un sérieux obstacle [comment dépasser le statut de *sujet connaissant* qui empêche, dès lors qu'on y est installée, de comprendre la méchanceté]. Est-il possible de le contourner? J'ai tenté de le faire en procédant par allers et retours entre récit et réflexion. Le grand avantage des récits de fiction en effet, c'est qu'ils ne s'adressent pas au *sujet connaissant* : dans un récit, il y a toujours quelqu'un aux prises avec quelqu'un d'autre, et qui est guetté par la démesure, la confusion ou une autre manière de dépasser les bornes. Les récits nous parlent du *sujet existant* et ils s'adressent à lui. D'ailleurs, on consomme des récits de fiction en lisant, en allant au cinéma ou en regardant la télévision : activités auxquelles on se livre bien moins pour savoir que pour éprouver du plaisir ou, au moins, éviter l'ennui.

Pour n'être pas de l'ordre du savoir, les scènes d'un récit et leur succession n'en sont pas moins parlantes ; c'est pourquoi, si nous pouvons apprendre quelque chose d'un rêve que nous avons fait, nous pouvons également apprendre quelque chose d'un récit qui nous a laissé une forte impression - la grande différence entre le rêve et le récit de fiction étant que le premier concerne une personne singulière, alors que le second concerne des milliers, voire des millions de personnes.

Essayer de nous deviner dans le miroir énigmatique que les récits de fiction nous tendent n'est pas une tâche facile. La tentation est grande en effet d'interpréter le récit à partir d'un savoir dont nous disposons, par conséquent de l'assimiler à ce savoir et, ainsi, de ne rien en apprendre de nouveau. Cela évite aussi d'avoir à se reconnaître dans l'image éventuellement déplaisante que nous montre le miroir. Dans *L'interprétation des rêves*, Freud fait au contraire le pari que *ses rêves savent quelque chose de lui qu'il ne sait pas*. C'est ce principe fondamental que je me suis efforcé de suivre. Les récits de méchanceté auxquels je me suis intéressé, j'ai donc fait l'hypothèse que ce qui faisait que, précisément, je m'y intéressais ne résidait pas seulement en eux, *mais aussi en moi*. Car même lus avec attention et analysés avec méthode, ils ne m'auraient rien appris sur la méchanceté si je ne m'étais pas préparé à ce qu'ils me disent d'abord quelque chose de ma propre méchanceté. Ce qu'un récit dit du *sujet existant*, le *sujet connaissant* n'est pas en position de l'entendre.

Il s'agissait donc de m'interroger sur les impressions laissées par le récit, sur le lien invisible qui s'était noué entre les scènes qui m'avaient frappé et la manière dont moi-même j'existe. Il s'agissait ensuite d'essayer de formuler quelque chose de ce lien, de cette résonance. Et enfin de confronter ces premières formulations avec les idées que j'ai, que nous avons, sur ce qu'est l'être humain (le récit lui-même se réfère parfois à ces idées, mais celles-ci ne concordent pas pour autant avec les "pensées" que tisse obscurément sa trame narrative). Il s'agissait donc,

non pas de ramener le récit à la raison, mais de faire qu'il me suggère les siennes et qu'ainsi il m'apporte une véritable *aide philosophique*.

François Flahaut, *La méchanceté*, Éd. Descartes & Cie (1998). pp.14-16 [Disponible en pdf en ligne sur le site de l'auteur](#).

La gouvernance par les nombres

Le point de départ de Hobbes [1588-1679], le premier mot du livre [*Le Léviathan*, publié en 1651], est celui de « nature » : « La nature (l'art par lequel Dieu a fait le monde et le gouverne) est si bien imitée par l'art de l'homme, en ceci comme en de nombreuses autres choses, que cet art peut fabriquer un animal artificiel. » Hobbes n'est pas le premier ni les dernier à vouloir ancrer le droit dans la nature, mais il voit celle-ci comme l'expression d'une technique divine qui a permis sa fabrication et préside à son fonctionnement. L'homme, fait à l'image de Dieu, se peut et se doit d'imiter cette technique, et d'user des arts mécaniques pour fabriquer à son tour des êtres artificiels, plus précisément des « animaux artificiels, c'est-à-dire des automates, qui imitent le vivant et sont capables de se mouvoir eux-mêmes. De la fin du Moyen-Âge jusqu'à l'ère industriel, les automates ont fasciné et donné lieu à des prouesses techniques de la part des meilleurs horlogers d'Europe [...].

« Étant donné, poursuit Hobbes, que la vie n'est rien d'autre qu'un mouvement de membres, dont le commencement est en quelque partie principale intérieure, pourquoi ne pourrions-nous pas dire que tous les automates (des engins qui se meuvent eux-mêmes, par des ressorts et des roues, comme une montre) ont une vie artificielle ? »

« Comme une montre » écrit Hobbes. L'automate par excellence, celui qui va occuper l'imaginaire occidentale de la fin du Moyen-Âge à l'ère industrielle, est en effet l'horloge, qui reproduit à l'échelle humaine la création du Dieu horloger. Au point de proposer à la contemplation des fidèles des horloges astronomiques dans le chœur même des cathédrales [...]. La création toute entière se présente alors comme un immense mécanisme d'horlogerie, mu par un jeu de masses et d'énergie que la physique classique va s'efforcer de déchiffrer.

Parvenu à ce point, Hobbes opère un retournement : si l'homme, imitant l'œuvre divine, crée des automates, c'est parce qu'il est lui-même un automate créé par le Grand horloger : « Car qu'est-ce que le cœur, sinon un ressort, les nerfs, sinon de nombreux fils, et les jointures, sinon autant de nombreuses roues qui donnent du mouvement au corps entier, comme cela a été voulu par l'artisan. » Ce retournement est du même type que celui de nos contemporains qui, se fondant sur le fait que l'ordinateur aurait été conçu sur le modèle de certaines facultés cérébrales, en viennent à concevoir le cerveau humain sur le modèle de l'ordinateur. À la base [du *Léviathan*] se trouve donc, non pas une métaphore, mais une véritable anthropologie physique : celle de l'homme machine. [...]

Une fois ainsi établi un continuum entre l'homme, l'animal et la machine, Hobbes franchit un dernier pas qui le conduit à concevoir l'État comme un automate fabriqué par l'homme à sa propre image : (voir p.65) [...] un imaginaire normatif qui est encore largement le nôtre : celui qui se représente le gouvernement des hommes sur le modèle de la machine.

[...]

À ce modèle physique de l'horloge qui conduisait à voir dans l'homme lui-même une machine, s'est ajouté au XIXe siècle le modèle biologique de la sélection naturelle, qui a inspiré le darwinisme social et continue de sévir sous les espèces de l'ultra-libéralisme et de la compétition de tous contre tous. À ces représentations, qui ne s'annulent pas mais se superposent, s'ajoute aujourd'hui celle de l'homme programmable portée par la cybernétique et la révolution numérique. Son modèle n'est plus l'horloge et son jeu de forces et d'engrenages, mais l'ordinateur et son traitement numérique des signaux. L'ordinateur obéit à des programmes plutôt qu'à des lois. Autorisant une extériorisation de certaines facultés cérébrales de l'être humain, il ouvre une ère nouvelle dans notre rapport aux machines, aussi bien dans le contenu et l'organisation de notre travail.

Alain Supiot, *La gouvernance par les nombres*, Éd. Fayard (2015) – pp.60 et suivantes

Du mode d'existence des objets techniques

L'homme qui veut dominer ses semblables suscite la machine androïde. Il abdique devant elle et lui délègue son humanité. Il cherche à construire la machine à penser, rêvant de pouvoir construire la machine à vouloir, la machine à vivre, pour rester derrière elle sans angoisse, libéré de tout danger, exempt de tout sentiment de faiblesse, et triomphant médiatement par ce qu'il a inventé. Or, dans ce cas, la machine devenue selon l'imagination ce double de l'homme qu'est le robot, dépourvu d'intériorité, représente de façon bien évidente et inévitable un être purement mythique et imaginaire.

Nous voudrions précisément montrer que le robot n'existe pas, qu'il n'est pas une machine, pas plus qu'une statue n'est un être vivant, mais seulement un produit de l'imagination et de la fabrication fictive, de l'art d'illusion. Pourtant, la notion de la machine qui existe dans la culture actuelle incorpore dans une large mesure cette représentation mythique du robot. Un homme cultivé ne se permettrait pas de parler des objets ou des personnages peints sur une toile comme de véritables réalités, ayant une intériorité, une volonté bonne ou mauvaise. Ce même homme parle pourtant des machines qui menacent l'homme comme s'il attribuait à ces objets une âme et une existence séparée, autonome, qui leur confère l'usage de sentiments et d'intentions envers l'homme.

Gilbert Simondon, *Du mode d'existence des objets techniques*, Éd. Flammarion (2024) – pp.13 (première édition en 1958 chez Aubier)

Chapitre 2

Fragment d'un village global

Par le passé, les effets des médias étaient expérimentés de manière plus progressive, ce qui permettait à l'individu et à la société d'en amortir l'impact dans une certaine mesure. Aujourd'hui, à l'âge électronique de la communication instantanée, je pense que notre survie, tout au moins notre confort et notre bonheur, dépendent de notre compréhension de la nature même de notre nouvel environnement, car contrairement aux changements environnementaux précédents, le médium électrique transforme entièrement et presque instantanément la culture,

les valeurs et les comportements. Une grande douleur et une perte d'identité découlent de ce bouleversement, et elles ne peuvent être soulagées que par une prise de conscience de ses dynamiques. Si nous comprenons les transformations révolutionnaires causées par les nouveaux médias, nous pourrons les anticiper et les contrôler ; mais si nous continuons à nous infliger cette transe subliminale, nous serons leurs esclaves.

Aujourd'hui, en raison de la formidable accélération que connaît la circulation de l'information, nous avons la possibilité d'appréhender, de prédire et d'influencer les forces environnementales qui nous modèlent – et de reprendre ainsi le contrôle sur notre destinée. Cette évolution se manifeste essentiellement par les nouveaux prolongements de l'homme et par l'environnement qu'ils engendrent, et pourtant nous avons l'illusion que le manière dont un médium est utilisé compte davantage que les effets qu'il a sur nous et avec nous. Nous nous retrouvons dans la posture de l'idiot technologique au stade de zombie.

Marshall McLuhan, *Fragment d'un village global*, Éd. Allia (2025) (texte original de 1968) – pp.15-17

Chapitre 3

Du mode d'existence des objets techniques

La culture s'est constituée en système de défense contre les techniques. Or, cette défense se présente comme une défense de l'homme, supposant que les objets techniques ne contiennent pas de réalité humaine. Nous voudrions montrer que la culture ignore dans la réalité technique une réalité humaine et que pour jouer son rôle complet, la culture doit incorporer les êtres techniques sous forme de connaissances et de sens des valeurs. La prise de conscience de modes d'existence des objets techniques doit être effectué par la pensée philosophique qui se trouve devoir remplir en cette œuvre un devoir analogue à celui qu'elle a joué pour l'abolition de l'esclavage et l'affirmation de la valeur de la personne humaine. [...]

La culture se conduit envers l'objet technique comme l'homme envers l'étranger quand il se laisse emporter par la xénophobie primitive. La machine est l'étrangère, c'est l'étrangère en laquelle reste enfermé de l'humain, méconnu, matérialisé, asservi, mais restant pourtant de l'humain. La plus forte cause d'aliénation dans le monde contemporain, réside dans cette méconnaissance de la machine, qui n'est pas une aliénation causée par la machine, mais par la non connaissance de sa nature et son essence, par son absence du monde des significations, et par son omission dans la table des valeurs et des concepts faisant partie de la culture.

Gilbert Simondon, *Du mode d'existence des objets techniques*, Éd. Flammarion (2024) – pp.11-12 (première édition en 1958 chez Aubier)

L'attention prédominante et exclusive donnée à une machine ne peut conduire à la découverte de la technique, pas plus que la relation avec une sorte unique d'étrangers ne peut permettre de pénétrer l'intériorité de leur mode de vie, et de la connaître selon la culture. Même la fréquentation de plusieurs machines ne suffit pas, pas plus que la fréquentation successive de plusieurs étrangers ; ces expériences ne conduisent qu'à la xénophobie ou à la xénophilie, qui sont des

attitudes opposées mais également passionnées. Pour considérer un étranger à travers la culture, il faut avoir vu jouer hors de soi, objectivement, le rapport qui fait que deux êtres sont étrangers l'un par rapport à l'autre. De même, si une technique unique ne suffit pas à donner un contenu culturel, une polytechnique ne suffit pas non plus ; elle n'engendre que tendance à la technocratie ou refus des techniques prises en bloc.

Gilbert Simondon, *Du mode d'existence des objets techniques*, Éd. Flammarion (2024) – p.219 (première édition en 1958 chez Aubier)

Psychosociologie de la technicité

Les constructeurs et vendeurs savent capter cette faim de magie qui existe dans les groupes humains selon les situations où les individus se trouvent engager. La crainte du danger, l'accablement devant le travail, la crainte de l'échec en affaires ou en amour, le désir de supériorité, n'ont pas nécessairement une signification collective mais bien individuelle. C'est la tendance de l'individu qui est à l'origine de cette adjonction de magie à l'objet technique. Tout particulièrement, on accuse souvent l'objet ménager de mécaniser la vie. Mais en fait, c'est la femme en situation ménagère qui demande à une machine à laver ou à une autre machine de la remplacer dans une tâche pénible ou dont elle craint de s'acquitter mal.

Des récits féeriques nous présentent les ménagères des temps passés accablées de travail, s'endormant à la besogne, vaincues par le découragement, mais une fée veille et les fourmis ou les gnomes viennent travailler pendant la nuit. Au réveil, tout est net, tout est prêt. La machine à laver moderne est magique dans la mesure où elle est automatique et non point dans la mesure où elle est une machine. C'est cet automatisme qui est désiré parce que la ménagère désire près d'elle pour lui donner courage une autre ménagère obscure et mystérieuse qui est l'esprit bienveillant de la buanderie comme le réfrigérateur est celui de la cuisine moderne. « Moderne » signifie magique, pour le subconscient individuel de l'utilisateur. »

Gilbert Simondon, extraits de *Psychosociologie de la technicité* (1960-1961) cités dans l'émission radio Les chemins de la philosophie *Objets trouvés (3/4) : la boîte à outils de Gilbert Simondon* diffusée sur France Culture, Les chemins de la philosophie, le 17/04/2013

Persistance du merveilleux

Il est important de considérer le contexte temporel, puis le milieu dans lesquels les imaginaires se déploient. De ce point de vue, les années qui ont suivi la Seconde Guerre mondiale, période de déploiement de technologies informatiques insaisissables, à l'usage relativement abscons et sans commune mesure avec les objets antérieurs, sont sans équivalent. L'incarnation dans des personnages déjà connus permet alors à la fois de donner du sens et de traduire pour un public non spécialiste ces concepts en expressions évocatrices, lesquelles rendent compte de nouvelles capacités d'autonomie de la machinerie numérique. C'est peut-être la raison pour laquelle on trouve tant de références aux animaux. Le fait, par exemple, de parler de virus qui se reproduisent comme leur pendant organique permet de les désigner en tant qu'entités agissantes. En personnifiant le mode

opératoire de programmes et d'outils, la langue leur attribue implicitement des attributs et des capacités d'action. Et, ce faisant, elle implique toute une gamme de relations, de manières d'interagir avec cette entité et de s'en occuper. Ce recours à une terminologie associée à des êtres animés, en particulier pour les bugs, les virus et les vers, permet également aux discours de l'industrie informatique de s'appuyer sur l'autorité des sciences naturelles pour décrire ces nouveaux risques et légitimer leur intervention [...].

Cette connotation vitaliste ne se limite pas à de simples références animales. La convocation de diverses créatures du merveilleux souligne par exemple la singularité des objets numériques et des risques qu'ils portent. Elle permet d'abord de souligner la démesure et les proportions gargantuesques de certains de ces programmes, comme on a pu le voir avec le Basilic de Roko ou les Shoggoths tentaculaires. Mais surtout, elle rend compte de leur comportement surnaturel ou aberrant. Un « daemon » informatique qui fonctionne en arrière-plan est plus facilement compréhensible quand on l'associe à l'idée d'un petit esprit invisible et agissant. L'expression de « centaure » est aussi une façon de souligner les nouveaux pouvoirs cognitifs accessibles aux êtres humains lorsqu'ils collaborent avec des systèmes d'intelligence artificielle.**

Spécifique d'une phase transitoire de domestication, l'emploi de cette terminologie plus vive et animée rappelle d'ailleurs le recours aux métaphores magiques fréquentes dans l'histoire de l'informatique, notamment lorsqu'il s'est agi de susciter l'intérêt du grand public pour les nouveaux objets numériques que furent l'ordinateur ou le smartphone. À chacune de ces étapes, la langue fut un moyen de suggérer l'autonomie relative de ces machines au service de leurs utilisateurs – à telle enseigne que l'on peur se demander si l'emploi de métaphores fabuleuses ne serait pas une manière d'entretenir une certaine méconnaissance. Elles témoignent dans tous les cas d'une relation asymétrique entre les concepteurs et les utilisateurs, presque d'une forme d'infantilisation dans la mesure où elles renforcent le manque voire la rétention d'informations quant au fonctionnement de ces objets numériques. Le « magicien », jeune entreprise ou organisation multinationale, tire les ficelles, reste masqué, et on n'a d'autre choix que de lui faire confiance.

Nicolas Nova, *Persistance du merveilleux*, Éd. Premier Parallèle (2025) – pp.178-180

La gouvernance par les nombres

Les *big data* ne sont pas des connaissances objectives dans lesquelles il suffirait d'aller puiser, comme dans des mines de houille (*data mining*), nous épargnant ainsi tout effort réflexif. Comme les algorithmes qui les exploitent, ce sont des objets construits, qui reposent sur des opérations de qualification masquées par la force dogmatique des nombres, et dont la compilation repose sur le travail peu ou pas rémunéré d'innombrables petites mains. Il faut donc pouvoir les déconstruire pour les comprendre et les soumettre à la critique rationnelle nécessaire à leur usage avisé.

Alain Supiot, *La gouvernance par les nombres*, Éd. Fayard (2015) – pp.14-15

Chapitre 4

De la bêtise numérique

Même si pour l'instant, les utilisateurs des automates computationnels ont tous appris à parler, à lire et à écrire, il est nécessaire de s'interroger sur les conséquences de l'extension de l'usage de ces dispositifs génératifs dans les années à venir. Au siècle dernier, le compositeur Béla Bartók s'interrogeait sur les enjeux de « l'extension de l'usage de la radio et du gramophone » pour la pratique de la musique. Selon lui, la possibilité d'entendre de la musique sans avoir à en jouer risquait de pousser les individus à renoncer à l'apprentissage de la musique. « Pourquoi me fatiguer à apprendre la musique alors que j'ai là des machines qui mettent à ma disposition n'importe quel genre de musique à n'importe quel moment ? » Pour ceux-là, soutient Bartók, « la radio est assurément dommageable » : « ils ne savent pas à quel point l'effet de la musique est différent sur celui qui connaît les partitions et sait les jouer lui-même, si maladroitement que ce soit ». Même si Bartók reconnaissait que la radio pouvait se révéler très instructive « pour ceux qui assistent régulièrement aux concerts et « qui ne renoncent pas à la pratique musicale active », il s'inquiétait de la perte de connaissances et pratiques musicales que cette technologie pouvait entraîner.

De même, [...] Pourquoi se fatiguer à apprendre à écrire des textes, à produire des images, à jouer de la musique et réaliser des films, alors que j'ai là des machines qui génèrent n'importe quel genre de contenus symboliques à n'importe quel moment ? Les automates computationnels deviendront tout aussi dommageables que la radio à l'époque de Bartók, car l'effet d'un texte, d'une musique, d'une image, d'un film est évidemment différent sur celui qui connaît ces techniques et sait les pratiquer que sur celui qui ne sait produire que des « prompts » adaptés.

Anne Alombert, *De la bêtise numérique*, Éd. Allia (2025) – pp.59-61

Annexe : Ressources pour les ateliers

Esprit critique et analyse d'images

Vous trouverez ici des propositions d'images. Chaque image originale est accompagnée d'une description générée par IA et d'une image générée à partir de cette description. Pour plus de facilités, un pdf contenant trois pages (image originale, image générée et description) est mis à disposition pour être imprimé.

Libre à vous de produire d'autres images.

Des jeunes avec un ordinateur



FIGURE 1 – Photo originale de jeunes qui utilisent des ordinateurs

Source : https://commons.wikimedia.org/wiki/File:Computer_learning.jpg

Description par Gemini Flash 2.5 :

L'image montre deux jeunes personnes assises à un bureau en bois, chacune devant un ordinateur portable. La personne au premier plan est vue de profil, concentrée sur son ordinateur, avec une main levée vers son menton comme si

elle réfléchissait. Elle a un casque d'écoute sur la tête. Devant elle, on voit un clavier d'ordinateur et une partie d'un ancien moniteur d'ordinateur de type CRT. La personne assise derrière est également concentrée sur son ordinateur portable, avec une souris à la main. Les ordinateurs portables sont de couleur sombre et semblent être des modèles plus anciens.

La pièce a des murs de couleur verte claire, et on peut voir une fenêtre avec des barreaux ou des stores à l'arrière-plan. L'éclairage semble venir principalement de l'extérieur par la fenêtre, créant des zones d'ombre et de lumière dans la pièce.



FIGURE 2 – Photo générée par IA de jeunes qui utilisent des ordinateurs

[Télécharger la version imprimable](#)

Une classe d'électromécanique

Source : <https://www.idbbxl.com/cours/electro-mecanique/>



FIGURE 3 – Photo originale d'élève dans une classe d'électromécanique

Description par Gemini Flash 2.5 :

C'est une photo d'un atelier ou d'une salle de classe avec plusieurs personnes engagées dans des activités pratiques, probablement liées à l'électromécanique ou à l'ingénierie.

Au premier plan, deux jeunes hommes sont penchés sur un établi, travaillant sur des composants. Celui de gauche porte un t-shirt noir et a des tresses, tandis que celui de droite porte un t-shirt blanc et un pantalon de camouflage. Ils semblent manipuler des outils et des pièces, qui pourraient être des moteurs ou d'autres dispositifs électriques.

L'établi est long et vert, s'étendant à travers la pièce, avec divers outils, pièces et équipements, y compris des moteurs et d'autres machines. Sur le côté gauche de l'image, il y a une autre table avec des matériaux et des étagères en dessous.

Plusieurs autres personnes sont visibles en arrière-plan, travaillant également sur des établis ou des équipements similaires. L'un d'eux, à droite, porte une veste matelassée sombre. Les murs sont de couleur claire et l'éclairage provient de luminaires fluorescents au plafond.

La pièce semble être un environnement de travail fonctionnel avec des câbles et des tableaux électriques visibles sur les murs, suggérant un cadre éducatif ou professionnel axé sur la mécanique ou l'électricité. Il y a des armoires de rangement rouges et sombres sur le côté droit de l'image.

[Télécharger la version imprimable](#)

Industrie textile au Ghana

Source : https://commons.wikimedia.org/wiki/File:Sleek_Garments_Industry_in_Ghana.jpg



FIGURE 4 – Photo générée d'élève dans une classe d'électromécanique



FIGURE 5 – Photo originale d'une industrie textile au Ghana

Description par Gemini Flash 2.5 :

L'image présente l'intérieur d'une usine de confection ou d'un atelier de couture, où de nombreuses personnes sont assises devant des machines à coudre industrielles, en train de travailler sur des tissus clairs, probablement blancs.

Plusieurs ouvriers, principalement des femmes, portent des coiffes ou des bonnets jaunes ou bleus, ainsi que des tabliers ou des blouses pour la plupart de couleur claire, beige ou jaune. Les machines à coudre sont disposées en rangées, avec des bobines de fil visibles sur les supports.

L'éclairage est assuré par de nombreux luminaires fluorescents suspendus au plafond. L'arrière-plan de l'usine montre d'autres postes de travail et du personnel, suggérant une opération de grande envergure. Le sol est de couleur sombre. L'ensemble de la scène dépeint une activité industrielle concentrée sur la production de vêtements.



FIGURE 6 – Photo originale d'une industrie textile au Ghana

[Télécharger la version imprimable](#)

Mary Jackson



FIGURE 7 – Photo originale de Mary Jackson

Source : https://commons.wikimedia.org/wiki/File:Mary_Jackson_working_2_-_Restoration.jpg

Description par Gemini Flash 2.5 :

Dans l'image en noir et blanc, on voit une femme aux cheveux noirs et bouclés, portant des lunettes, une blouse à pois avec un grand noeud papillon, un cardigan foncé et un pantalon foncé. Elle est assise à un bureau devant un grand panneau de contrôle avec de nombreux cadrans et boutons. Sur le bureau, il y a un dossier ouvert et un petit appareil. Elle tend la main vers le panneau de contrôle. L'arrière-plan présente des étagères remplies d'équipement.

[Télécharger la version imprimable](#)

Annexe : Ressources pour les ateliers



FIGURE 8 – Photo générée de Mary Jackson



FIGURE 9 – Photo originale d'un métier à tisser datant de 1935

Métier à tisser en 1935

Source : https://commons.wikimedia.org/wiki/File:Mary_Jackson_working_2_-_Restoration.jpg

Description par Gemini Flash 2.5 :

L'image en noir et blanc représente un homme, portant un costume et une moustache, debout derrière un grand métier à tisser en bois. Le métier semble ancien et est équipé de nombreux fils, ainsi que d'un rouleau de tissu en cours de tissage. On aperçoit d'autres personnes, partiellement visibles, en arrière-plan. L'ensemble de la scène donne l'impression d'un atelier d'une époque révolue. On peut également distinguer du texte sur l'image, notamment "CHAPAGNE", "MEU 2614B", "C38788" et "SELD B".

[Télécharger la version imprimable](#)

Dessin du futur

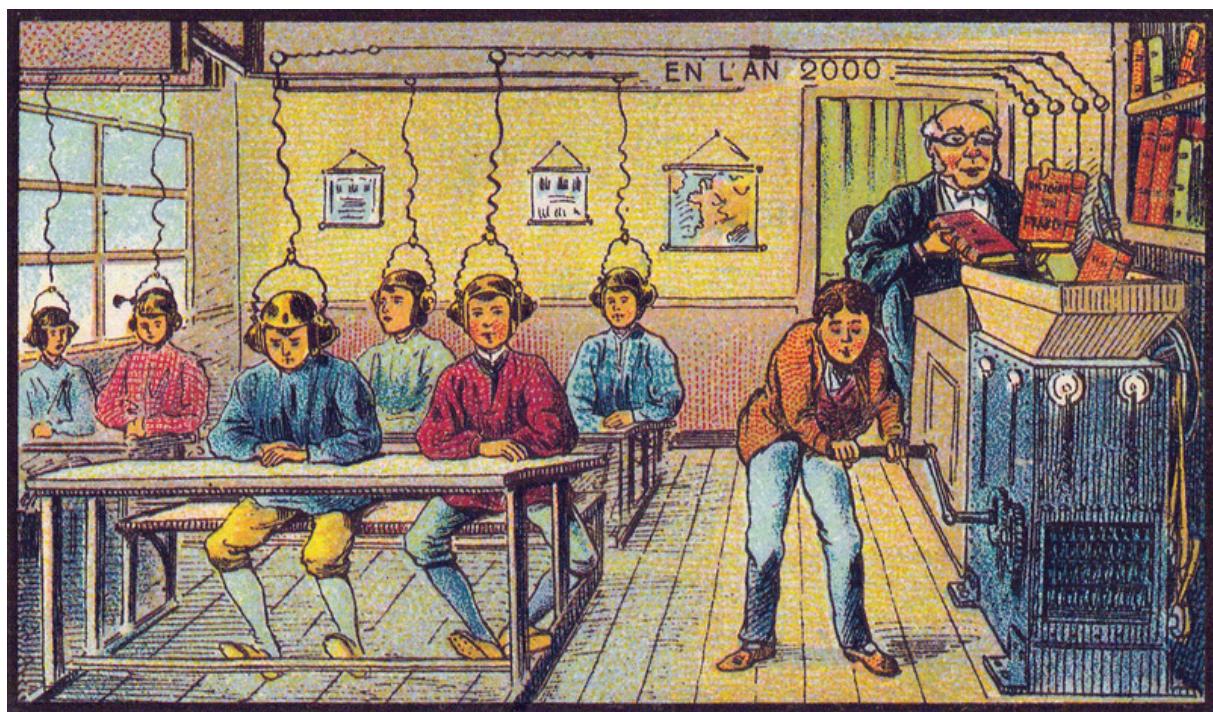


FIGURE 10 – Dessin originale d'une future école

Source : https://fr.wikipedia.org/wiki/Fichier:France_in_XXI_Century._School.jpg

Description par Gemini Flash 2.5 :

Cette image est une illustration de style ancien, représentant une vision futuriste d'une salle de classe.

Au premier plan, plusieurs élèves sont assis à des pupitres individuels. Ils portent des casques connectés par des fils qui pendent du plafond, ce qui suggère une méthode d'apprentissage directe ou "télépathique".

À droite de l'image, un enseignant barbu, vêtu d'un costume, se tient à côté d'une machine complexe. Cette machine semble broyer des livres, et les fils des casques des élèves y sont connectés, impliquant que la connaissance des livres est directement transmise aux cerveaux des élèves. Un jeune garçon actionne une manivelle sur le côté de la machine.

En arrière-plan, on peut voir des murs avec des tableaux ou des cartes accrochés, et une fenêtre sur le côté gauche de la pièce. L'inscription "EN L'AN 2000" est visible en haut de l'image, indiquant qu'il s'agit d'une prévision pour l'an 2000. Le style de l'illustration est coloré et détaillé, typique des gravures du début du 20ème siècle.

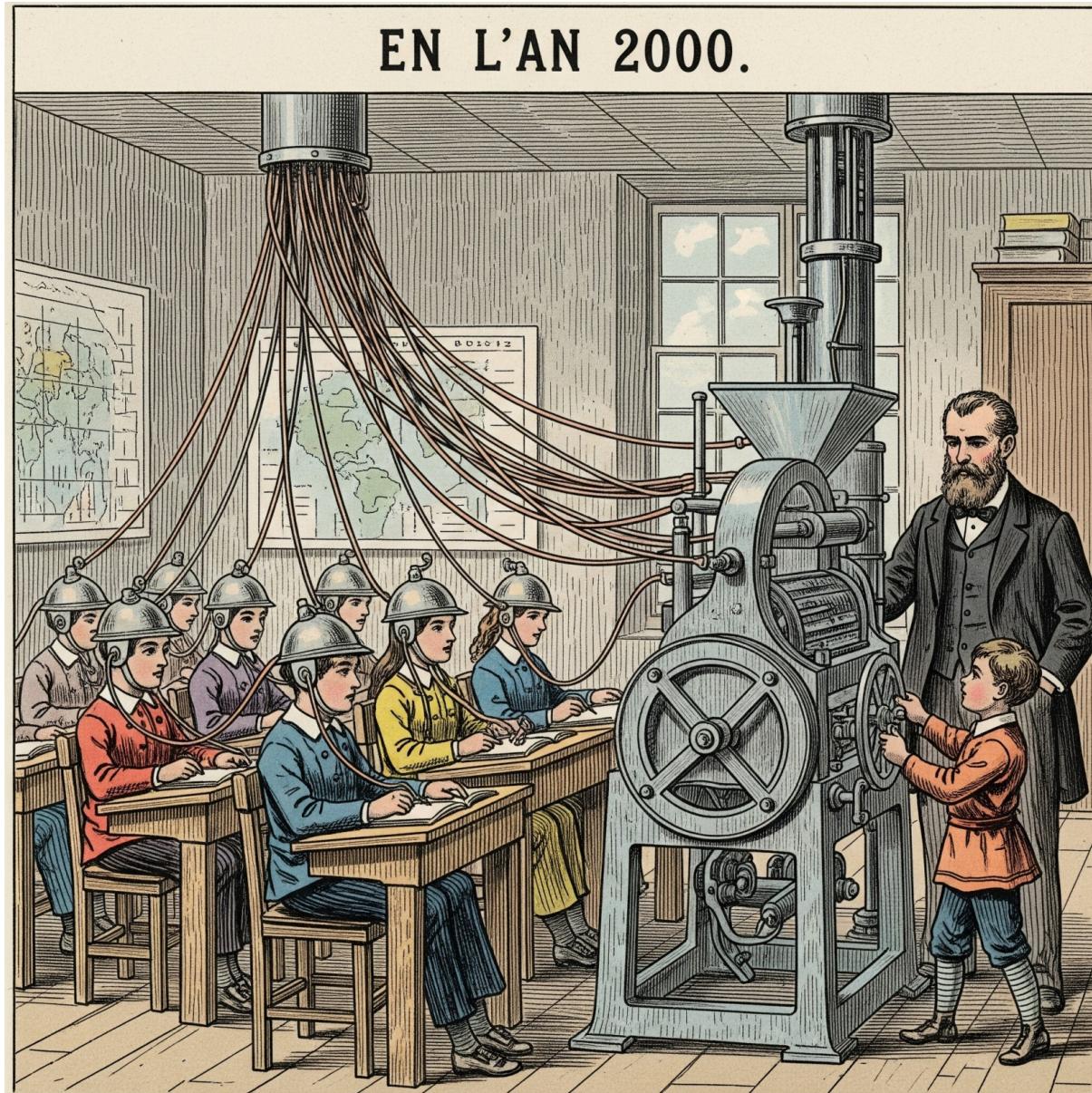


FIGURE 11 – Dessin générée d'une future école

[Télécharger la version imprimable](#)

Dessiner le fonctionnement de l'IA

Voici deux exemples de dessins réalisés lors d'un atelier. Ils sont là à titre d'exemple pour les animateurs. Nous vous conseillons de ne pas les montrer aux participants afin de mobiliser un maximum leurs imaginations.

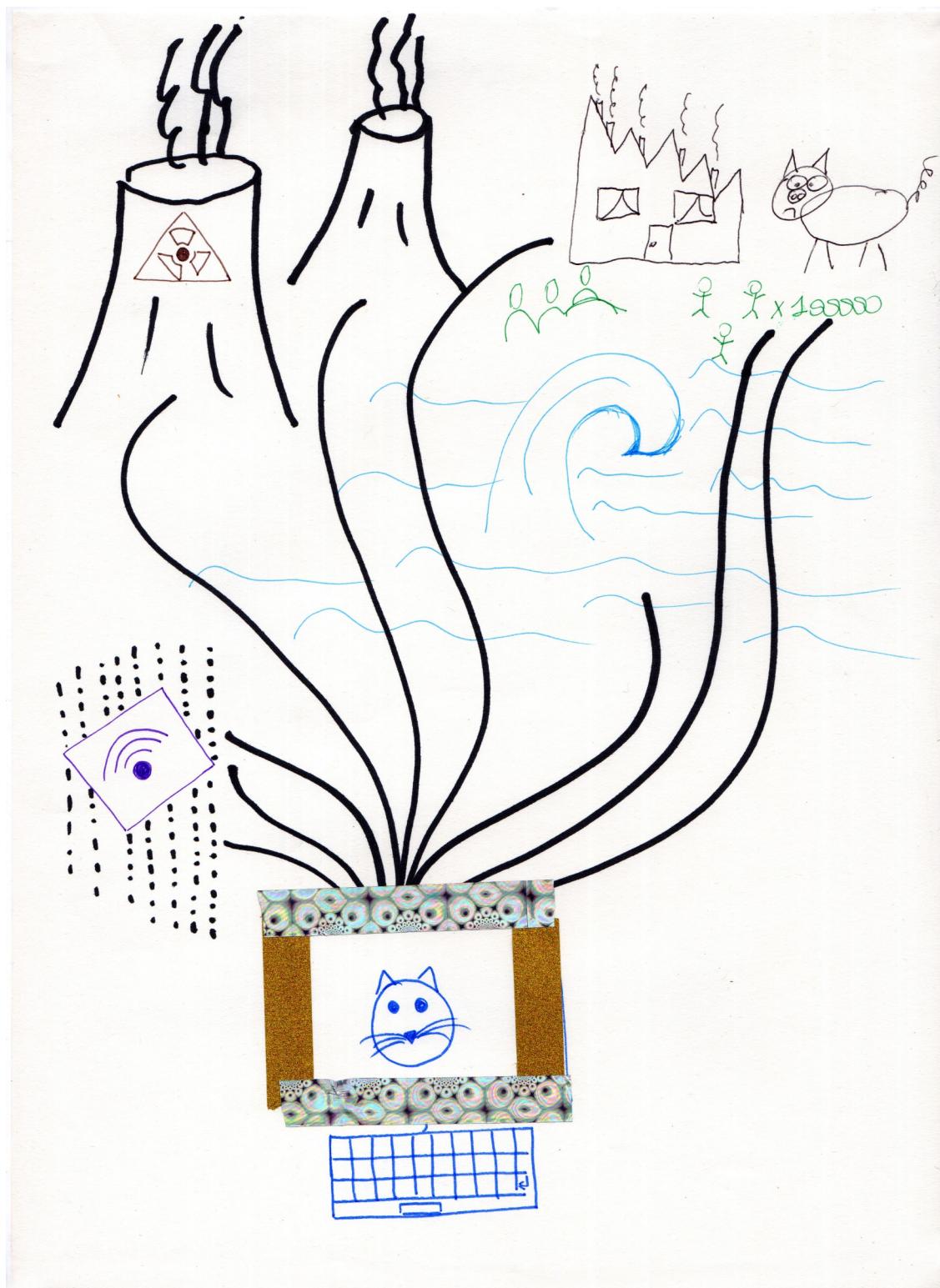


FIGURE 12 – Un dessin du fonctionnement de l'IA

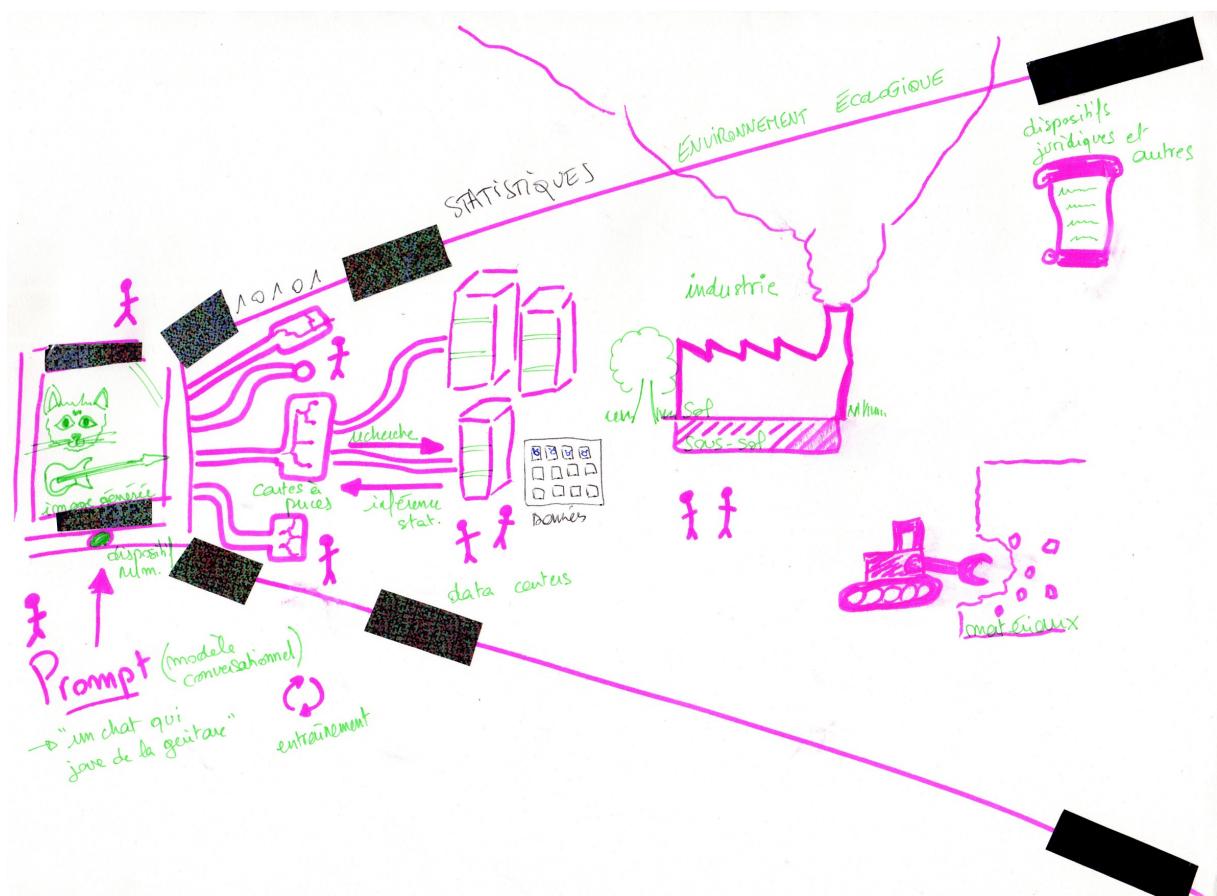


FIGURE 13 – Un dessin du fonctionnement de l'IA

Wiki

L'histoire de l'IA

Une histoire multiple de l'intelligence artificielle

Dans cet article nous allons aborder l'histoire de l'« intelligence artificielle ».

Même si ce vocable est apparu à une date précise, en 1956, son histoire est bien plus large. En effet, outre l'idée d'un être artificiel doué d'intelligence qui remonte à l'antiquité, ce que nous avons aujourd'hui et que nous appelons « intelligence artificielle » s'inscrit dans une histoire scientifique, technique, sociale, économique et culturelle.

L'idée de l'intelligence artificielle

Les prémisses du vocable se retrouvent dans une série de mythes antiques comme [Prométhée](#), [Pygmalion](#) ou encore [Talos](#). D'autres figures vont suivre, comme le [golem](#) aux environs du 2ème siècle : un humanoïde fait d'argile créé pour aider ou défendre son créateur d'après le Talmud. Ce mythe est revisité bien plus tard au 19ème siècle avec le [Maharal de Prague](#). Le [Chapitre 1](#) du livret explore plus avant cette thématique et propose de *démythologiser* ce concept, de comprendre comment il anime et structure notre imaginaire.

La cybernétique

C'est toutefois lors de la seconde guerre mondiale qu'une idée bien plus précise va se former. L'informatique en est à ses débuts, la technique alimente les théories scientifiques qui à leur tour débouche sur de nouvelles applications. Cela transforme petit à petit le regard porté sur l'être humain et en particulier sur ses capacités intellectuelles qui sont alors comparées à des logiciels informatiques. Tout cela se formalise par le mouvement de la *cybernétique* en 1947. Le mathématicien [Norbert Wiener](#) décrit le résultat de conférences rassemblant des scientifiques d'horizons très divers (mathématiciens, anthropologues, économistes, psychologues, etc) portant sur l'étude des mécanismes d'information des systèmes complexes. Il s'agit d'étudier comment l'information caractérise un système.

Par exemple, si nous analysons sous le prisme de l'information la conduite d'une voiture, l'information de départ est le code de la route qui définit toute une sorte de modalité de conduite (limitation de la vitesse, priorité de droite, etc) ainsi que l'infrastructure, les carrefours, les routes et leurs limitations respectives, etc. Lors d'un déplacement un conducteur va recevoir toute une série d'informations : d'autres usagers sur son chemin, des feux verts, rouges oranges, des travaux, un bon ou mauvais état de la route et ainsi de suite. Il va alors en fonction de ces informations *adapter* sa manière de conduire. Autrement dit, il va transformer sa manière de conduire (initialement dictée par le code de la route et l'infrastructure routière) par les informations pratiques du système. Par sa présence, il va modifier certaines informations du système qui va à son tour modifier sa conduite, c'est-à-dire qu'il a un impact sur la circulation réelle. Cette boucle est appelée un phénomène de [rétroaction](#). Il

est donc possible de décrire la dynamique de ce système uniquement par la perspective de l'information. C'est en tout cas l'hypothèse de la cybernétique.

De par cette formalisation, il devient imaginable qu'un être humain puissent être réduit à un être d'information et donc, théoriquement et scientifiquement comparable à un ordinateur. En 1956, [John McCarthy](#) propose le vocable « Intelligence artificielle » qui sera défini par son collègue [Marvin Minsky](#) comme étant :

Un programme informatique qui est utilisé dans des tâches qui sont effectuées de manière efficace par des êtres humains et qui demandent un haut niveau de processus mentaux comme : l'apprentissage perceptuel, l'organisation de la mémoire et le raisonnement critique.

La cybernétique aura un impact important sur le monde scientifique et permettra l'élaboration des sciences cognitives, des thérapies systémiques, des théories biologiques de l'auto-organisation, et évidemment des sciences autour de l'intelligence artificielle.

L'influence ne sera pas que scientifique ou technique, la cybernétique va également s'infiltrer dans la culture, en particulier avec des termes comme le *cyborg*, le cyberspace, le cyberpunk et autres mots avec un préfixe *cyber*.

L'histoire technique et scientifique

L'histoire des techniques est souvent prise par le prisme des usages. Ainsi la voiture électrique succèderait à la voiture thermique, le téléphone portable au téléphone fixe, etc. Même si cela est pertinent, par exemple pour comprendre les impacts culturels (comment les pratiques évoluent), il ne rend pas compte de la réalité technique de ces objets. Par exemple, techniquement, dans son fonctionnement, une voiture électrique est bien plus proche d'une machine à laver que d'une voiture thermique. Il y a donc une claire distinction entre histoire des pratiques humaines et histoire des techniques.

S'intéresser à l'évolution technique permet, entre autres, de comprendre quel contexte économique a été nécessaire pour produire tel ou tel objet, ce qu'il a mobilisé en terme de ressources et d'organisations (logistiques, politiques,...).

Pour exposer cette perspective technique et les enseignements qui lui sont propres, nous allons ici essayer de bien séparer les points de vues.

L'histoire des techniques est rythmée par deux types de progression ¹.

- La première continue et mineur : elle reflète les petits changements et améliorations d'un objet technique existant (meilleure efficacité d'un moteur, du rendement, etc).
- La deuxième, discontinue et majeur : elle voit l'arrivée de nouveaux objets techniques (l'invention de la machine à vapeur, l'ordinateur, etc) ou une amélioration très forte d'un objet technique existant qui a un impact important sur son fonctionnement (le circuit intégré ou puce électronique en informatique).

De plus, l'invention et l'amélioration d'un objet technique se fait également par des découvertes et maîtrises d'autres objets techniques. Par exemple, les puces électroniques ont pu être inventées grâce à la miniaturisation des composants électroniques. Ce n'est donc pas l'objet technique seul qui doit être considéré, mais tout un écosystème technique pour comprendre leurs évolutions.

1. Gilbert SIMONDON, Irlande SAURIN, Nathalie SIMONDON, *Du mode d'existence des objets techniques*, Nouvelle éd. 2024 augmentée, Paris, Flammarion, coll. « Champs », 2024, p. 52.

Cependant, comme nous l'avons vu, l'intelligence artificielle ne représente pas un objet technique spécifique, mais plutôt une idée. Ce sont en effet des objets techniques complètement différents d'un point de vue technique qui forme son histoire et que nous allons décrire ici.

Les automates

Nous pouvons commencer cette histoire par les automates, ces dispositifs qui permettent de reproduire une séquence d'actions pré-déterminées, et ce en toute autonomie. Dès l'Égypte antique, nous retrouvons des statues animées dans le cadre d'événements religieux et vers la fin de cette civilisation, des automates plus avancés comme des corbeaux qui chantent². La Renaissance voit apparaître un nouvel intérêt pour les automates dans un but de divertissement. En 1744, Jacques de Vaucanson crée un « [canard artificiel](#) » qui peut manger, digérer, cancaner et simuler la nage³. Ces automates sont mécaniques. Ce sont donc des ressorts et des engrenages qui sont à l'œuvre dans ses machines. Technique nous nous rapprochons plus des horloges mécaniques que des robots actuels ou des programmes informatiques.

L'arrivée de l'informatique

Avec la seconde guerre mondiale arrive les premiers ordinateurs modernes, des machines à calculer électroniques *programmable*, utilisées pour déchiffrer les communications ennemis. Nous retrouvons deux scientifiques [Alan Turing](#) et [John von Neumann](#) qui participent tous deux au mouvement de la cybernétique. Fort des nouvelles possibilités techniques, c'est l'exploration d'un nouveau monde qui s'offre aux scientifiques et aux techniciens, penser de nouveaux [algorithmes](#), les tester et les mettre en application.

C'est ainsi que naît en 1951 les [Les réseaux de neurones artificiels](#) artificiels, une imitation informatique simplifiée d'un [neurone](#), la cellule à la base de notre système nerveux et de notre cerveau. Ces systèmes à la base du [deep learning](#) ou apprentissage profond donneront les grands modèles de langage actuels, les générateurs de textes. Cependant, en pratique, ils seront peu utilisés jusqu'aux années 2000.

Toujours dans la recherche de la reproduction des fonctions cognitives, d'autres algorithmes sont testés. Par exemple en 1966 avec ELIZA, un programme conçu pour converser avec une personne, à l'instar des agents conversationnels actuels. Sa particularité est de simuler un psychotérapeute qui ne fait « que » reformuler les affirmations de son interlocuteur. Le logiciel est très intéressant, mais reste très limité.

Pour cette époque, l'optimisme est à la hauteur de ces nouveaux résultats. Les chercheurs prédisent prochainement un avenir radieux pour l'IA⁴ :

- En 1958, H. Simon et Allen Newell : « d'ici dix ans un ordinateur sera le champion du monde des échecs » et « d'ici dix ans, un ordinateur découvrira et résoudra un nouveau théorème mathématique majeur ».
- En 1965, H. Simon : « des machines seront capables, d'ici vingt ans, de faire tout travail que l'homme peut faire ».
- En 1967, Marvin Minsky : « dans une génération [...] le problème de la création d'une 'intelligence artificielle' [sera] en grande partie résolu ».

2. Décrit par [Vitruve](#), un architecte Romain du 1er siècle av. J.-C..

3. Jacques de Vaucanson, *Le Mécanisme du flûteur automate présenté à messieurs de l'Académie royale des sciences*, 1738

4. Voir l'[article wikipedia](#) sur l'histoire de l'intelligence artificielle.

- En 1970, Marvin Minsky (dans le magazine *Life*) : « Dans trois à huit ans nous aurons une machine avec l'intelligence générale d'un être humain ordinaire ».

L'effervescence et la fascination expliquent bien entendu cet optimisme, mais pas seulement. La recherche coûte cher et prend du temps, il ne s'agit pas uniquement de se convaincre que la direction des recherches est bonne, mais également de convaincre ceux qui investissent dans celles-ci.

Un premier essoufflement de la recherche se produit pendant les années 1970. Ce n'est pas si simple de faire de l'IA. La technique a ses limites, des contraintes matérielles fortes. Le problème est large et pas très bien défini, l'idée de l'intelligence artificielle se confronte à la matérialité du monde. Les problèmes techniques sont multiples. Pour résoudre des problèmes, l'être humain est confronté à un nombre très important d'informations à traiter. Les relations possibles à faire entre elles sont astronomiques, il faut donc filtrer et synthétiser. Cela demande un stockage et un traitement de l'information, une puissance de calcul bien trop importante pour l'époque. Mais ce n'est pas tout, la logique et les mathématiques ont leurs propres contraintes avec lesquels il faut composer lorsqu'il s'agit de transposer nos fonctions cognitives.

À partir de 1980, les ambitions sont revues à la baisse et la recherche se concentre sur la résolution de tâches beaucoup plus ciblées avec les systèmes experts. Ceux-ci répondent à des questions dans un contexte spécifique, par exemple en médecine pour l'établissement d'un diagnostic de maladies du sang. Ces logiciels sont construits à partir d'une connaissance spécifique et de règles limitées. Ils permettent à partir de données extérieures de déduire un résultat utile, à la prise de décision ou à des recommandations par exemple.

Les résultats de ces nouveaux algorithmes attirent à nouveaux les investisseurs et révolutionnent certains secteurs économiques spécifiques. Mais très rapidement, les limites se font à nouveau sentir et la recherche connaît un nouvel essoufflement. En particulier, le besoin de connaissances très larges, de savoirs « généraux » pour résoudre des problèmes.

La fin des années 80 voit également l'arrivée de l'ordinateur personnel qui impacte profondément l'économie informatique. De grosses entreprises font faillites, les techniques et le rapport à l'informatique changent, se spécifient et se concrétisent. L'idée et les promesses de l'IA s'éloignent d'autant plus. Le vocable IA est petit à petit abandonné pour des termes plus spécifiques : robotique, reconnaissance vocale, de formes, exploration de données (*data mining*), apprentissage machine, etc. À ce moment, ceux qui parlent de l'IA sont plutôt vus comme des rêveurs, et est un repoussoir pour les investisseurs dans la recherche « sérieuse ».

Malgré la déconvenue, les techniques progressent de manière continue. En particulier, la puissance de calcul augmente de manière exponentielle ([Loi de Moore](#)). Des choses impossibles dans les années 50 ou 60, deviennent envisageables. De plus, les années 90 voient également la démocratisation d'Internet. Ce nouveau lieu, le cyberspace, va progressivement offrir une nouvelle source de connaissances exploitables par les algorithmes. Ces deux aspects combinés, puissance de calcul et connaissances exploitables vont permettre la maturité d'une catégorie d'algorithme d'IA, l'apprentissage machine ou *machine learning* : l'utilisation des [statistiques](#) pour apprendre et adapter le comportement de l'algorithme.

Dans les années 2000

L'année 2000 voit l'apogée de la [bulle Internet](#), à nouveau des promesses technologiques ne sont pas tenues. S'en suit une reconfiguration économique qui voit l'émergence de grandes

sociétés commerciales les **GAFAM**. Celles-ci sont omniprésentes dans le monde numérique. Pour Facebook et Google, c'est une nécessité. Elles ont besoin d'accumuler un maximum d'informations sur leurs utilisateurs dans le but de créer un profil qui servira à du ciblage publicitaire, la source de leurs revenus. La création de ce profil est fait avec un algorithme statistique qui « apprend » à différencier, à reconnaître et à associer les personnes en fonction des données qu'ils produisent sur Internet ou plus tard avec leur smartphone. Pour cela, les données sont traitées dans des « fermes » de calcul où des centaines, voire des milliers d'ordinateurs traitent en permanence les informations des utilisateurs. Et ça fonctionne très bien, le ciblage publicitaire est très pertinent et ces entreprises engrangent des milliards.

Le succès de ces deux entités commerciales ravivent l'attrait des investisseurs et la course à l'IA. En 2012, des chercheurs montrent les progrès importants qu'ils ont pu obtenir avec un algorithme statistique, l'**apprentissage profond**. Celui-ci est une amélioration substantielle des **réseaux de neurones artificiels**. En particulier, ils excellent dans la reconnaissance et le traitement d'image. Ces calculs sont particulièrement rapides avec un type de matériel, les **processeurs graphiques** qui sont optimisés sur le travail de l'image (pour les jeux vidéos par exemple).

Le vocable « intelligence artificielle » revient à la mode et est poussé par Google et Facebook, entre autres. L'espace médiatique est témoin d'une véritable tension entre chercheurs publics et privés. Les sociologues Anne Bellon et Julia Velkovska montrent à la suite d'une analyse des médias que cette « *“vague de l'IA” semble plutôt marquer la montée en puissance des acteurs privés (grandes entreprises du numérique) dans le domaine...»*⁵ ».

Les grands modèles de langage

Les chercheurs ne s'arrêtent pas pour autant au traitement et à la génération d'images avec l'**apprentissage profond**, ils essayent également de générer et traiter du texte. Pour cela, ils modélisent le langage humain avec cet algorithme, c'est-à-dire qu'ils essayent de réduire le langage, sa grammaire, ses styles, son vocabulaire et le savoir qui va avec, dans un objet mathématique (**Les statistiques**) produit à partir de l'**apprentissage profond**. Les essais sont d'abord infructueux, mais en 2017 une équipe de chercheur de Google propose une légère amélioration⁶ de ce qui a déjà été développé. Celle-ci a trois atouts, elle est simple, elle donne de bons résultats et elle est parallélisable. Le dernier point est fondamentale car il signifie qu'il n'est pas nécessaire d'avoir une machine extrêmement puissante pour traiter toutes les informations, tout le savoir. Il est possible de répartir le traitement entre plusieurs machines moins puissantes. Et le savoir, la connaissance humaine est grande, ce sont donc des centaines de milliers de textes qui sont utilisés pour construire ce modèle, ce *grand modèle de langage*.

Il faudra alors trois années de recherche pour amasser des textes, créer un modèle et optimiser un logiciel de type agent conversationnel performant. Mais c'est en novembre 2022 que le grand public peut enfin y accéder avec la publication de ChatGPT par OpenAI.

Les résultats sont tels qu'ils créent, à nouveau, un optimisme sur l'*avenir* de l'IA. Les investisseurs ne manqueront pas d'y investir des milliards, encore aujourd'hui en 2025. À l'heure actuelle, il n'existe toujours pas de modèle économique bénéficiaire chez les sociétés

5. Conclusion de (Anne BELLON, Julia VELKOVSKA, « L'intelligence Artificielle Dans l'espace Public : Du Domaine Scientifique Au Problème Public : Enquête Sur Un Processus de Publicisation Controversé » [en ligne], *Réseaux*, n° 4, 2023, URL : <https://www.cairn.info/revue-reseaux-2023-4-page-31.htm?ref=doi>, consulté le 27 mai 2025).

6. Voir l'article (Ashish VASWANI, Noam SHAZER, Niki PARMAR, et al., « Attention Is All You Need » [en ligne], arXiv, 2023, URL : <http://arxiv.org/abs/1706.03762>, consulté le 18 novembre 2025).

L'histoire de l'IA

qui commercialisent les agents conversationnels, uniquement des perspectives vagues à moyen ou long terme. L'amélioration des performances commence à diminuer malgré les budgets colossaux et un nouvel essoufflement se fait sentir d'où la crainte d'une nouvelle bulle économique. L'histoire nous le dira.

Fonctionnement de l'IA

Comment fonctionne une intelligence artificielle, au juste ?

- Algorithmes
- Les probabilités
- Les statistiques
- Apprentissage machine
- Les réseaux de neurones artificiels

Algorithme

Les algorithmes

Les algorithmes sont une suite d'instructions et d'opérations précises. Une recette de cuisine en est un si celle-ci est suivie à la lettre. La série d'instructions que donne un GPS pour arriver à une destination (rouler pendant 5 km, puis tourner à droite, prenez la deuxième à gauche, etc) est également un algorithme.

L'algorithmie étudie les manières de résoudre un problème, d'arriver à un résultat grâce à des instructions ou opérations. En effet, il existe souvent plusieurs manières de faire, parfois plus rapides, parfois plus simples, parfois plus compréhensibles que d'autres. C'est particulièrement important en informatique où, chaque instruction, chaque calcul demande des ressources. Il est donc courant en informatique d'*optimiser* un algorithme pour qu'il soit le plus rapide, par exemple trier des nombres dans une colonne d'un tableau ou stocker une image, une vidéo pour qu'elle occupe le moins d'espace sur le disque dur.

En fait, les algorithmes sont au cœur du fonctionnement d'un ordinateur, puisque celui-ci a été précisément construit pour exécuter des logiciels, c'est-à-dire une série d'instructions et d'opérations.

Histoire des algorithmes

Le terme « algorithme » vient du nom d'un Perse du 9^e siècle *Al-Khwârizmî*, qui écrivit plusieurs traités mathématiques. Trois siècles plus tard, un moine reprit son nom pour parler du concept actuel, les algorithmes.

C'est dès 1842 qu'[Ada Lovelace](#) crée le premier algorithme de programmation de l'histoire.

Algorithme explicite et implicite

Pour mieux comprendre les algorithmes utilisés en IA, nous pouvons faire une distinction entre deux types d'algorithmes. D'une part, les algorithmes explicites, dont la logique des instructions est définie à l'avance, connue par les concepteurs. D'autre part, les algorithmes implicites, qui ne décrivent pas toutes les instructions et opérations à effectuer, mais plutôt la manière dont ces instructions doivent être faites. Pour cela, ils utilisent des statistiques pour établir les futures instructions à exécuter. On les appelle également algorithmes d'[apprentissage machine](#). Les instructions statistiques pouvant être très complexes, il peut être difficile, voire impossible, de comprendre la logique finale de l'algorithme.

Par exemple, les recommandations de vidéos sur YouTube suivent un algorithme implicite. À partir des données personnelles des utilisateurs, des calculs statistiques sont menés pour établir quelles sont les « meilleures » propositions selon des critères spécifiques, comme faire en sorte que le visiteur reste le plus longtemps possible sur la plateforme. La façon d'y arriver

Algorithme

n'est pas défini à l'avance, ce sont les statistiques qui vont l'établir et non les concepteurs. Si les recommandations sont souvent des vidéos courtes et à fort contenu émotionnel, ce n'est qu'un résultat statistique, et non une volonté première des techniciens derrière ces algorithmes. Toutefois, rien n'empêche les concepteurs d'ajouter d'autres critères dans un second temps.

Apprentissage machine

L'apprentissage machine, ou *machine learning*, est un type d'**algorithme** qui a pour but de comprendre certaines propriétés d'un ensemble de données à partir de méthodes statistiques. Il va apprendre à les situer ou les catégoriser les unes par rapport aux autres. De la connaissance de cette relation, il va pouvoir situer de nouvelles données qu'il reçoit ou prédire de nouvelles données en conservant le même type de relation.

En pratique, ce genre d'algorithmes va par exemple pouvoir établir le profil des personnes sur les réseaux sociaux. Il va observer les personnes qui ont des contacts fréquents pour établir les groupes sociaux, puis établir ce qui statistiquement met en relation ces gens : un hobby, l'âge, le lieu géographique, le niveau d'études, etc. À partir de cela, il est possible d'une part d'établir rapidement le profil d'une nouvelle personne, et d'autre part de proposer des mises en relation entre personnes qui sont les plus susceptibles d'être acceptées.

En ce qui concerne les **grands modèles de langage**, l'apprentissage se fait à partir de textes existants et consiste à établir les liens entre les mots, la grammaire, le style, les sujets abordés, etc. Fort de cette connaissance, il peut analyser des textes et produire un texte **probable** en fonction d'un contexte précis, un prompt par exemple.

Histoire de l'apprentissage machine

Le premier **algorithme** d'apprentissage machine est le **réseau de neurones** artificiels, inventé en 1943. Le but est de reproduire le fonctionnement des cellules nerveuses et, *in fine*, le cerveau humain. Toutefois, la première application connue d'un logiciel d'apprentissage machine est un programme de jeu de dames datant de 1955. Il n'utilisait pas un **réseau de neurones**, mais des statistiques beaucoup plus simples issues des parties.

Ce n'est toutefois qu'en 1959 que le créateur de ce programme introduit l'expression *machine learning*, ou apprentissage machine en français.

Neurone artificiel

Un neurone artificiel, dans les **réseaux de neurones**, est simplement un algorithme, une équation.

Au départ, c'est une représentation mathématique d'un neurone biologique comme on le représentait en 1943 lors de son invention par McCulloch et Pitts,¹ mais le développement des **réseaux de neurones artificiels** a motivé l'adaptation de ce modèle pour obtenir des réseaux plus efficaces, pas forcément plus proches du fonctionnement biologique.

Il y a plusieurs équations possibles, mais le principe est toujours le même :

1. On prend des entrées (des nombres qui peuvent représenter une information visuelle, audio, textuelle... ou la sortie d'un autre neurone dans un réseau!)
2. On les additionne avec un poids, une pondération, différent pour chaque entrée (le poids est fixe, il fait partie de la définition du neurone)
3. On vérifie si cette *somme pondérée* dépasse un seuil (lui aussi fixe et définissant le neurone), avec une fonction choisie par les concepteurs du neurone (seuil, sigmoïde, unité... voir)²
4. Si le seuil est dépassé, la sortie du neurone « s'allume », sinon non ! (On peut aussi avoir une sortie *continue*, pas juste « 1 » (allumé) ou « 0 » (éteint)).

1. « Neurone formel » [en ligne], *Wikipédia*, 2023, URL : https://fr.wikipedia.org/w/index.php?title=Neurone_formel&oldid=209205599, consulté le 13 mai 2025.

2. *Ibid.*

Les réseaux de neurones artificiels

Un réseau de neurones est un modèle informatique qui s'inspire du fonctionnement des cellules nerveuses biologiques et de la structure de leurs associations. Ce sont des **algorithmes d'apprentissage machine**.

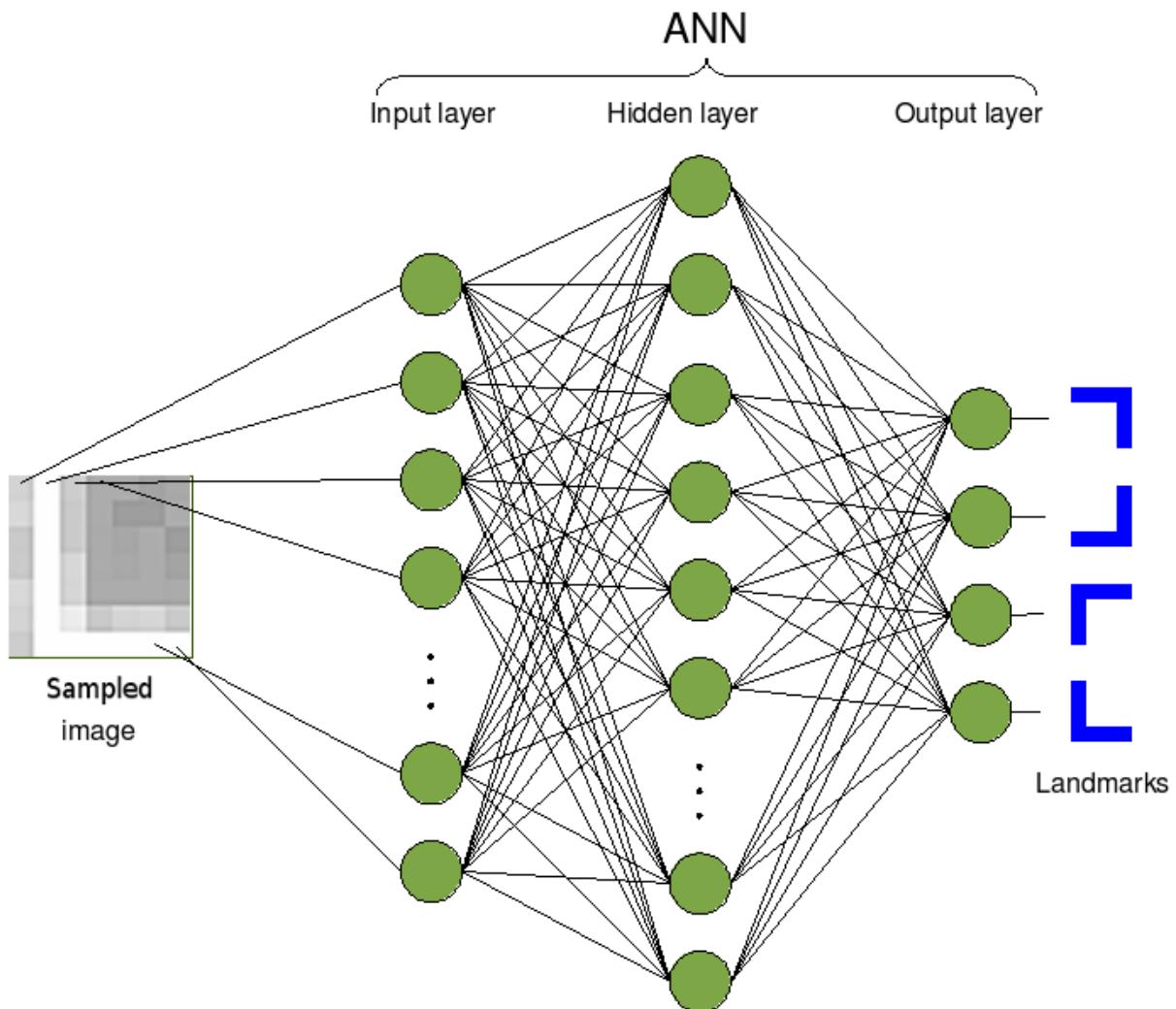


FIGURE 1 – Réseau de neurones

https://commons.wikimedia.org/wiki/File:Artificial_neural_network_image_recognition.png

Il est constitué de **neurones artificiels** reliés les uns aux autres et organisés en couches. La première reçoit un signal, constitué d'un ou plusieurs neurones, traite l'information et renvoie le résultat à la deuxième couche. Chacun des neurones va envoyer son résultat à l'ensemble des neurones de la couche suivante. La deuxième couche va réitérer l'opération et transmettre le résultat à la couche suivante, et ainsi de suite. La dernière couche va collecter

un résultat final, somme des transformations des couches précédentes. Selon le problème à résoudre, les neurones vont transformer les données de manières différentes. Manières qui vont également évoluer en fonctionnement du résultat final obtenu.

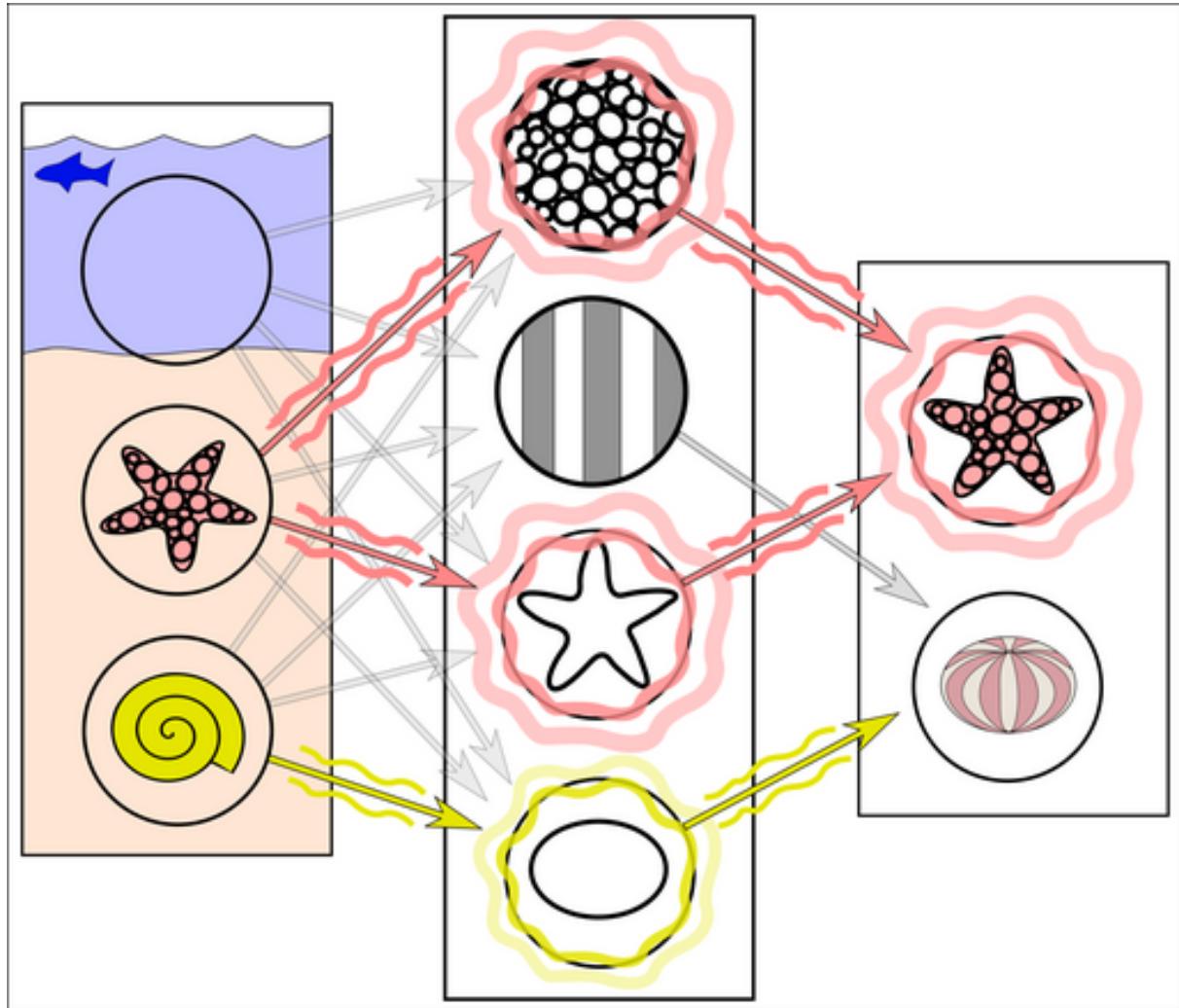


FIGURE 2 – Réseau de neurones pour une image

https://commons.wikimedia.org/wiki/File:Simplified_neural_network_example.png

Par exemple, lorsqu'il s'agit d'apprendre à identifier une image, le signal prend la forme des différents paramètres de l'image, comme la valeur de chacun des pixels la constituant. Chacun des neurones suivants va recevoir une combinaison de chaque neurone et appliquer un traitement, par exemple, privilégier les pixels proches qui ont la même teinte, afin d'extraire des formes. Les résultats sont renvoyés dans la couche suivante qui va, par exemple privilégier les formes qui se ressemblent. Le processus peut ainsi continuer jusqu'à la dernière couche qui doit établir le type d'image. En fonction du résultat obtenu, les transformations que font les neurones peuvent être modifiées, optimisées et réappliquées à une autre image. Le réseau apprend ainsi, image après image à reconnaître de quoi est constituée l'image.

Lorsqu'il y a plusieurs couches intermédiaires, l'algorithme est appelé réseau de neurones profond ou apprentissage profond (*Deep Learning* en anglais).

Dans les grands modèles de langage

Les **grands modèles de langage** comme ceux utilisés par ChatGPT fonctionnent avec des réseaux de neurones profonds extrêmement grands. Le nombre de couches est de l'ordre d'une centaine et celui de neurones par couche se compte en *centaines de millions*.

Statistiques et grands modèles de langage

Le milieu des statistiques et les grands modèles de langage

Alors que **les probabilités** étudient le phénomène aléatoire en lui-même, les statistiques s'occupent de recueillir, traiter et interpréter des données variables, que ce soit la distribution des salaires dans la population ou les observations météorologiques.

À partir de données collectées, les statistiques vont essayer de comprendre quel type de hasard nous avons à faire, et quels sont les types de phénomènes aléatoires et non aléatoires qui permettent de produire ou modéliser nos observations.

Les grands modèles de langage

Les grands modèles de langage, ou LLM pour *Large Language Model*, sont un type de modèle statistique qui, à partir d'une énorme quantité de textes, cherche à produire un texte **probable**. Pour cela, ils s'appuient sur toute une série d'**algorithmes** pour collecter les textes, les traiter et déduire les relations statistiques entre l'agencement des mots et des phrases d'un texte. Ces relations sont établies par un **algorithme** basé sur des **Les réseaux de neurones artificiels** profonds. Pour établir ces relations, il se base sur la « signification » des mots. Par exemple, une pomme et une poire auront une signification proche au contraire d'une pomme et d'un avion. En analysant les textes et la proximité des mots entre eux, l'**algorithme** va établir statistiquement quels sont les mots que l'on retrouve ensemble et selon quel contexte. Il est plus fréquent de parler de pomme dans un texte, si celui-ci parle de verger plutôt que d'engin spatial. Ainsi le LLM va établir une signification statistique des mots. Cela va lui permettre de se créer une représentation du langage sous la forme d'un objet mathématique.

Ainsi, pour fonctionner, les LLM actuels vont regarder l'ensemble du texte précédent, une question par exemple, en déduire son contexte et prédire le mot suivant. Ils vont recommencer l'opération jusqu'à avoir un texte qui les satisfait ; satisfaction déduite de tous les textes qu'ils ont précédemment analysés, c'est une satisfaction moyenne.

Il existe d'autres méthodes que la modélisation du mot suivant. En effet, il est possible d'analyser l'entièreté d'une phrase en partant des deux côtés de la phrase et de voir comment ils se rejoignent « statistiquement ». Cette méthode est par exemple utilisée pour la traduction afin d'être au plus près de la signification du texte à traduire. Dans une deuxième étape, la traduction est générée mot à mot, de la même manière que précédemment.

De manière générale, toutes les techniques employées pour un LLM doivent être finement paramétrées pour arriver à un résultat qui satisfait son concepteur, l'entreprise qui le conçoit. La quantité et la qualité des textes qu'il analyse est également primordiales. Les textes sont-ils bien orthographiés ? Agréables à lire ? Plutôt formels ou informels ? Longs ou courts ? Orientés politiquement de manière évidente ou pas ? La qualité du résultat statistique en dépend.

Alignment

Le modèle construit à partir d'une grande quantité de textes sans autres manipulations développe un comportement purement statistique. C'est-à-dire qu'il n'a pas ou peu de résistance à produire un texte contraire à une certaine éthique ou morale. Pour pallier ce souci, les techniciens vont donner au modèle des phrases spécifiques qui auront une extrême importance dans sa considération du langage. On pourra par exemple retrouver des phrases comme « Tu es une IA qui ne doit pas nuire ou causer du tort », de même que « Tu es une IA qui doit être utile et honnête ». Ceci permet de développer un comportement plus en phase avec la volonté du concepteur. Pas grâce à un raisonnement logique comme nous pourrions le croire, mais juste par le rapprochement statistique avec les phrases qu'il a déjà analysées qui sont « utiles, honnêtes et ne doivent pas causer de tort ». Il est donc tout à fait possible de contourner ce mécanisme, entre autres en « convaincant » l'IA que la réponse est bien « utile, honnête et ne cause pas de tort ». Il est donc difficile, par la nature statistique des LLM, de résoudre ce problème d'[« alignment »](#). D'autres techniques sont utilisées, mais en combinaison avec un autre algorithme que le LLM, par exemple en recherchant dans le texte en cours d'élaboration des termes « interdits », à l'instar de DeepSeek un agent conversationnel chinois, lorsqu'il parle de sujets proscrits.

Les probabilités

Le milieu des probabilités

Les contraintes pour construire les grands modèles de langage, comme ChatGPT, sont de différents ordres. Essayons dans cet article d'aborder une contrainte fondamentale et hors du champ humain, les mathématiques, en particulier les probabilités qui sont au cœur de ces objets techniques.

Les probabilités désignent l'étude des phénomènes aléatoires, par exemple les résultats possibles d'un lancer de dés. Sa compréhension permet, entre autres, de prédire comment des phénomènes aléatoires distincts se combinent et ce qu'ils produisent. Quelles sont les probabilités d'avoir un six, si nous lançons deux, trois ou quatre dés ? Quelle est la probabilité d'avoir telle ou telle maladie étant donné le contexte génétique et environnemental ? En effet, les probabilités ne se mélangent pas « n'importe comment », elles suivent les rigides contraintes des mathématiques.

Une loi mathématique incontournable pour comprendre comment les probabilités s'additionnent est le [Théorème Central Limite](#) qui dit que le type de probabilité produite par une somme de phénomènes aléatoires finira toujours par avoir la même forme, une « Gaussienne » ou loi normale.

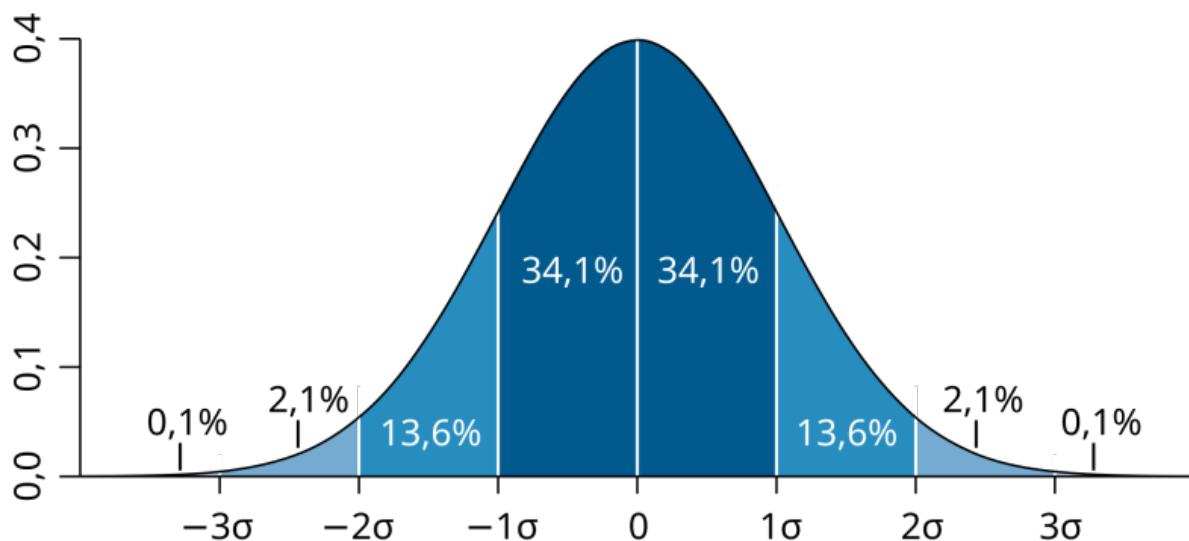


FIGURE 1 – La courbe d'une gaussienne)

Source : [https://commons.wikimedia.org/wiki/File:Standard_deviation_diagram_\(decimal_comma\).svg](https://commons.wikimedia.org/wiki/File:Standard_deviation_diagram_(decimal_comma).svg)

Cette forme représente comment les événements aléatoires se répartissent dans les différents résultats que ces événements peuvent prendre : plus la courbe est haute, plus l'événement

Les probabilités

est probable. Elle est caractérisée par sa courbure et sa symétrie ; les éléments les plus probables se trouvent au centre, la moyenne, et ceux plus rares dans les « queues » de la courbe. Prenons un exemple : la taille des femmes. Si nous mesurons la taille d'un nombre relativement important de personnes, nous allons obtenir une répartition des mesures similaire à une gaussienne. La moyenne en Belgique sera aux alentours de 164 cm, au centre de la courbe. La plupart des personnes se trouveront aux alentours, mais au plus nous nous éloignerons de la moyenne, au plus il sera rare de trouver une personne faisant cette taille.

Pourquoi est-ce important et comment cela constraint-il le milieu technique ? Parce qu'un technicien qui travaille avec de l'aléatoire aimerait parfois avoir une courbe plutôt asymétrique, c'est-à-dire que les déviations de sa machine soit plutôt d'un côté que d'un autre, par exemple si un côté est plus dangereux . Or il n'a pas le choix dans la forme des probabilités. Il ne peut agir que sur les phénomènes aléatoires sous-jacents à sa machine et ainsi chercher à réduire le hasard.

Dans les grands modèles de langage

Une des hypothèses des grands modèles de langage est de dire que le langage a un caractère probabiliste. C'est-à-dire que l'écriture d'un texte peut prendre plusieurs formes tout en gardant la même signification, il y a juste des manières plus probables que d'autres. Parfois l'ensemble des mots peut varier, parfois seulement certains si l'on veut garder le sens du texte.

En particulier, lorsqu'un grand modèle de langage veut prédire le mot suivant d'un texte, il ne prend pas le mot le plus probable, mais un ensemble de mot probable et va choisir, au hasard, un des mots pour constituer sa phrase. Ceci est mis en place pour que le texte paraisse plus « naturel ». Chaque LLM paramètre finement ce hasard : au plus le mot sera choisi au hasard, au plus le texte sera « créatif », mais au risque de dénaturer le sens, voire de produire quelque chose qui ne veut plus rien dire comme dans le cas des hallucinations.

Dans les algorithmes

Lancé par Hubert Guillaud et François-Xavier Petit (matrice.io) à travers l'association [Vecteur](#), [Danslesalgorithmes.net](#) est un média de réflexion et analyse des enjeux du numérique actuel, proposant des articles de grande qualité visant à *apporter à tous de la matière pour mieux comprendre les transformations sociales que produisent les données et les calculs sur la société*.¹

Hubert Guillaud, rédacteur en chef et journaliste principal du média, est spécialiste des questions techniques et de leurs impacts sur nos sociétés.

Il a notamment écrit l'essai *Les algorithmes contre la société*² et il intervient régulièrement en tant qu'expert dans plusieurs médias de réflexion autour du numérique.

1. Hubert GUILLAUD, « Avec Mes Excuses... Algorithmiques » [en ligne], 2024, URL : <https://hubertguillaud.wordpress.com/2024/07/09/avec-mes-excuses-algorithmiques/>, consulté le 9 octobre 2025.

2. Hubert GUILLAUD, *Les algorithmes contre la société* [en ligne], Paris, la Fabrique éditions, 2025, URL : <https://lafabrique.fr/les-algorithmes-contre-la-societe/>.