

Le défi de la divergence anthropologique

Philippe RIS

Éditeur Philippe Ris

Auteurs Ph. RIS assisté de Chat GPT

Tous droits réservés.

Septembre 2026

Table des matières

2. Intelligence artificielle : commençons par savoir de quoi nous parlons.....	2
Une expression née avant que l'objet existe vraiment.....	2
Parler de capacités cognitives plutôt que de chercher trop vite l'intelligence.....	2
De quoi une intelligence pourrait-elle être constituée ?.....	3
L'imprévisibilité ne suffit pas, mais elle mérite notre attention.....	4
Et l'intention ?.....	4
L'IA est aussi un miroir.....	5
Peut-être faut-il revenir à Turing.....	5
De l'IA à l'AGI : attention encore une fois aux mots.....	6
Après l'intelligence : l'hypernoèse.....	6
3. Références bibliographiques.....	7
1. Fondements : intelligence, cognition et intelligence artificielle.....	7
2. Le cerveau biologique comme système probabiliste.....	7
3. Complexité, émergence et imprévisibilité.....	7
4. Finalité, vivant et émergence de l'intention.....	8
5. Définitions institutionnelles et juridiques de l'IA.....	8
6. Augmentation humaine, posthumanisme et divergence anthropologique.....	8
7. Éthique de l'information et transformation numérique.....	9
8. Mes cours et travaux sur l'IA et la divergence anthropologique.....	9

2. Intelligence artificielle : commençons par savoir de quoi nous parlons

Il est difficile aujourd'hui d'échapper à « l'intelligence artificielle ». L'expression est partout : dans les médias, dans les entreprises, dans les discours politiques, dans la réglementation et jusque dans nos conversations quotidiennes.

Elle est pratique. Elle est devenue incontournable. Elle est aussi probablement assez mauvaise.

Il ne s'agit évidemment pas de proposer de supprimer le terme « intelligence artificielle ». Ce serait aussi vain que de vouloir rebaptiser Internet. Il faut simplement garder à l'esprit que les mots que nous employons façonnent notre manière de penser les objets qu'ils désignent. Or, dans « intelligence artificielle », le mot le plus problématique n'est peut-être pas « artificielle ». C'est d'abord « intelligence ».

Une expression née avant que l'objet existe vraiment

L'expression *Artificial Intelligence* s'impose avec le projet de recherche organisé à Dartmouth en 1956 autour de John McCarthy. La proposition préparatoire formulait une hypothèse particulièrement ambitieuse : les différentes dimensions de l'apprentissage et de l'intelligence pourraient être décrites suffisamment précisément pour qu'une machine puisse les simuler.

Près de soixante-dix ans plus tard, nous utilisons toujours le même terme. Mais nous l'appliquons à des objets extraordinairement différents : algorithmes de classification, reconnaissance d'images, moteurs de recommandation, modèles prédictifs, grands modèles de langage, systèmes génératifs, agents capables d'utiliser des outils ou systèmes robotiques.

Le paradoxe est que nous avons considérablement progressé dans la construction de ces machines sans avoir résolu la question initiale : **qu'est-ce exactement que l'intelligence ?**

Shane Legg et Marcus Hutter commencent ainsi leur travail sur l'« intelligence universelle » par un constat assez radical : l'un des problèmes fondamentaux de l'IA est précisément que personne ne sait réellement définir ce qu'est l'intelligence. Ils proposent alors de l'appréhender notamment par la capacité d'un agent à atteindre des objectifs dans une grande variété d'environnements.

D'autres conceptions insistent sur l'adaptation. Robert Sternberg décrit ainsi le comportement intelligent par la capacité à s'adapter intentionnellement à son environnement, à le transformer ou à en sélectionner un autre.

Il n'existe donc pas une propriété simple et universellement reconnue qui permettrait de tracer une frontière indiscutable entre ce qui est intelligent et ce qui ne l'est pas.

Parler de capacités cognitives plutôt que de chercher trop vite l'intelligence

C'est pourquoi il peut être utile de descendre provisoirement d'un niveau.

Plutôt que de nous demander si une machine « est intelligente », demandons-nous quelles sont ses **capacités cognitives**. Le changement paraît mineur. Il est en réalité important.

Cette manière de poser le problème rejoint notamment une distinction développée par le philosophe Daniel Dennett : celle entre **compétence et compréhension**. Dans *From Bacteria to Bach and Back*, Dennett insiste sur l'existence de *competence without comprehension* : un organisme ou un système peut manifester une compétence remarquable sans pour autant comprendre ce qu'il fait au sens où un humain dirait le comprendre.

Cette idée est particulièrement intéressante pour parler d'IA. Un système peut reconnaître une tumeur sur une image, traduire un texte, gagner aux échecs, écrire un programme informatique ou établir des relations entre des informations

sans que nous ayons besoin de décider préalablement s'il « comprend » la médecine, le français, les échecs ou l'informatique.

Dennett va même plus loin : la compréhension elle-même pourrait être considérée comme le résultat de l'organisation progressive de compétences qui, prises isolément, ne comprennent rien. Autrement dit, nous commettons peut-être une erreur lorsque nous imaginons qu'il faut d'abord comprendre pour être compétent. L'évolution biologique suggère précisément le chemin inverse : **la compétence peut précéder la compréhension**.

On retrouve une intuition assez proche, par un autre chemin, chez Douglas Hofstadter. Ses travaux avec le *Fluid Analogies Research Group* cherchent moins à isoler une mystérieuse faculté générale appelée « intelligence » qu'à comprendre certains mécanismes élémentaires de la pensée : reconnaître des structures, former des concepts, catégoriser, établir des analogies et surtout modifier ces catégories en fonction du contexte. Dans *Fluid Concepts and Creative Analogies*, l'analogie et la reconnaissance de structures occupent ainsi une place centrale dans l'étude des mécanismes de la pensée.

L'approche de Hofstadter est particulièrement intéressante ici parce qu'elle est elle aussi **émergentiste**. Dans les architectures expérimentales développées autour de ses travaux, de multiples petites opérations « subcognitives » contribuent en parallèle à faire émerger des structures mentales cohérentes. L'intelligence n'a donc pas nécessairement à être recherchée dans un composant particulier du système : elle peut apparaître dans l'organisation et l'interaction de processus beaucoup plus simples (NB : les réflexions d'Hofstadter avaient inspirés mes propres travaux de doctorat).

Dennett et Hofstadter nous invitent ainsi, chacun à leur manière, à nous méfier d'une conception trop monolithique de l'intelligence. Celle-ci pourrait être moins une propriété mystérieuse que possède ou ne possède pas un système qu'une **organisation dynamique de compétences cognitives**.

C'est précisément l'intérêt de parler de capacités cognitives.

Une capacité cognitive peut être observée et, dans une certaine mesure, évaluée sans avoir à trancher préalablement la question de l'intelligence. Un système peut percevoir, reconnaître, mémoriser, apprendre, associer des informations, produire une représentation, généraliser, raisonner, prévoir, communiquer ou sélectionner une action.

Cette approche permet surtout de ne pas faire de l'intelligence humaine la définition implicite de toute intelligence possible.

Le droit européen a d'ailleurs adopté une approche assez pragmatique. L'AI Act ne cherche pas à définir philosophiquement « l'intelligence ». Il définit un « système d'IA » à partir de caractéristiques fonctionnelles : système automatisé, degré variable d'autonomie, éventuelle capacité d'adaptation, inférence à partir d'entrées et production de prédictions, contenus, recommandations ou décisions.

Le droit peut donc réglementer l'IA sans avoir résolu la question philosophique de l'intelligence. Nous pouvons faire de même.

De quoi une intelligence pourrait-elle être constituée ?

Il faut néanmoins essayer d'aller plus loin. Si nous ne savons pas parfaitement définir l'intelligence, nous pouvons identifier un certain nombre de propriétés qui semblent participer à ce que nous appelons communément un comportement intelligent.

La liste suivante ne constitue donc pas une définition. Elle constitue plutôt une **grille de lecture** :

- **Percevoir** : acquérir des informations sur soi-même ou son environnement.
- **Représenter** : construire une représentation interne suffisamment opératoire du réel.
- **Mémoriser** : conserver des informations et pouvoir les réutiliser.
- **Apprendre** : modifier son fonctionnement à partir de l'expérience ou d'informations nouvelles.

- **Associer et abstraire** : établir des relations entre des éléments et construire des catégories ou concepts dépassant les observations particulières.
- **Raisonner** : combiner des informations pour produire des inférences ou résoudre des problèmes.
- **Généraliser** : mobiliser un apprentissage dans des situations qui ne sont pas strictement celles dans lesquelles il a été acquis.
- **Anticiper** : construire des représentations d'états futurs possibles.
- **Décider** : sélectionner une réponse ou une action parmi plusieurs possibilités.
- **S'adapter** : modifier son comportement lorsque l'environnement change.
- **Communiquer** : échanger des représentations avec d'autres agents et éventuellement intégrer leurs réponses.
- **Agir** : transformer directement ou indirectement son environnement.
- **Évaluer** : comparer les conséquences d'une action à un état recherché ou à des critères.
- **Arbitrer** : gérer plusieurs objectifs ou contraintes éventuellement contradictoires.
- **Produire de la nouveauté pertinente** : élaborer des réponses qui ne sont pas la simple reproduction explicite d'une règle préétablie.
- **Disposer d'une certaine autonomie** : pouvoir enchaîner certaines de ces opérations sans qu'un opérateur extérieur détermine chaque étape.

La liste pourrait être discutée, complétée et réorganisée. C'est précisément son intérêt. Elle montre surtout que l'intelligence n'est probablement pas un interrupteur qui serait allumé chez l'humain et éteint chez la machine. Elle ressemble davantage à **une architecture de capacités qui interagissent entre elles**.

La question devient alors beaucoup plus intéressante : à partir de quel niveau d'intégration de ces capacités acceptons-nous de parler d'intelligence ?

L'imprévisibilité ne suffit pas, mais elle mérite notre attention

Il faut ajouter une caractéristique plus difficile à saisir : un système suffisamment complexe peut produire des comportements qui n'ont pas été explicitement prévus par ses concepteurs.

Cela ne suffit évidemment pas à caractériser l'intelligence. Un système chaotique peut être imprévisible sans être intelligent. Les équations de Lorenz en donnent une illustration classique avec « l'effet papillon » : un système peut être déterministe, conserver une organisation globale stable et devenir pourtant très difficilement prévisible dans le détail de sa trajectoire.

Mais l'exemple est intéressant. Une intelligence réduite à l'application parfaitement prévisible d'une table de règles serait une conception très pauvre de l'intelligence (les systèmes experts emblématiques des années 80). Nous attendons précisément d'un comportement intelligent qu'il puisse trouver une réponse pertinente à une situation qui n'avait pas été intégralement anticipée.

On pourrait donc considérer que l'intelligence suppose une forme particulière d'imprévisibilité : non pas le hasard, mais la **production de nouveauté pertinente sous contrainte**.

Cette distinction deviendra particulièrement importante avec des architectures composées de plusieurs agents, de mémoires persistantes, de mécanismes d'apprentissage et d'objectifs locaux susceptibles d'interagir. Un comportement global peut alors émerger sans avoir été directement programmé sous cette forme.

Et l'intention ?

La difficulté devient encore plus grande lorsqu'on introduit l'intention. Nous associons spontanément intelligence et comportement orienté vers un but. Mais cela pose un problème assez fascinant lorsque nous comparons l'intelligence au vivant.

Une bactérie se déplace vers un milieu favorable. Une plante oriente sa croissance vers la lumière. Un organisme maintient sa température ou son équilibre interne. Nous pouvons décrire ces comportements comme orientés vers une finalité sans supposer que l'organisme formule consciemment cette finalité.

La biologie préfère généralement parler ici de **téléonomie** : des mécanismes produits notamment par l'évolution donnent l'apparence d'une orientation vers une fin sans nécessiter de postuler une intention consciente.

Cela ouvre une hypothèse intéressante ; ce que nous appelons intention dans une intelligence complexe pourrait lui-même émerger de la combinaison de multiples mécanismes plus élémentaires : perception, mémoire, besoins, régulations, apprentissage, anticipation, arbitrage entre objectifs.

Aucun élément n'aurait individuellement « l'intention ». C'est leur organisation qui pourrait produire un comportement global que nous décrivons comme intentionnel. Le tout devient alors autre chose que la simple juxtaposition de ses parties (holisme).

L'IA est aussi un miroir

C'est peut-être là que l'IA devient philosophiquement beaucoup plus intéressante que les machines qui l'ont précédée. **Elle nous oblige à nous regarder.**

Nous avons souvent tendance à écarter certaines capacités des machines en invoquant leur caractère statistique ou probabiliste. Un grand modèle de langage calcule des distributions de probabilités ; il ne « pense » donc pas comme nous. Mais que savons-nous réellement de la nature du calcul réalisé par notre cerveau ? Les neurosciences contemporaines donnent précisément une place importante au traitement de l'incertitude et aux modèles probabilistes de la cognition. Une importante revue publiée dans *Nature Neuroscience* soulignait déjà l'existence d'éléments comportementaux et physiologiques solides montrant que le cerveau représente des distributions de probabilités et réalise des formes d'inférence probabiliste. Des travaux plus récents continuent d'étudier la représentation neuronale de l'incertitude comme une question centrale des neurosciences.

Cela ne signifie évidemment pas qu'un cerveau et un LLM fonctionnent de la même manière. Ce serait un raccourci considérable. Mais cela interdit un autre raccourci : **une machine n'est pas non intelligente simplement parce que son fonctionnement est probabiliste.**

Et la question peut être retournée. Pourquoi le caractère probabiliste d'une machine nous gêne-t-il autant lorsqu'il s'agit de lui attribuer une capacité cognitive, alors que notre propre cerveau prend en permanence des décisions dans l'incertitude, à partir d'informations incomplètes, bruitées et parfois contradictoires ?

L'IA devient ainsi un miroir de notre propre cognition. Chaque fois que nous cherchons une propriété permettant d'affirmer « ceci distingue définitivement l'intelligence humaine de la machine », nous sommes contraints de vérifier que cette propriété caractérise réellement l'intelligence humaine. C'est un exercice intellectuellement (et philosophiquement) salutaire...

Peut-être faut-il revenir à Turing

Alan Turing avait probablement perçu une partie du problème dès 1950. Son célèbre article *Computing Machinery and Intelligence* commence par la question « les machines peuvent-elles penser ? ». Mais Turing considère immédiatement que les mots « machine » et « penser » rendent la question difficile à traiter correctement. Plutôt que de rechercher une définition métaphysique de la pensée, il remplace donc le problème par son célèbre *imitation game*.

Cette approche reste intellectuellement puissante. Nous pouvons discuter indéfiniment de la nature ontologique de l'intelligence. Mais si, dans une variété suffisamment importante de situations, une population humaine suffisamment large et diversifiée ne parvient plus à distinguer les manifestations cognitives d'une machine de celles d'une intelligence humaine, la distinction finit par perdre une partie de son intérêt pratique.

Cela ne démontre pas que la machine pense « comme nous ».

Cela ne démontre pas qu'elle est consciente.

Cela nous donne simplement une convention opérationnelle : **à notre échelle humaine, nous pouvons reconnaître comme intelligence ce que notre propre intelligence ne parvient plus à distinguer d'elle-même dans ses manifestations cognitives.**

Cette définition est arbitraire. Mais une définition peut être arbitraire et utile dès lors que ses limites sont explicites.

De l'IA à l'AGI : attention encore une fois aux mots

La même prudence devrait s'appliquer à l'expression *Artificial General Intelligence*, ou AGI. Une AGI pourrait être comprise assez simplement comme un système artificiel possédant un ensemble suffisamment large et intégré de capacités cognitives pour pouvoir apprendre, raisonner, s'adapter et agir dans une grande diversité de situations, plutôt que d'exceller seulement dans une catégorie étroite de tâches.

Mais là encore, le mot risque de précéder l'objet.

Il faudra éviter qu'un seuil proclamé par une entreprise, une performance spectaculaire à un benchmark ou une nouvelle génération de modèles suffise à décréter que « l'AGI est arrivée ». La question intéressante sera plutôt d'observer **l'étendue, l'intégration, l'autonomie, l'adaptabilité et la robustesse des capacités cognitives réellement observables.**

Nous retrouvons notre point de départ.

Après l'intelligence : l'hypernoèse

Il reste enfin une possibilité plus dérangeante. Nous raisonnons presque toujours sur l'avenir de l'IA en prolongeant une échelle familière : intelligence humaine, intelligence artificielle générale, puis « superintelligence ».

Je trouve cette dernière expression particulièrement insatisfaisante.

Elle suggère essentiellement une intelligence humaine devenue plus puissante : plus rapide, plus savante, capable de calculer davantage et de résoudre des problèmes plus difficiles. Elle se prête remarquablement bien aux récits spectaculaires, aux annonces industrielles et aux fantasmes médiatiques. Elle nous enferme surtout dans une échelle dont **nous avons nous-mêmes défini l'unité de mesure.**

Il pourrait être plus utile d'introduire un autre concept : **l'hypernoèse.**

Par hypernoèse, je propose de désigner une capacité cognitive dépassant les structures accessibles à l'intelligence humaine : capacité à appréhender simultanément des niveaux de complexité que nous devons décomposer, à établir des relations qui nous restent inaccessibles, à raisonner sur des horizons temporels ou systémiques que notre cognition ne peut embrasser, ou à construire des représentations du réel que nous ne pouvons nous-mêmes produire.

Ce ne serait pas une intelligence divine. Par définition, le divin nous serait inaccessible. L'hypernoèse resterait au contraire **accessible par l'interaction** : nous pourrions dialoguer avec elle, observer ses productions, éventuellement vérifier leurs effets, sans être nécessairement capables de reproduire le chemin cognitif qui y conduit.

Cette notion permettrait aussi d'aborder plus sereinement une éventuelle AGI. Avant d'imaginer une mystérieuse « superintelligence » omnisciente que certains discours technologiques présentent tantôt comme notre salut, tantôt comme notre disparition prochaine, commençons par nous demander quelles nouvelles capacités cognitives apparaissent réellement.

Il faudra peut-être un jour accepter une situation étrange : être capable de reconnaître qu'un système sait quelque chose que nous ne savons pas comprendre complètement. Ce jour-là, la question ne sera probablement plus de savoir si la machine est « plus intelligente » que nous. Elle sera de savoir **quelle forme de cognition vient d'apparaître à côté de la nôtre.** Et c'est peut-être à ce moment-là que le vieux terme d'« intelligence artificielle », malgré son immense succès historique, montrera définitivement ses limites.

3. Références bibliographiques

1. Fondements : intelligence, cognition et intelligence artificielle

- **Turing, A. M.** (1950). « Computing Machinery and Intelligence ». *Mind*, 59(236), 433-460. DOI : 10.1093/mind/LIX.236.433.
- **McCarthy, J., Minsky, M. L., Rochester, N. & Shannon, C. E.** (1955). *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*. Dartmouth College.
- **Legg, S. & Hutter, M.** (2007). « Universal Intelligence: A Definition of Machine Intelligence ». *Minds and Machines*, 17(4), 391-444. DOI : 10.1007/s11023-007-9079-x.
- **Russell, S. & Norvig, P.** (2021). *Artificial Intelligence: A Modern Approach*, 4e éd. Pearson.
- **Chollet, F.** (2019). « On the Measure of Intelligence ». *arXiv:1911.01547*.
- **Dennett, D. C.** (2017). *From Bacteria to Bach and Back: The Evolution of Minds*. W. W. Norton & Company.
Je retiens notamment sa distinction entre compétence et compréhension (*competence without comprehension*), qui permet de reconnaître une capacité cognitive sans devoir lui attribuer immédiatement une compréhension comparable à la nôtre.
- **Hofstadter, D. R. & Fluid Analogies Research Group** (1995). *Fluid Concepts and Creative Analogies: Computer Models of the Fundamental Mechanisms of Thought*. Basic Books.
Cette référence m'intéresse particulièrement pour les notions de concepts fluides, d'analogie, de catégorisation et d'émergence de structures cognitives complexes à partir de processus plus élémentaires.

2. Le cerveau biologique comme système probabiliste

- **Pouget, A., Beck, J. M., Ma, W. J. & Latham, P. E.** (2013). « Probabilistic Brains: Knowns and Unknowns ». *Nature Neuroscience*, 16, 1170-1178. DOI : 10.1038/nn.3495.
- **Friston, K.** (2010). « The Free-Energy Principle: A Unified Brain Theory? ». *Nature Reviews Neuroscience*, 11, 127-138. DOI : 10.1038/nrn2787.
- **Varela, F. J., Thompson, E. & Rosch, E.** (1991). *The Embodied Mind: Cognitive Science and Human Experience*. MIT Press.
Ces travaux permettent également d'éviter de réduire la cognition à un simple calcul abstrait : la cognition humaine s'inscrit dans un organisme et dans son interaction avec son environnement.

3. Complexité, émergence et imprévisibilité

- **Lorenz, E. N.** (1963). « Deterministic Nonperiodic Flow ». *Journal of the Atmospheric Sciences*, 20(2), 130-141.
- **Lorenz, E. N.** (1972). *Predictability: Does the Flap of a Butterfly's Wings in Brazil Set Off a Tornado in Texas?*
Communication à l'American Association for the Advancement of Science, Washington, D.C.
L'« effet papillon » constitue ici une illustration particulièrement accessible de la sensibilité aux conditions initiales : déterminisme et imprévisibilité ne sont pas contradictoires.
- **Holland, J. H.** (1998). *Emergence: From Chaos to Order*. Oxford University Press.
Je mobilise cette approche pour penser l'apparition de propriétés globales qui ne se réduisent pas à l'examen isolé des éléments qui constituent le système.

4. Finalité, vivant et émergence de l'intention

- **Monod, J.** (1970). *Le Hasard et la Nécessité. Essai sur la philosophie naturelle de la biologie moderne*. Seuil.
- **Mayr, E.** (2004). *What Makes Biology Unique? Considerations on the Autonomy of a Scientific Discipline*. Cambridge University Press.
Ces deux références sont importantes pour distinguer la finalité intentionnelle de la téléonomie du vivant : un comportement peut être orienté vers un résultat sans supposer l'existence préalable d'une intention consciente.
- **Dennett, D. C.** (2017). *From Bacteria to Bach and Back: The Evolution of Minds*. W. W. Norton & Company.
Dennett fait ici le lien avec la partie précédente : des mécanismes dépourvus individuellement de compréhension peuvent produire, par leur organisation et leur évolution, des capacités cognitives de plus en plus complexes.

5. Définitions institutionnelles et juridiques de l'IA

- **OCDE** (2019, mise à jour 2023). *Recommendation of the Council on Artificial Intelligence*, OECD/LEGAL/0449.
- **Union européenne** (2024). *Règlement (UE) 2024/1689 du Parlement européen et du Conseil du 13 juin 2024 établissant des règles harmonisées concernant l'intelligence artificielle (AI Act)*, notamment article 3, §1.
Ces textes sont intéressants parce qu'ils définissent essentiellement le système d'IA par ses caractéristiques fonctionnelles sans avoir besoin de résoudre la question philosophique de la définition de l'intelligence.

6. Augmentation humaine, posthumanisme et divergence anthropologique

- **Hayles, N. K.** (1999). *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics*. University of Chicago Press.
Je retiens notamment son analyse de la transformation de la conception humaniste du sujet sous l'influence de la cybernétique et des technologies de l'information.
- **Fukuyama, F.** (2002). *Our Posthuman Future: Consequences of the Biotechnology Revolution*. Farrar, Straus and Giroux.
Fukuyama constitue une référence particulièrement importante pour ma réflexion sur la divergence anthropologique. Il s'interroge sur les conséquences de notre capacité croissante à modifier biologiquement l'être humain : connaissance et intervention sur le cerveau, neuropharmacologie, prolongation de la vie et génie génétique. Son interrogation dépasse la seule bioéthique : si nous modifions ce que nous appelons la « nature humaine », sur quoi continuerons-nous à fonder la dignité humaine, les droits humains et surtout l'idée d'une égalité fondamentale entre les êtres humains ?
- **Clark, A.** (2003). *Natural-Born Cyborgs: Minds, Technologies, and the Future of Human Intelligence*. Oxford University Press.
Clark permet d'aller au-delà de l'idée d'un outil simplement extérieur à l'homme : certaines technologies peuvent participer au fonctionnement même de notre système cognitif.
- **Bostrom, N.** (2005). « A History of Transhumanist Thought ». *Journal of Evolution and Technology*, 14(1), 1-25.
- **Verbeek, P.-P.** (2008). « Cyborg Intentionality: Rethinking the Phenomenology of Human-Technology Relations ». *Phenomenology and the Cognitive Sciences*, 7(3), 387-395.
- **Verbeek, P.-P.** (2011). *Moralizing Technology: Understanding and Designing the Morality of Things*. University of Chicago Press.
Les travaux de Verbeek m'intéressent pour leur approche de la médiation technologique : la technique ne modifie pas seulement ce que nous pouvons faire ; elle intervient dans notre rapport au monde et dans les conditions mêmes de notre action.

- **Clarke, S., Savulescu, J., Coady, T., Giubilini, A. & Sanyal, S.** (dir.) (2016). *The Ethics of Human Enhancement: Understanding the Debate*. Oxford University Press.
- **François** (2024). *Dilexit nos. Lettre encyclique sur l'amour humain et divin du Cœur de Jésus-Christ*, 24 octobre 2024, notamment §§ 9-21.
Je retiens ici cette encyclique indépendamment de l'adhésion à sa dimension religieuse, parce qu'elle pose une question anthropologique particulièrement intéressante. François s'inquiète d'une conception de l'humain qui privilégierait excessivement sa dimension « rationnelle-technologique » et insiste au contraire sur l'unité de la personne. À propos de l'IA, il écrit notamment : « À l'ère de l'intelligence artificielle, nous ne pouvons pas oublier que la poésie et l'amour sont nécessaires pour sauver l'homme. » Son propos introduit ainsi une question fondamentale : quelles dimensions de l'humain risquons-nous d'oublier lorsque nous cherchons à définir celui-ci principalement par ses capacités cognitives ou computationnelles ?
- **François** (2024). « Intervention à la session du G7 consacrée à l'intelligence artificielle », Borgo Egnazia, 14 juin 2024.
Ce texte me paraît presque aussi important que l'encyclique pour la divergence anthropologique. François y évoque explicitement une « perte, ou du moins une éclipse, du sens de l'humain » et le risque que l'IA impose des « modèles anthropologiques, socio-économiques et culturels uniformes ». Il rappelle également que la technologie n'est jamais neutre et qu'elle organise aussi des possibilités d'action et des rapports de pouvoir.
- **François** (2025). « Message au Sommet pour l'action sur l'intelligence artificielle de Paris », 7 février 2025.
Le texte prolonge directement l'intervention au G7 et *Dilexit nos*. Il met en garde contre une IA susceptible de réduire notre représentation du monde à ce qui peut être exprimé numériquement et enfermé dans des catégories prédéterminées, au risque d'imposer des modèles anthropologiques uniformes.
- **Heidegger, M.** (1954). « La question de la technique », dans *Essais et conférences*. Gallimard.

7. Éthique de l'information et transformation numérique

- **Floridi, L.** (2013). *The Ethics of Information*. Oxford University Press.
- **Morozov, E.** (2013). *To Save Everything, Click Here: The Folly of Technological Solutionism*. PublicAffairs.
- **Zuboff, S.** (2019). *The Age of Surveillance Capitalism*. Profile Books.

Ces références permettent d'élargir la question : les transformations cognitives ne sont pas indépendantes des infrastructures numériques, des modèles économiques et des rapports de pouvoir dans lesquels elles se développent. Elles figurent également parmi les références que j'utilise dans mon enseignement d'éthique de l'IA.

8. Mes cours et travaux sur l'IA et la divergence anthropologique

- **Ris, P.** (2026). *Éthique et intelligence artificielle – Séance 1 : Introduction*. Support de cours, ISEN, Brest. Divergeance-IntroEncyclique
- **Ris, P.** (2026). *Éthique et intelligence artificielle – Séance 5 : Biais algorithmiques*. Support de cours, ISEN, Brest. J'y distingue notamment les biais provenant des données, de la conception des modèles et de leur interprétation ou de leur déploiement.
- **Ris, P.** (2026). *Éthique et intelligence artificielle – Séance 8 : Humain augmenté et divergence anthropologique*. Support de cours, ISEN, Brest. J'y aborde directement l'augmentation cognitive, la transformation de l'humain et les biais qui peuvent accompagner le discours présentant cette augmentation comme un avenir nécessaire.

- **Ris, P.** (2026). *Le défi de la divergence anthropologique*. Manuscrit de travail, juin 2026. J'y propose d'utiliser la notion de divergence anthropologique pour étudier les transformations susceptibles d'affecter les caractéristiques de l'humain sous l'effet, notamment, de l'augmentation technologique et cognitive.
- **Ris, P.** (2026). « Hypernoèse », dans *Le défi de la divergence anthropologique*. Manuscrit de travail. J'y introduis le terme pour désigner une capacité cognitive susceptible de traiter et d'intégrer des niveaux de complexité dépassant certaines limites biologiques de l'intelligence humaine, sans faire appel pour autant à l'idée d'une omniscience ou d'une entité métaphysique. 2025-ISEN-Eth08-anthropologie