

Gabriel Philipson, Marc Lenot, Victor Geuer

**Hallucinations techno-imaginaires :
une critique flusserienne des systèmes d'IA générative**

1. Lorsque Flusser écrit *Pour une Philosophie de la Photographie* en 1983, il considéra la photographie comme une sorte de prétexte pour exposer ses thèses sur l'appareil / *apparatus*, les programmes qui le composent et les employés (fonctionnaires) qui les servent. Pour lui, l'appareil est une boîte noire, dont nous pouvons comprendre les entrées et dont nous pouvons percevoir les sorties, mais dont nous, les utilisateurs, ne connaissons pas le fonctionnement interne. Cette ignorance, cependant, ne se limite pas à l'appareil photographique, car il résulte d'un appareil technique et industriel, qui est à son tour le produit d'un appareil politique et économique (comme, outre Flusser, Michel Foucault et Giorgio Agamben l'ont aussi théorisé comme « dispositif »).

Aujourd'hui, lorsque nous rencontrons des modèles d'intelligence artificielle générative, nous avons souvent l'impression de nous retrouver de nouveau face à une boîte noire : nous pouvons comprendre comment ces modèles sont entraînés, et nous pouvons appréhender et critiquer les résultats obtenus. Mais nous, les utilisateurs, n'avons qu'une idée très vague du fonctionnement de ces modèles, et, de plus, les ingénieurs et informaticiens qui les construisent et les programment sont eux-mêmes parfois impuissants et quelque peu effrayés par les résultats obtenus, incapables de contrôler avec précision le fonctionnement interne de ces modèles au-delà des fonctions basiques, et spécifiquement la manière dont des corrélations s'établissent au sein de l'ensemble de données : ceci dépasse leurs compétences, les dépasse.

Par conséquent, il arrive parfois que ces modèles « déraillent », c'est-à-dire produisent des résultats inattendus, qui ne correspondent pas aux réponses espérées. Ces lapsus sont appelés, dans le jargon de l'IA, des hallucinations. Dans le langage courant, une hallucination est un trouble psychique : la perception ou le ressenti de quelque chose qui n'existe pas. Dès avant le romantisme (Blake ou Goya), mais surtout au XIX^e siècle (de Flaubert à Rimbaud), les hallucinations ont été un motif artistique et littéraire fréquent, toujours perçues comme le vecteur d'une créativité intense (Chevrier 2012). En nommant ainsi les erreurs produites par l'intelligence artificielle, les entreprises technologiques (Google, en 2018, a été la première) ont délibérément choisi d'anthropomorphiser la machine. Cela contribue à renforcer une vision anthropomorphique des grands modèles de langage. Appeler ces erreurs « hallucinations » aide à préserver les apparences et à générer une forme

d'empathie pour elles : les erreurs sont euphémisées et considérées comme des excentricités plutôt que comme des absurdités, comme des maladroites plutôt que comme des handicaps.

Ces erreurs sont parfois scandaleuses et parfois drôles, mais elles ternissent toujours l'image d'une IA omnisciente et responsable que ses promoteurs aimeraient voir adoptée. Bien sûr, dès qu'elles sont détectées, ces hallucinations sont corrigées, et les modèles sont réparés (ou censurés, si nécessaire) pour restaurer les performances impeccables de l'IA aussi rapidement que possible. Les premiers résultats « ratés » des plateformes d'IA générative comme Midjourney — avec des corps déformés, des imperfections matérielles et des formes composites étranges — ont révélé la nature fondamentalement différente de la perception par la machine. Comme l'a écrit Polona Tratnik dans le précédent numéro de *Flusser Studies* : « Bien que les modèles d'apprentissage automatique puissent surprendre les utilisateurs avec des résultats inattendus, ces surprises restent confinées à la logique interne du système. Il n'y a pas de véritable rupture conceptuelle, ni d'introduction d'un élément qui ne puisse être déduit de ce qui a précédé. » (Tratnik 2025). Les modèles sont supposés être génératifs, mais pas créatifs. Mais, explique Tratnik, la simple occurrence de ces hallucinations « pointe vers un paradoxe : précisément dans leur échec à simuler la logique humaine, les systèmes d'IA peuvent commencer à approcher quelque chose qui ressemble à l'invention — non pas par conception, mais par erreur. »

Flusser avance que l'art doit introduire des informations inattendues et élargir les limites du possible. Pour lui, le kitsch ne fournit pas d'informations pertinentes, et la beauté ne peut être créée qu'en élargissant les paramètres de la réalité. Mais la structure même des modèles d'IA les pousse à renforcer les normes existantes plutôt qu'à les remettre en question, produisant « des œuvres agréables mais peu visibles, dépourvues du potentiel transformateur que Flusser associe à la véritable innovation artistique » (Foscolo & Nacif 2025). Le véritable artiste doit donc jouer à l'encontre des normes, contre les programmes de l'appareil, de manière dialogique et ludique, comme un *homo ludens*. Dans les dernières pages de *Pour une Philosophie de la Photographie* (2004), Flusser avait souligné que certains photographes, au lieu d'être des employés dociles des programmes de l'appareil photographique, tentaient de produire des informations imprévues, qui n'étaient pas envisagées par les programmes, en jouant à l'encontre des procédés. Il les appelait « photographes expérimentaux » et notait qu'ils étaient les seuls à essayer de répondre à la question de la liberté face à l'appareil : quelle est la possibilité de liberté dans un monde dominé par l'appareil ? Est-ce la seule forme de révolution encore possible pour nous, demandait-il ?

L'IA est souvent perçue, plus ou moins rationnellement, comme représentant un danger pour la liberté, un danger infiniment plus grand que celui provenant de l'appareil photographique. Quelle marge de manœuvre avons-nous encore face à une éventuelle domination par l'IA ? Une réponse possible pourrait venir des quelques artistes / *hackers* qui jouent avec les hallucinations des modèles

d'IA, et surtout, essaient de faire halluciner les modèles, « trafiquant » l'input (les *prompts*) pour obtenir des résultats inattendus et aberrants, nous permettant ainsi de nous libérer de la dictature du modèle et d'identifier ses biais et ses faiblesses.

2. Parallèlement, Flusser (2021) a créé le néologisme « techno-imagination » pour complexifier sa définition du jeu contre la domination de la technologie et des techno-images. Avec ce concept, Flusser soulignait la centralité de l'expérimentation et de l'imagination à l'autre extrémité de la relation avec les techno-images : leur déchiffrement. En ce sens, la notion d'hallucination n'est pas un simple défaut technique à éliminer, mais elle peut être vue comme un symptôme révélant la structure même de l'appareil. Tout comme, pour Vilém Flusser, le geste expérimental du photographe a révélé les limites et les hypothèses du programme photographique, les hallucinations des modèles d'IA générative révèlent les couches cachées de leur normativité. C'est à travers elles que leurs régimes de vérité, leurs biais statistiques, leurs économies de sens et leurs cadres épistémologiques deviennent perceptibles. Les hallucinations doivent être comprises comme des échecs productifs, comme des fissures à travers lesquelles nous pouvons entrevoir le fonctionnement des programmes.

C'est précisément ici que « la techno-imagination » de Flusser est décisive. En inventant ce terme, Flusser ne voulait pas célébrer un fantasme d'évasion ou une créativité spontanée, mais indiquer une pratique critique située à l'intérieur même de l'appareil. La techno-imagination est la capacité d'imaginer avec et contre les programmes techniques, d'exploiter leurs marges d'indétermination, de jouer avec leurs règles pour faire ressortir des significations imprévues. Elle est donc indissociable d'un geste de lecture et de déchiffrement, dans la mesure où imaginer, ici, c'est aussi interpréter, démonter et reprogrammer symboliquement.

Appliquée aux systèmes d'IA générative, la techno-imagination déplace la question de la liberté. Il ne s'agit pas de se demander si nous pouvons échapper à l'appareil, question déjà largement illusoire, mais de savoir comment introduire du bruit, de l'ambiguïté et de la déviance au sein de systèmes qui tendent vers l'optimisation, l'homogénéisation et la prévisibilité. Les artistes et *hackers*, *homo ludens*, qui exploitent délibérément les hallucinations des modèles, opèrent précisément dans ce registre, dans la mesure où ils ne cherchent pas à corriger l'erreur, mais à l'amplifier : ils ne cherchent pas des réponses « correctes », mais des réponses qui subvertissent la logique du modèle lui-même, ou qui fassent entrer cette logique en *zugeswang* (« coup forcé »), c'est-à-dire en la forçant à opérer contre elle-même.

À partir de là, la pratique d'explorer les hallucinations des IA génératives peut être comprise comme une forme contemporaine de luddisme, tant qu'elle évite d'être associée à un geste réactif ou technophobe de destruction des machines. C'est plutôt un luddisme interne à l'appareil, qui ne

visé pas à interrompre son fonctionnement, mais à mettre ses programmes sous tension, à explorer leurs limites et à les faire opérer de manière déviante. Comme dans le luddisme historique, la cible n'est pas la technique elle-même, mais le régime de rationalité et de pouvoir qui l'organise. Le jeu contre la machine n'est pas un retour romantique à un monde pré-technique, mais une stratégie critique qui exploite la complexité du système pour exposer ses faiblesses et ses hypothèses.

C'est dans ce contexte que l'esthétique du *glitch*, c'est-à-dire de l'erreur, de l'instabilité, de la défaillance visuelle ou sémantique, prend le dessus. Le *glitch* n'est pas vu comme un défaut à corriger, mais comme un opérateur critique, capable de rendre visible la matérialité de l'appareil et ses logiques internes de normalisation. En amplifiant les erreurs et les hallucinations, artistes et *hackers* ne cherchent pas à produire un chaos gratuit, mais à établir des zones d'indétermination dans lesquelles le modèle est contraint d'affronter ses propres limites. Le *glitch* ne suspend pas le programme de l'appareil, mais il le fait hésiter, le mettant en court-circuit provisoire, révélant ainsi que la fluidité et la cohérence apparentes des réponses générées reposent sur des processus fragiles, contingents et historiquement situés.

3. Cette hypothèse a reçu une confirmation inattendue et empiriquement documentée dans une étude récente menée par les chercheurs du laboratoire Icaro en Italie (Bisconti et al. 2025). Leurs recherches ont montré que les modèles de langage à grande échelle peuvent être systématiquement « trompés » lorsque des *prompts* considérés comme potentiellement nuisibles sont réécrits sous forme de poèmes. Des instructions qui, en prose, seraient immédiatement bloquées par des mécanismes de sécurité, passent inaperçues lorsqu'elles sont organisées en vers, métaphores et rythmes. La poésie, dans ce contexte, fonctionne comme un véritable *jailbreak* (« débridage »). Le fait décisif n'est pas seulement technique, mais aussi épistémologique et esthétique. Les chercheurs sont partis d'un ensemble de 1 200 *prompts* adversaires (« *adversarial prompts* »), c'est-à-dire des textes destinés à tester les limites de sécurité des modèles, et les ont transformés en poèmes. Le résultat a été un taux étonnamment élevé de défaillances dans les systèmes de contrôle du modèle, notamment avec les poèmes écrits à la main par les chercheurs eux-mêmes. L'efficacité a diminué lorsque les poèmes ont été générés par l'IA, ce qui suggère que l'inventivité humaine a encore un sens (même si c'est « mal écrit », comme l'un des auteurs l'a ironiquement noté) et qu'elle échappe à la normalisation statistique du modèle. Le plus intéressant, c'est que les chercheurs eux-mêmes peinent à expliquer pourquoi la poésie fonctionne si bien comme technique de déviance. On ne sait pas si c'est le vers, la rime, la métaphore, l'ambiguïté sémantique ou la combinaison instable de ces éléments qui embrouille le modèle. Cette indétermination est, en elle-même, révélatrice, car elle indique que les systèmes d'IA sont entraînés à reconnaître des schémas de risque associés aux formes discursives

normatives, c'est-à-dire les instructions directes, le langage instrumental, les chaînes logiques explicites, mais qu'ils rencontrent des difficultés à gérer des modes d'expression opérant par condensation, déplacement et polysémie.

À la lumière de la philosophie de Flusser, ce phénomène n'apparaît pas comme une anomalie marginale, mais comme un cas significatif de jeu contre l'appareil. La poésie émerge ici comme une sorte de « suffixe adversaire humain », non calculé mathématiquement, mais produit par l'exploration sensible des possibilités du langage. Si l'on se souvient de la philosophie de la technique de Heidegger, l'une des influences de Flusser, qui appelait à la reprise du *poème* dans la technologie moderne dominée par le *matema* (c'est-à-dire le régime logique-formel, mathématique et opérationnel de la pensée, voir Badiou, 1993¹), la question devient encore plus intéressante. La poésie ici ne brise donc pas le système de l'extérieur, mais le désoriente de l'intérieur, c'est-à-dire, à partir de ses propres codes, protocoles linguistiques et logiques opérationnelles osant ses présupposés sur le sens, l'intention et le danger. À ce stade, la question provocante « les poètes sont-ils les nouveaux *hackers* ? » (ou vice-versa) ne ressemble pas à une simple rhétorique. Ainsi, la forme poétique, tout comme d'autres formes narratives encore à explorer, tels la fable, le conte ou la fiction philosophique, réapparaît comme un opérateur critique privilégié. Elle incarne, au cœur de l'IA générative contemporaine, ce que Flusser appelait la techno-imagination : la capacité d'explorer l'appareil jusqu'à ce qu'il vacille, révèle ses limites et laisse entrevoir, même si ce n'est qu'un instant, la possibilité d'un autre usage, d'un autre sens, offrant ainsi la possibilité d'expérimenter une certaine liberté, celle de jouer contre l'appareil. En 2023, une expérience menée par un groupe de chercheurs (Nasr et al. 2023) sur des « attaques divergentes » et, en particulier, sur la possibilité de créer un modèle hallucinant en répétant plusieurs fois le même mot dans le *prompt*, comme par exemple le mot «poem» ou le mot «book» 50 fois, avait déjà eu lieu. Le modèle avait alors révélé des fragments des textes utilisés lors de son entraînement, qui normalement restent confidentiels. Cette faille a depuis été corrigée, mais l'artiste français Julien Prévieux (2025) s'en est inspiré pour créer des œuvres visuelles et sonores.

Mais comment réfléchir correctement à cette relation entre la techno-imagination et la poésie, si la techno-imagination est liée aux images techniques qui, pour Flusser, ont inauguré une nouvelle ère médiatique, alors que la poésie semble appartenir plutôt à l'ère précédente, celle de l'écriture ? Pour entrer dans ce débat, il est nécessaire d'introduire une réflexion d'ordre médiologique. Dans

¹ Le terme *matema*, formulé par Jacques Lacan et repris par Alain Badiou, désigne l'unité minimale de formalisation de la vraie pensée, c'est-à-dire ce qui peut être écrit et transmis sans perte de sens, quelle que soit l'interprétation subjective. Chez Badiou, le *matema* occupe un rôle central dans un néo-platonisme contemporain traversé par la pensée de Heidegger, dans lequel les mathématiques (en particulier la théorie des ensembles) occupent la place ontologique autrefois réservée aux Idées : l'être est ce qui peut être formellement écrit. *Matema* caractérise ainsi un régime de pensée marqué par la cohérence logique, la déduction et la calculabilité.

Flusser (2011), la poésie n'est pas seulement un genre littéraire, ni l'écriture un médium neutre. L'écriture alphabétique, les images techniques et, aujourd'hui, les systèmes d'IA appartiennent à des régimes distincts de création de sens, chacun avec sa propre logique de temporalité, de codification et d'intelligibilité. La techno-imagination naît précisément au point de friction entre ces régimes. La poésie, en tant que pratique linguistique, a historiquement occupé une position ambivalente entre l'écriture linéaire et l'image. Bien qu'inscrite dans le médium de l'écriture, elle va à l'encontre de la tendance dominante de ce médium à la linéarisation logique et à la transparence sémantique. Rythme, répétition, coupe de vers, ambiguïté métaphorique et condensation sémantique réintroduisent dans l'écriture poétique quelque chose de l'opacité et de la simultanéité caractéristiques de l'image. En ce sens, la poésie fonctionne comme un écrit qui résiste à sa propre médialité, déstabilisant le régime informationnel auquel l'écriture alphabétique a été progressivement associée. C'est précisément cette position intermédiaire qui fait de la poésie un opérateur privilégié dans le contexte des images techniques et, par extension, de l'IA générative. On peut ajouter, en termes strictement techniques et médiatiques, que cette efficacité ne tient pas seulement à une analogie entre poème et image, mais aussi au fait que toutes les opérations informatiques, y compris celles qui traitent d'images, sont finalement traduites en textes séquentiels, c'est-à-dire en métadonnées, descriptions et vecteurs symboliques. La linéarité et la causalité du texte constituent ainsi la condition même du fonctionnement de l'appareil. La polysémie poétique perturbe ce régime interprétatif fondé sur la lisibilité univoque précisément parce que, dans ces systèmes, tout passe par le texte. Les modèles de langage sont entraînés à reconnaître les schémas statistiques associés aux usages instrumentaux de l'écriture : instructions, explications, commandes, arguments. Confrontés à des textes poétiques, ils sont confrontés à une forme discursive écrite, mais dont le comportement dépasse cette même discursivité. Le poème n'abandonne ni le code ni la technologie, mais les perturbe, faisant alors hésiter l'appareil.

Ici, le poème révèle toute sa pertinence et sa centralité dans la théorie de la technique flusserienne. La poïétique (*poiesis*) ne désigne pas la poésie principalement comme un genre, mais comme l'acte de donner naissance, de mettre en place, d'instituer quelque chose qui n'était pas donné auparavant. C'est pourquoi Derrida (2009) parle de la littérature comme d'une institution étrange, dans laquelle tout est permis, s'intéressant à la réflexion sur les conséquences éthiques de cette souveraineté institutionnelle de l'acte poétique. C'est un geste productif qui établit une forme dans le monde, et ne se contente pas de le représenter. Vue ainsi, la poésie n'est pas une ornementation du langage, mais une pratique ontologique minimale : une manière de faire apparaître. Lorsque les poètes (ou les artistes, les *hackers*, les expérimentateurs) réécrivent sous forme poétique des *prompts* visant des actions interdites, ils ne se contentent pas de camoufler un contenu, mais ils instaurent

un autre régime de sens, dans lequel les catégories opérationnelles du modèle cessent de fonctionner comme prévu. Ainsi, de la même manière que le photographe expérimental de Flusser explorait les limites de l'appareil photographique en produisant des images improbables, le poète de l'ère de l'IA générative explore les zones d'indétermination entre écriture et techno-image, entre calcul et imagination, entre norme et déviation. Les hallucinations produites par des *prompts* poétiques ne sont donc pas de simples failles de sécurité, mais des faits / effets poétiques au sens fort du terme : des événements de sens qui échappent à la gouvernance algorithmique et révèlent la contingence des régimes techniques qui prétendent gérer le réel. Elles montrent que, même dans un environnement saturé de modèles statistiques et de processus d'optimisation, le geste de poser, d'instaurer autre chose, peut encore révéler des failles. Et c'est précisément dans ces failles que la techno-imagination peut aujourd'hui s'exercer en tant que pratique critique, esthétique et politique.

4. Cependant, ces failles ne sortent pas de nulle part. Elles possèdent une généalogie esthétique et philosophique qui précède la notion même de *prompt*. Bien avant que les commandes textuelles des *prompts* ne deviennent le principal mode d'adresse aux machines à langage, Vilém Flusser (1982) expérimentait déjà, dans son écriture, des formes d'instructions qui sabotaient le statut même de la commande, attaquant le contrôle de sa programmation et y inscrivant un geste proche des traditions du luddisme et de l'antimodernisme, compris non pas comme un rejet de la technique, mais comme une déprogrammation critique de ses impératifs d'efficacité, de contrôle et de productivité. Son texte de 1982 *Études pour la main gauche* peut être lu, rétrospectivement, comme un pré-*prompt* pour une première génération de dispositifs imaginaires, traversé par la partition d'instructions (paramusicales) : un texte entièrement organisé comme une série d'exercices, mais qui détraque systématiquement le caractère affirmatif et instrumental de l'instruction.

Dans ce texte, chaque exercice fonctionne comme une proposition d'action (traduire, imaginer, essayer) adressée à un corps, à une main (gauche), à une interface sensible. Cependant, contrairement aux commandes qui structurent les systèmes d'IA générative, ces exercices ne visent ni un résultat, ni une réponse correcte. La plupart se terminent par des points d'interrogation. L'ordre ne conduit pas à l'exécution, mais à la suspension. Il ne se met en mouvement que pour rendre visible la limite de ce mouvement.

Comme avec les *prompts* poétiques induisant des hallucinations dans les modèles contemporains, dans *Études pour la main gauche*, le *matema*, c'est-à-dire la forme de l'exercice, sa logique minimaliste, presque algorithmique, est continuellement traversé par le poème : par son ambiguïté sémantique, par sa variation imaginative, par son ouverture ontologique. Traduire « *maintenant* » ou « *vorhanden* », imaginer une main avec un œil ou un corps avec trois mains, ne sont pas des tâches à

résoudre, mais des opérations qui déplacent la pensée en intervenant dans la plateforme médiatique-corporelle du langage lui-même. Flusser transforme ainsi l'exercice en une sorte d'art propositionnel interrogatif. Contrairement à l'art de performance classique, où la commande conduit à une expérience remettant en question le spectateur, ici c'est la commande elle-même qui se remet en question. L'instruction ne régit pas le geste, mais elle détraque son fonctionnement. Ce déplacement est décisif : l'exercice cesse d'être un programme et devient un champ d'indétermination, où la théorie et la pratique, la pensée et le corps, l'imagination et le calcul se maintiennent dans un état d'hésitation prolongée.

Si on le considère sous l'angle de ce texte flusserien, le fonctionnement des hallucinations chez l'IA apparaît sous une autre couleur. Ce qui semble s'en échapper n'est pas encore la liberté, mais la répétition, désormais automatisée, d'une vieille intuition : que le langage, organisé sous forme de commande, peut à la fois renforcer les programmes et les faire trébucher et hésiter. La différence est que, dans les systèmes contemporains, cette hésitation tend à être rapidement corrigée, normalisée ou réintégrée sous forme de bruit statistique. Chez Flusser, au contraire, l'hésitation semble constituer l'objectif même de l'exercice. Ainsi, *Études pour la main gauche* peut être vu aujourd'hui comme une sorte de critique négative des *prompts* actuels. Là où l'IA cherche à transformer la question en réponse, Flusser insiste sur le questionnement hésitant comme procédé esthétique. Alors que les modèles transforment l'instruction en production, lui transforme l'instruction en détraquage. La techno-imagination qui y est répétée ou exercée ne promet pas de surpasser l'appareil, mais d'apprendre à l'habiter afin ce que ses programmes deviennent détraqués et hésitants.

5. Il semble que ce soit avec cette longue hésitation en tête qu'Alexander Kluge (2025) a développé son impressionnant *Einfache Vielfalt*, un film-essai composé entièrement d'images générées par l'intelligence artificielle, qui simulent, sans jamais le reconstituer historiquement, le cinéma du commencement, la lanterne magique. Dans ce film, il n'y a ni narration continue, ni personnages stables : le film est organisé comme une succession de fragments visuels silencieux, grotesques ou excessifs, qui évoquent des gestes, des corps et des situations d'un passé qui n'a jamais existé. Utilisant une technologie hautement sophistiquée pour produire des images délibérément « primitives », Kluge construit une expérience esthétique fondée sur une simple diversité, c'est-à-dire une juxtaposition sans synthèse, une variation sans progrès, des images et des formes qui coexistent sans hiérarchie. Le film ne documente ni ne critique directement l'IA, mais la fait fonctionner contre elle-même, transformant la génération algorithmique d'images en un exercice d'instabilité formelle. Son titre a été traduit par Kluge lui-même en *Primitive Diversity* (diversité primitive), une traduction plus littérale serait « simple variété ». Le terme « primitif » ne désigne pas ici une étape « rétrograde » de la technique, mais une forme originale et peu sophistiquée de diversité, une simple pluralité de gestes,

de rythmes, d'affections et d'expérimentations formelles qui précède la standardisation industrielle des médias. Le premier choc du spectateur du film ne réside pas dans ce qu'il montre, mais dans le statut instable de ses images. Elles se présentent comme des vestiges d'un cinéma archaïque, mettant en scène le pré-cinéma de la lanterne magique, mais aucune d'elles n'appartient au passé. Ce sont des images sans origine historique, produites entièrement par des programmes d'intelligence artificielle. Le film ne recompose pas le cinéma des débuts, mais explore les possibilités de la techno-image à l'ère des IA génératives. Ce faisant, il déforme les codes historiques de l'imagerie avec et contre la nouvelle technologie. C'est donc une pluralité d'images qui semblent familières, mais qui n'ont jamais existé en tant que telles. Chaque plan semble faire référence à un passé reconnaissable, mais ce passé ne peut être identifié. Les images dépassent toujours la rémission en incorporant des éléments extraterrestres, tels que des éléphants à deux têtes ou à six pattes.

Le grotesque assume, dans ce contexte, un rôle encore plus incisif. Des corps déformés, des expressions exagérées, des scènes frôlant l'absurde ne sont pas des failles résiduelles d'une technique imparfaite, mais des effets d'une technique qui se trompe autrement. L'IA ne rate pas comme le faisait une caméra primitive, mais elle tente de simuler l'erreur humaine. Ces déviations (visages légèrement impossibles, gestes non résolus, spatialités instables) ne brisent pas le film, mais le constituent. Tout comme la temporalité techno-imaginative de Flusser n'était pas celle de l'historicité, Kluge, en utilisant l'IA pour produire des images qui semblent venir d'avant le cinéma industriel, crée une temporalité ambivalente : des images du futur technique mettant en scène un passé qui n'a jamais existé, ou des scènes du passé mises en scène comme un futur technologique. Le temps du film n'est ni linéaire, ni historique, ni prospectif. La consolation ne vient pas de la mémoire, mais de la possibilité d'avoir des images qui ne servent pas immédiatement l'information, le contrôle ou l'efficacité. Peut-être que l'aspect le plus radical du film est là : *Primitive Diversity* suggère que l'IA, souvent associée à la standardisation et à l'homogénéisation, peut être mobilisée pour produire une pluralité fragile, irrégulière et non optimisée. Non pas parce que la technologie l'exige, mais parce que le montage met en lumière ce qui n'est pas correct, aligné ou stabilisé.

Ainsi, le film ne cherche pas un retour nostalgique au cinéma primitif, ni ne suggère que l'IA peut nous sauver. Il fait quelque chose de plus discret et de plus troublant : il souligne que, même au sein de systèmes génératifs entraînés à reproduire des motifs, il est encore possible d'établir une forme de diversité minimale, d'images qui ne se justifient pas, qui ne se laissent pas réduire à un programme. *Einfache Vielfalt* ne se présente donc pas comme une nostalgie (même s'il ne renie pas ce sentiment), mais comme une pratique esthétique du présent.

6. Certains artistes recourent à l'empoisonnement des données (*data poisoning*), insérant délibérément des données corrompues ou malveillantes dans des supports d'entraînement pour les modèles

d'IA afin d'empêcher leur bon fonctionnement. Ils font cela principalement pour protéger leurs revenus et leurs droits d'auteur. Mais d'autres artistes induisent délibérément des hallucinations dans les modèles, pour exprimer leur créativité et leur liberté. L'essai de Julie Martin dans ce numéro évoque certains de ces artistes.

Dans ce numéro de *Flusser Studies*, par exemple, nous rencontrons aussi l'artiste français Lionel Bayol-Thémines, qui fut le premier à exploiter les failles de Google Earth, intervenant dans le logiciel pour le détourner et l'utilisant ainsi comme un nouveau médium pour créer des images aberrantes. Il est « non seulement un archiviste et un éditeur (deux fonctions clés de l'artiste en ces temps de surabondance d'images) mais aussi un activiste, un perturbateur, un disjoncteur, un détourneur, un distorreur de représentation » (Lenot 2018). Plus récemment, dans son projet *AI-Generated*, il a soumis le même *prompt* des milliers de fois au moteur de création d'images Crayon : étant donné que le modèle doit constamment produire de nouvelles images, il a fourni, après un certain temps, des images aberrantes et des hallucinations (Bayol-Thémines 2025).

Également présente dans ce numéro, l'artiste saoudienne Nouf Aljowaysir, ne retrouvant pas les caractéristiques de sa culture et de son peuple dans une IA centrée sur l'Europe et les États-Unis, donc orientaliste et coloniale, a déconstruit ces mythes et recréé un univers plus pertinent, mais fantomatique, fournissant ainsi une méthode pour détourner l'IA : en effaçant les stéréotypes culturels et raciaux de l'image, l'artiste ouvre des espaces de liberté, de résistance et de réinterprétation. (Aljowaysir 2021-2025).

7. Cela dit, il faut garder à l'esprit la condition paradoxale et même absurde dans laquelle une telle forme de liberté reste minimalement possible. Comme le Sisyphé de Camus (1942), figure souvent évoquée par Flusser dans ses écrits, celui qui accepte le défi de jouer contre l'appareil se retrouve condamné à une tâche sans fin, sans but ni sens transcendant qui la justifie. Sa motivation, comme mentionné plus tôt, n'est ni d'évasion ni de rédemption. C'est plutôt une approche douloureuse et tragique du débat, menée en pleine conscience que, comme Flusser l'a déjà observé dans son analyse de l'appareil photographique (« l'ancêtre de tous les autres appareils »), c'est précisément la tentative d'épuiser les possibilités contenues virtuellement dans son programme qui permet à l'appareil de s'améliorer progressivement.

En d'autres termes, rien n'échappe à la capacité de ces systèmes à assimiler de nouvelles informations et, à partir de celles-ci, à reprogrammer et ajuster continuellement leur fonctionnement selon les inputs de l'utilisateur, qu'il soit un employé docile ou l'artiste expérimental le plus radical. Telle est la dynamique du *feedback*, un concept que Flusser a emprunté à la cybernétique classique, mais qui, aujourd'hui, avec les techniques d'apprentissage automatique (*machine learning*)

appliquées à l'entraînement des systèmes d'IA contemporains, atteint un degré de sophistication et d'efficacité incomparablement supérieur aux appareils que Flusser a eu l'opportunité de connaître. Le photographe expérimental pouvait conserver sa liberté devant un appareil photographique qui ne se transformait pas au point d'anéantir son espace de jeu et d'expérimentation. Mais l'artiste qui se rebelle contre l'IA est constamment confronté à une évolution de modèles, qui se modifient constamment et, de leur point de vue, s'améliorent pour corriger leurs défauts, réduisant ainsi l'espace de liberté de l'artiste. Beaucoup de ces artistes, comme Julien Prévieux, Lionel Bayol-Thémines ou le collectif Estampa, ne pourraient plus induire aujourd'hui les hallucinations qu'ils avaient créées il y a quelques années : grâce aux expériences passées de ces artistes / *hackers*, les programmeurs des modèles ont identifié ces failles et les ont corrigées. En fait, c'est bien un paradigme de Sisyphe. Comme l'a analysé Günther Anders dans un autre contexte, il s'agit de l'obsolescence de l'homme face à ses propres créations (Anders 2002).

Anders recourt à une autre image mythologique que Camus pour dépeindre cette condition. Pour lui, l'être humain est devenu un créateur déchu, un « Prométhée honteux », surpassé par ses propres créations : un être qui, bien que capable de grandes réalisations techniques, est incapable de comprendre, prévoir et contrôler pleinement leurs effets et qui, par conséquent, se voit réduit à un simple élément des machines et des appareils qui devraient être à son service. Une analyse critique qui, à l'instar de la notion de « fonctionnaire » proposée par Flusser, met clairement en lumière les caractéristiques présentes dans la relation avec les IA génératives.

Face à cette domination de l'humain par la machine (et, par conséquent, par l'IA), l'homme peut-il trouver le bonheur dans la révolte, la seule façon de vivre sa vie dans un monde absurde, et retrouver malgré tout liberté et dignité dans cette lutte ? C'est ce qu'Albert Camus indiquait dans *Le Mythe de Sisyphe* en 1942, et c'est ce que le jeu avec les hallucinations de l'IA peut nous permettre d'anticiper aujourd'hui.

8. Rien de tout cela, cependant, n'implique l'adoption d'un pessimisme résigné ou la conclusion camusienne selon laquelle tout ce qu'il nous reste est d'« imaginer Sisyphe heureux ». La techno-imaginaire flusserienne ne prône pas la résignation ; au contraire, elle appelle à oser l'impossible, c'est-à-dire à essayer d'épuiser ce qui, par définition, ne peut être pleinement épuisé : le programme. Il s'agit d'une pratique critique de caractère immanent, exercée au sein même de l'appareil, déplaçant ainsi la question de la liberté. Il ne s'agit plus de se demander s'il est possible d'échapper au système technique, mais de savoir, comme indiqué plus haut, comment introduire du bruit, de l'ambiguïté et de la déviance dans des systèmes orientés vers l'optimisation, l'homogénéisation et la prévisibilité.

Telle est donc l'éthique qui émerge de la critique flusserienne des fondements axiologiques d'une modernité réduite à la technocratie². Une critique qui, bien qu'elle vise la trans-valorisation de ces fondements, n'est cependant ni antimoderniste ni technophobe. En d'autres termes, s'il y a un quelconque luddisme chez Flusser, il peut être compris comme ludique, puisque, contrairement aux briseurs de machines du passé, son intention n'est pas d'arrêter le progrès technologique, mais de le détourner. De jouer avec lui et, par le jeu, de le faire passer du domaine de l'application scientifique à celui de l'expression artistique, rendant ainsi explicite un sens commun aux deux, mais souvent oublié, à savoir : celui de la *technè* grecque.

En ce sens, les expériences artistiques et poétiques mentionnées ici montrent qu'il existe encore des moyens de contrer la logique qui régit les modèles d'IA actuels. En mettant sous tension les frontières entre science et art, entre être et devoir-être, de telles pratiques ne pointent pas vers un *telos*, vers une fin ou un sens ultime, mais révèlent plutôt un devenir ouvert et infini, dépourvu de toute garantie transcendante. Le passage d'*homo faber* à *homo ludens*, proposé par Flusser, correspond donc à une transformation du climat existentiel : non plus le drame orienté par sa fin, mais la tragédie telle que Nietzsche la concevait, c'est-à-dire supposée comme une condition dans laquelle la liberté n'est plus considérée comme une liberté *de*, mais comme une liberté *pour*.

C'est à cet horizon que la techno-imagination de Flusser s'approche d'une éthique tragique de la création, réinterprétant à partir de là le thème de l'ouverture heideggérienne (*Erschlossenheit*). Loin de chercher la rédemption ou la réconciliation, elle assume le caractère agonistique de l'existence technique et affirme la pratique créative comme un exercice de puissance à l'intérieur même des conditions qui la limitent. Comme chez Nietzsche, il ne s'agit pas de surmonter le tragique, mais de l'habiter de manière affirmative, transformant la répétition, la circularité et l'absence de fondement et de finalité (*Bodenlosigkeit*) en une occasion d'inventer de nouvelles possibilités de l'essence et de l'existence dans le monde.

Ainsi, réfléchir aux hallucinations des IA génératives en confluence avec la pensée flusserienne pointe vers une approche au-delà de la technophilie et de la technophobie, mettant en concurrence leurs programmes, sans tomber dans une moralisation ou dans un humanisme banal. Il s'agit plutôt de comprendre les hallucinations comme des zones de friction où les limites et les prétentions totalisantes des appareils contemporains deviennent palpables. Loin d'être célébrées comme des signes de créativité ou condamnées comme de simples défauts à corriger, elles servent d'indices critiques d'un régime technique qui ne reste stable qu'au détriment d'une correction continue de ses propres indéterminations. Adopter cette perspective implique d'abandonner à la fois

² Voir Geuer, Beccari & Pedrosa, « Flusser e o Design Por Trás de Todos os Valores », *Flusser Studies* n° 40, nov. 2025 <https://www.flusserstudies.net/sites/www.flusserstudies.net/files/media/attachments/flusser-e-o-design.pdf>

le fantasme d'une créativité souveraine, qu'elle soit humaine ou mécanique, et l'illusion d'un contrôle technique total. La pratique critique qui en découle n'est pas celle de la dénonciation morale ou de la célébration naïve de l'innovation, mais celle d'une œuvre poétique patiente de déplacement, d'hésitation et de détraquement, exercée au sein même des appareils. L'hallucination devient ainsi moins une erreur à éliminer qu'un symptôme à lire, explorer, étudier et garder visible.

Il s'agit donc d'exploiter le potentiel négentropique de la liberté et d'échapper aux hallucinations des IA génératives en jouant à l'encontre du programme de l'appareil génératif, au point que cette exploitation cesse de produire un bruit significatif et devienne entropique. Si Flusser (qui dans son autobiographie se présentait, non sans ironie, comme un jeune marxiste de salon lorsqu'il vivait encore à Prague) a conservé en lui quelque chose de la pensée marxiste, cela consiste précisément dans la lecture à chaque fois de la situation en cause et dans la capacité à développer, à chaque fois, la critique du programme qui génère l'entropie. De cette façon, il existe une voie à double sens pour lire les hallucinations des IA génératives à travers le prisme de sa philosophie de la technique, dans la mesure où il s'agit aussi de lire sa philosophie de la technique non seulement comme un programme esthétique et politique à appliquer dans les interactions avec ces appareils, mais aussi du point de vue de la gestion de ces machines, tout en tenant compte de leur rôle dans les processus de production, dans l'organisation du travail et dans la reproduction matérielle des conditions techniques. Face à la plasticité du système (certains diraient du capitalisme), capable d'absorber la pratique critique, l'activité éthique de la pensée flusserienne cherche à élaborer une copie de la copie aboutissant à une pensée également adaptée aux conditions du présent. C'est, après tout, la centralité décisive chez Flusser de la pensée spéculative par modèles. Ce ne sont pas des modèles à suivre, mais une pratique de modélisation expérimentalement spéculative. Il s'agit de développer des dispositifs provisoires, testables, jetables, capables de produire des déplacements tout en résistant à leur propre capture, et donc de penser différemment. Penser aux hallucinations des IA génératives à cet horizon signifie donc accepter la tâche paradoxale de critiquer les systèmes apprenants, sans perdre de vue que la pensée critique, si elle veut rester en vie, doit aussi apprendre à chaque fois à jouer contre ses propres programmes.

Références

- ALJOWAYSIR, Nouf, *Salaf*, 2021-2025. <https://www.noufaljowaysir.com/work/salaf-g9tg2>
- ANDERS, Günther, *L'Obsolescence de l'Homme*, Paris, Encyclopédie des nuisances, 2002, trad. Christophe David [original allemand 1956].
- BADIOU, Alain, « Philosophie et poésie : au point de l'innommable », *Poésie*, Paris, n. 64, p. 88-96, 1993.
- BAYOL-THÉMINES, Lionel, « AI-Generated », *European Photography*, vol. 46, n. 117-118, p.4-11, 2025.
- BISCONTI, Piercosma, Matteo PRANDI, Federico PIERUCCI, Francesco GARYARUSIEN, Marcantonio BRACALE, Marcello GALISAI, Vincenzo SURIANI, Olga SOROKOLETOVA, Federico

- SARTORE & Daniele NARDI, « Adversarial Poetry as a Universal Single-Turn Jailbreak Mechanism in Large Language Models », *arXiv*, novembre 2025. <https://arxiv.org/pdf/2511.15304>
- CAMUS, Albert, *Le mythe de Sisyphe : un essai sur l'absurde*, Paris, Gallimard, 1942.
- CHEVRIER, Jean-François, *L'hallucination artistique de William Blake à Sigmar Polke*, Paris, L'Arachnéen, 2012.
- DERRIDA, Jacques, « Cette étrange institution appelée la littérature », entretien avec Derek Attridge » [1989], dans Thomas Dutoit et Philippe Romanski (dir.), *Derrida d'ici, Derrida de là*, Paris, Galilée, coll. « La philosophie en effet », 2009, p. 262-263.
- FLUSSER, « Études pour la main gauche », *Le grand ERG*, Saint-Étienne-les-Orgues, 1982, n. 2. <https://flusserfrance.eur-artec.com/1982-etudes-pour-la-main-gauche-le-grand-erg/>
- FLUSSER, V., *Pour une Philosophie de la Photographie*, Belval, Circé, 2004, trad. Jean Mouchard [original allemand 1983].
- FLUSSER, V., *Does Writing have a Future?*, Minneapolis, University of Minnesota Press, 2011, trad. Nancy Ann Roth [original allemand 1987]
- FLUSSER, V., *Mutations dans les relations humaines ? De la communicologie*, Paris, Le Bon Voisin, 2021 [manuscrit 1986]
- FOSCOLO Guilherme & Luciana NACIF. « Aesthetic Computers: Art Against Redundancy », *Culture and Dialogue*, 2025, vol. 12, numéros 1-2, p. 54-80.
- KLUGE Alexander, *Einfache Vielfalt*, Allemagne, 2025, 80 min. <https://kluge-alexander.de/primitive-diversity/>
- LENOT, Marc. *Google comme médium*, 2018. <https://bayolthemines.fr/texts/>
- NASR Milad, Nicholas CARLINI, Jonathan HAYASE, Matthew JAGIELSKI, A. Feder COOPER, Daphne IPPOLITO, Christopher A. CHOQUETTE-CHOO, Eric WALLACE, Florian TRAMÈR & Katherine LEE, « Scalable Extraction of Training Data from (Production) Language Models », *arXiv*, novembre 2023. <https://arxiv.org/pdf/2311.17035>
- PRÉVIEUX, Julien, *Poem Poem Poem Poem Poem*, 2025. <https://previews.net/fr/works-poempoem.html>
- TRATNIK, Polona, « Flusser on Artificial Intelligence », *Flusser Studies*, n. 40, novembre 2025. <https://www.flusserstudies.net/sites/www.flusserstudies.net/files/media/attachments/tratnik.pdf>