

Les trois régimes de l'IA : sens, communication, métabolisme

Architectonique de l'Intelligence artificielle

Jean-Marie Le Ray
Chercheur indépendant, Rome, Italie
jmleray@translation2.com

*Pensée, langage, représentation : trois médiations intemporelles du monde.
Sens, communication, métabolisme : trois régimes contemporains de l'IA.*

Résumé

En 2026, « intelligence artificielle » est devenu un mot-évidence. Mais ce mot ne décrit plus adéquatement la réalité qu'il recouvre : non une faculté unifiée logée dans des machines, ni un objet artificiel isolable, mais un milieu sociotechnique composé de modèles, de plateformes, d'infrastructures, de travail humain, de normes et d'usages — un ensemble massif, instable, difficile à gouverner.

Ce livre propose donc un déplacement conceptuel : parler d'intelligence artificielle. Il ne s'agit pas de décider si des systèmes sont « intelligents », mais de comprendre dans quelles conditions leurs productions deviennent intelligibles, crédibles et opérantes au point d'être reconnues comme telles. L'intelligence n'est pas ici une propriété des artefacts : elle est un effet architectonique, conditionnel et réversible, produit par des agencements techniques, symboliques, organisationnels et matériels.

Pour fonder cet objet, l'ouvrage développe une architectonique articulée autour de trois régimes irréductibles. Le régime du sens est abordé à partir de la responsabilité : la question du sens se déplace de la compréhension vers l'endossement au point d'usage, c'est-à-dire vers les conditions concrètes de reprise, de contestation et d'assomption. Le régime de la communication analyse la transformation contemporaine de la parole à l'ère de la circulation algorithmique, de la convertibilité généralisée et de l'autorité distribuée, où la crise du vrai/faux se double d'une crise fiduciaire. Le régime métabolique met en évidence la dette entropique située — accumulation irréversible de flux matériels, énergétiques, humains et organisationnels — qui impose des seuils ultimes de soutenabilité.

L'intelligence artificielle n'apparaît que lorsque ces trois régimes parviennent, temporairement, à se contraindre mutuellement sans se neutraliser. Elle peut alors être stabilisée et reconnue, mais demeure fragile, conflictuelle, réversible. Lorsque l'un des

régimes écrase les autres, l'IA peut continuer de fonctionner tout en cessant d'être gouvernable : sens sans soutenabilité, communication sans orientation, métabolisme sans signification.

À partir de ce diagnostic, l'ouvrage soutient que gouverner cette intelligence ne consiste ni à optimiser des performances ni à ajouter des normes après coup, mais à rendre visibles et négociables les arbitrages de sens, de circulation et de matérialité qui conditionnent son existence même. Cet essai, *Les trois régimes de l'IA*, ne propose pas une définition supplémentaire : il change le cadre à partir duquel l'intelligence artefactuelle devient pensable — et gouvernable. Il s'adresse aux chercheuses et chercheurs, décideurs publics, concepteurs, juristes et praticiens devant arbitrer, en situation, entre exigences de sens, conditions de circulation et contraintes métaboliques.

*

Mots clés

Intelligence artefactuelle, intelligence artificielle, philosophie de l'IA, sens, communication, métabolisme, architectonique, réponsabilité, responsabilité, dette entropique située, désassomption, dispositifs techniques, régimes, régimes sociotechniques, gouvernabilité, seuils, stabilisation, confiance, flux, arbitrage, autorité distribuée, mémoire computationnelle, point de reprise, visibilité calculée, co-appartenance, effet conditionnel

*

Note de l'auteur

Une fusion-acquisition nécessaire

Ce livre procède d'un geste simple dans son principe, mais exigeant dans son exécution : reprendre trois travaux distincts, déjà largement développés, pour n'en faire qu'un seul ensemble cohérent, en éliminant les redondances, les introductions multiples, les conclusions parallèles et les effets de série.

Les trois volets à l'origine de ce travail sont les suivants :

1. **Sens** — [*Pour une théorie et une pratique du sens à l'ère des modèles*](#)

Un travail consacré à l'actuelle crise du sens, analysée non comme un déficit de compréhension, mais comme une crise de la réponsabilité dans un régime de production discursive massive.

2. **Communication** — [*Le nouveau régime communicationnel de l'humanité*](#)

Une analyse des transformations contemporaines de la parole, de l'autorité et de la circulation des énoncés, à l'ère des plateformes, de la convertibilité généralisée et de l'oralité tertiaire.

3. **Métabolisme** — [*Le régime métabolique de l'IA*](#)

Une enquête sur les conditions matérielles, énergétiques, humaines et organisationnelles qui soutiennent les dispositifs dits intelligents, et sur les seuils à partir desquels ces dispositifs cessent de tenir.

Ces trois travaux représentent plus de cent quatre-vingt mille mots et ont été conçus initialement comme des essais autonomes. Le présent livre n'en est ni un résumé, ni une juxtaposition. Il procède d'une **fusion** : ne conserver de chacun que ce qui est conceptuellement nécessaire à l'architecture d'ensemble, en supprimant les doublons (voire les triplets) et en réorganisant les problématiques autour d'un tableau commun, en vue d'en faire un tout organisé, plus clair et plus cohérent que la simple addition des parties.

Mais cette fusion ne suffit pas à elle seule. Elle s'accompagne de l'**acquisition** explicite d'un cadre d'ensemble, destiné à poser les fondements de l'Intelligence artificielle (IA) en pensant conjointement ces trois régimes, non comme des dimensions additionnelles, mais comme des conditions irréductibles, liées entre elles par des contraintes réciproques.

C'est ce double mouvement — fusion des régimes, acquisition du cadre — qui permet de poser les fondements de ce que ce livre nomme [Intelligence artificielle](#). Celle-ci n'est pas présumée ; elle apparaît comme un effet, produit lorsque les régimes du sens, de la communication et du métabolisme parviennent, temporairement, à se contraindre mutuellement sans se neutraliser.

La note de méthode est donc claire : ce livre ne vise pas à enrichir un discours existant sur l'IA, mais à changer le cadre à partir duquel ce discours devient pensable. Il ne s'adresse pas seulement à celles et ceux qui s'intéressent aux dispositifs techniques, mais à qui cherche à comprendre dans quelles conditions une intelligence peut être reconnue, soutenue et assumée dans des agencements contemporains.

☞ Tous les schémas présentés dans cet essai ont été réalisés en faisant dialoguer **ChatGPT 5.2** et **Grok (automatique)**.

*

Préface

Pourquoi parler d'intelligence artificielle

En 2026, le terme « intelligence artificielle » est durablement installé dans l'usage. Il structure les discours médiatiques, les stratégies industrielles, les cadres réglementaires et les imaginaires collectifs. À ce titre, il serait illusoire de prétendre l'abandonner. Pourtant, ce terme ne correspond plus réellement à ce qu'il désigne. Non parce que les technologies auraient échoué, mais parce qu'elles ont changé de nature. [27][37]*

Ce que l'on appelle aujourd'hui « intelligence artificielle » n'est ni une intelligence au sens d'une faculté unifiée, ni un artefact autonome pouvant être isolé, évalué et gouverné comme tel. Il s'agit de systèmes distribués, dépendants d'infrastructures massives, de flux

* Le [Règlement européen sur l'intelligence artificielle \(AI Act\)](#) constitue aujourd'hui le cadre juridique le plus structurant de la gouvernance institutionnelle de l'IA. Une lecture architectonique de sa fonction d'interface normative est proposée à l'**annexe F**.

énergétiques, de chaînes de travail humain, de dispositifs organisationnels, de régimes de circulation symbolique et de cadres normatifs. Autrement dit, ce que nous continuons d'appeler « IA » fonctionne moins comme une machine intelligente que comme un milieu sociotechnique complexe, dont les effets dépassent largement les objets techniques qui le composent. [21][32]

Dans ce contexte, le terme « intelligence artificielle » devient conceptuellement trompeur. L'adjectif « artificiel » suggère une séparation entre la technique et le social, là où l'on observe une coproduction constante. Le substantif « intelligence » laisse entendre une propriété localisable — dans un modèle, un algorithme, un système — alors que ce qui est en jeu relève de conditions d'organisation, de stabilisation et de reconnaissance. Le vocabulaire hérité empêche ainsi de penser ce qui devient décisif : la tenue de ces systèmes dans le temps, leurs seuils de rupture, et les formes concrètes de responsabilité qu'ils rendent possibles — ou impossibles. [21][34]*

C'est pourquoi ce livre propose un déplacement conceptuel : passer de l'artefact à l'architectonique*. Il ne s'agit plus de se demander si tel ou tel système est intelligent, ni même comment mesurer son intelligence, mais de comprendre dans quelles conditions une intelligence peut émerger, se maintenir et être assumée. L'intelligence n'est ni une faculté interne — humaine ou machinique — ni une propriété technique attachée à un dispositif ; elle est un effet conditionnel, produit par l'agencement de médiations hétérogènes. Un système peut fonctionner sans être intelligible, produire sans être gouvernable, optimiser sans être soutenable. Dans ces cas, l'artefact opère, mais l'intelligence ne tient pas. [20][11]

C'est dans ce sens que nous adopterons ici la terminologie d'intelligence artefactuelle. Non pour désigner une intelligence « faite par des artefacts », mais pour nommer une intelligence faite d'artefacts, de dispositifs, de normes, d'usages et de flux, dont la cohérence dépend de leur organisation. L'intelligence artefactuelle n'est ni humaine ni machinique : elle est architectonique. « Artificielle » renvoie spontanément à une faculté interne, close et substitutive. « Artefactuelle » désigne au contraire une intelligence tenue par des médiations : elle n'existe qu'à travers des dispositifs, des institutions, des flux et des reprises. [19]

* Le cadre du [AI Risk Management Framework publié](#) par le *National Institute of Standards and Technology* (NIST) [34] constituera un point d'appui très souvent appelé dans les chapitres suivants ; sa structure et son statut d'interface de gouvernance sont explicités une fois pour toutes à l'**annexe D**.

* Architectonique vient du grec *arkhitektoniké* (ἀρχιτεκτονική), dérivé de *arkhitéktōn* (ἀρχιτέκτων) : le maître d'œuvre, celui qui coordonne (*arkhê*, principe, commandement) et construit (*téktōn*, artisan). Le terme ne désigne pas d'abord une architecture matérielle, mais un principe d'organisation des parties en vue d'un tout.

En philosophie, depuis Kant, l'architectonique désigne l'art de systématiser des éléments hétérogènes selon un principe de cohérence, sans les réduire à une substance unique. Elle s'oppose à une conception additive ou simplement technique de l'organisation.

L'usage du terme vise à souligner ici que l'intelligence artefactuelle est un effet d'organisation : elle apparaît lorsque des régimes irréductibles — de sens, de communication et de métabolisme — parviennent, temporairement, à coexister en se contraignant l'un l'autre sans se neutraliser. Parler d'architectonique permet ainsi de déplacer l'analyse de l'objet vers les conditions de tenue, de stabilité et de gouvernabilité de ce que l'on appelle, par convention, « intelligence artificielle ».

Ce livre soutient que ces régimes sont au nombre de trois : [le sens](#), [la communication](#) et [le métabolisme](#). Le régime du sens concerne la répondabilité et la stabilisation des interprétations au point d'usage. Le régime de la communication organise la circulation des énoncés, l'autorité, le crédit et la mémoire. Le régime métabolique soutient matériellement l'ensemble, en définissant des flux, des coûts et des seuils irréversibles. Ces régimes sont co-dépendants au sens fonctionnel, mais surtout co-appartenants : aucun ne constitue un simple contexte des autres, et aucun ne peut produire à lui seul une intelligence reconnaissable. L'intelligence artefactuelle n'existe que dans leur tenue conjointe, toujours tendue et conflictuelle.

L'architectonique proposée dans cet ouvrage n'est pas une description cumulative de dimensions, ni un essai de totalisation* théorique. Elle désigne une méthode critique : rendre visibles les contraintes réciproques entre régimes, les seuils à partir desquels ils se neutralisent ou se déforment, et les arbitrages — souvent invisibilisés — par lesquels ils tiennent provisoirement ensemble. Penser l'IA de manière architectonique, c'est refuser les explications par couches, les hiérarchies implicites et les promesses d'optimisation sans coûts, pour interroger les conditions réelles de gouvernabilité de l'intelligence artefactuelle. [21]

L'ambition de ce livre n'est donc pas d'enrichir un discours existant sur l'IA, ni d'en proposer une critique morale supplémentaire. Il s'agit de déplacer le cadre à partir duquel l'IA devient pensable : non comme une intelligence nouvelle à célébrer ou à redouter, mais comme un effet fragile, conditionnel et réversible, dont la reconnaissance engage des choix matériels, symboliques et politiques. C'est à partir de ce déplacement que ce que nous appelons encore, par convention, « intelligence artificielle » peut commencer à être comprise — et éventuellement gouvernée. [34][37]

*

Table des matières

Résumé.....	1
Mots clés.....	2
Note de l'auteur.....	2
Préface.....	3
Liste des figures et des encadrés.....	12
Introduction générale.....	12
Penser l'IA en régimes : méthode, portée, limites.....	12
Pourquoi l'objet « IA » est éclaté.....	12

* L'utilisation du terme « totaliser / totalisation » dans ce texte signifie la tentative de ramener un phénomène complexe à un principe unique, un critère souverain ou une grille générale capable d'épuiser l'ensemble de ses dimensions. Totaliser, ici, c'est vouloir résoudre la tension entre les régimes (sens, communication, métabolisme) en les subsumant sous une seule logique — technique, morale, juridique ou managériale — au prix d'une perte de gouvernabilité effective. La totalisation vise ainsi à neutraliser les tensions interrégimes en les dissolvant dans un critère unique.

Pourquoi l'unification échoue.....	13
Présentation de l'architectonique d'ensemble	13
Place du droit : institution de la contrainte, non fondement de l'intelligence.....	14
Portée et limites de la démarche	14
Organisation de l'ouvrage.....	15
PARTIE I — Principe d'architectonique : penser l'IA en régimes.....	15
1. Régimes, non couches	15
1.1 Historicité des médiations : pensée, langage, représentation	15
1.2 La persistance du modèle par couches.....	16
1.3 Ce que le modèle par couches rend invisible	16
1.4 Le concept de régime : définition générale.....	17
1.5 Irréductibilité, co-conditionnement, non-hiérarchie.....	17
1.6 Régimes et intelligence comme effet.....	17
1.7 Portée méthodologique.....	18
2. Schéma d'ensemble et règles de passage	18
2.1 Triangle des régimes et zones de stabilisation	18
2.2 Seuils, frictions et conflits interrégimes.....	19
2.3 Ce que gouverner l'IA veut dire	20
PARTIE II — Le régime du sens : répondre sans comprendre	21
3. Déplacement de la question du sens.....	21
3.1 De comprendre à répondre	22
3.2 Production massive sans intention ni expérience	23
3.3 Le sens comme stabilisation opératoire	24
4. Le sens comme régime.....	25
4.1 Du contenu au processus.....	26
4.2 Opérateurs du régime du sens	27
4.3 Le tournant architectural	29
5. Répondabilité et agentivité interprétative	30
5.1 Délégation interprétative	30
5.2 Imputabilité, répondabilité, responsabilité.....	31
5.3 Autorité interprétative et crise de la discussion	32
6. Architectures du sens : rendre la répondabilité praticable.....	34
6.1 Dispositifs à seuils multiples	34
6.2 Littératies du sens à l'ère des modèles.....	35
6.3 Écologie du sens répondable	36

PARTIE III — Le régime de la communication : circuler sans origine stable	38
7. Généalogie des régimes du langage	38
7.1 Oralité primaire et imputabilité située	39
7.2 Écriture, mémoire et prescriptions de lecture	39
7.3 Littératies et seuils de transformation	40
8. Convertibilité et textoralité	40
8.1 De la médiation à la convertibilité généralisée	41
8.2 Textoralité opérationnelle	41
8.3 Pragmatique de la parole machinique	42
9. Autorité, mémoire et crédit	42
9.1 Mémoire computationnelle : persistance et oubli	43
9.2 Autorité en chaîne et asymétrie structurelle	43
9.3 De la crise épistémique à la crise fiduciaire	44
10. Figures-limites du régime de la communication	44
10.1 Oralité tertiaire et transformation de la parrhêsia	44
10.2 Palimptextualité et circulation sans origine stable	45
10.3 Textoralité : milieu hybride de la circulation	45
10.4 Textautoralité et responsabilité requalifiée	46
PARTIE IV — Le régime métabolique : soutenir sans assumer	47
11. L'IA comme machine entropique	47
11.1 L'IA comme système de flux	47
11.2 L'extractivisme computationnel différé	51
11.3 Accumulation, invisibilisation et externalisation	53
12. Dette entropique située et seuils critiques	54
12.1 La dette entropique située : définition et portée	54
12.2 Typologie des seuils entropiques	56
12.3 Irréversibilité, hystérésis et points de non-retour	61
13. Conflits de répartition et désassomption	62
13.1 Qui paie la dette ? Répartition et invisibilisation de la dette entropique	62
13.2 Responsabilité comme conflit politico-technique	64
13.3 Rendre discutables les arbitrages	66
13.4 Arbitrages, seuils et gouvernabilité minimale	67
13.5 La limite métabolique et la flèche temporelle	68
PARTIE V — Co-appartenance, tensions et émergence de l'intelligence artefactuelle	71
14. Quand un régime écrase les autres	71

14.1 Sens sans soutenabilité.....	71
14.2 Communication sans orientation	73
14.3 Métabolisme sans signification.....	74
15. L'intelligence artefactuelle comme effet.....	75
15.1 Ni faculté, ni propriété, ni illusion.....	75
15.2 L'intelligence comme stabilisation reconnue.....	76
15.3 Conditions d'apparition, de maintien et de disparition	76
PARTIE VI — Opérateurs et interfaces de l'intelligence artefactuelle	77
16. Pourquoi des opérateurs.....	77
16.1 Concept, opérateur, indicateur.....	77
16.2 Rôle des opérateurs dans une architectonique	79
16.3 Opérateurs propres à chaque régime	80
17. Les opérateurs des trois régimes.....	80
17.1 Répondabilité (sens)	81
17.2 Fiducie et autorité distribuée (communication).....	82
17.3 Dette entropique située (métabolisme)	84
17.4 Pourquoi aucun opérateur n'est totalisant	86
18. Interfaces : traduire sans réduire	86
18.1 Pourquoi des interfaces.....	87
18.2 Interface technique	88
18.3 Interface éthique	89
18.4 Interface juridico-normative.....	91
18.5 Interface éthico-normative.....	92
18.6 Pourquoi les interfaces ne sont pas des régimes.....	94
18.7 Où et comment les interfaces échouent.....	95
PARTIE VII — Vers une gouvernance architectonique de l'intelligence artefactuelle	97
19. Gouverner sans unifier.....	97
19.1 Pourquoi la gouvernance par couches échoue	97
19.2 Pourquoi les principes abstraits ne suffisent pas.....	99
19.3 Gouverner : maintenir la co-contrainte	100
20. Seuils, arbitrages et désassomption	101
20.1 Gouverner par seuils, non par optimisation.....	102
20.2 Seuils de répondabilité, de fiducie, de désassomption.....	103
20.3 Conflictualité constitutive de la gouvernance.....	104
21. Méthodologies possibles (sans promesse)	106

21.1 Diagnostic interrégime	106
21.2 Traduction par interfaces.....	108
21.3 Décision située et réversibilité	109
21.4 Pourquoi aucune méthodologie n'est universelle	111
Conclusion générale	112
Trois régimes irréductibles	112
Une architectonique non totalisante.....	112
Gouverner sans promettre.....	113
Interfaces, seuils et responsabilité	113
Tenir ensemble	113
Postface.....	114
Annexes.....	115
Annexe A — Opérateurs et interfaces (fondation conceptuelle)	115
Statut de l'annexe.....	115
A.1 Pourquoi introduire des opérateurs	116
A.2 Opérateur principal et opérateurs secondaires	116
A.3 L'entropie comme opérateur limite (et non totalisant)	117
A.4 Pourquoi des interfaces.....	117
A.5 Articulation opérateurs / interfaces	119
Annexe B — Instruments de gouvernance architectonique de l'intelligence artefactuelle	120
Statut de l'annexe.....	120
B.1 Principes méthodologiques communs aux outils	120
B.2 Outil 1 — Matrice de diagnostic : <i>Le Triangle des déséquilibres</i>	120
B.3 Outil 2 — Protocole de réponsabilité.....	121
B.4 Outil 3 — Indicateur de charge métabolique réelle.....	122
B.5 Outil 4 — Test de fiducie communicationnelle.....	122
B.6 Outil 5 — Grille des agencements de gouvernance.....	123
B.7 Table de correspondance synthétique	124
Gabarits pratiques — Usage institutionnel.....	124
Conclusion de l'annexe	126
Annexe C — Les constitutions d'IA comme interfaces éthico-normatives	126
C.1 Qu'appelle-t-on une « constitution d'IA » ?	127
C.2 La constitution comme interface éthico-normative.....	127
C.3 Ce que gouverne effectivement une constitution d'IA	127

C.4 Ce que la constitution ne peut pas gouverner	128
C.5 Une pathologie typique : gouverner la conduite sans gouverner les conditions	129
C.6 Place des constitutions dans une architectonique de la gouvernance	129
Conclusion de l'annexe	130
Annexe D — Le NIST AI RMF comme interface de gouvernance (lecture architectonique)	130
Statut de l'annexe.....	130
D.1 Le RMF : une grammaire fonctionnelle plutôt qu'un principe	131
D.2 <i>GOVERN</i> : gouvernance comme fonction transversale	131
D.3 <i>MAP</i> : la cartographie située comme condition d'intervention.....	131
D.4 <i>MEASURE</i> : la puissance des métriques et leurs limites structurelles	132
D.5 <i>MANAGE</i> : gestion du risque ou arbitrage de seuil ?.....	132
D.6 Correspondances avec les trois régimes	133
D.7 Le RMF comme interface, non comme fondement.....	133
D.8 Conclusion : une grammaire précieuse, un horizon limité	134
Annexe E — ISO/IEC 42001 comme interface organisationnelle de gouvernance.....	134
Statut de l'annexe.....	134
E.1 ISO 42001 : gouverner par système de management	135
E.2 Activation indirecte de la réponsabilité	135
E.3 Fiducie et gouvernance procédurale.....	135
E.4 Limite majeure : le métabolisme comme angle mort	136
E.5 Pathologie typique : le juridisme managérial	136
Conclusion de l'annexe	136
Annexe F — L' <i>AI Act</i> européen comme interface juridico-normative (lecture architectonique)	137
Statut de l'annexe.....	137
F.1 Gouverner par niveaux de risque : traduction juridique d'un seuil	137
F.2 Réponsabilité : institution de devoirs formels.....	137
F.3 Fiducie : stabiliser la circulation par obligations de transparence	138
F.4 Limite métabolique : le grand impensé normatif.....	138
F.5 Pathologie possible : conformité sans gouvernabilité.....	138
Conclusion de l'annexe	138
Annexe G — Document-type de gouvernance architectonique située (exemple opérationnel, fictif)	139
Statut de l'annexe.....	139
G.1 Pourquoi un exemple documentaire ?.....	139

G.2 Exemple fictif mais réaliste : usage hospitalier d'un modèle génératif	140
Document-type — Fiche de décision architectonique située	140
1. Diagnostic interrégime (où se défait la co-contrainte ?).....	140
2. Seuil identifié (point de bascule)	141
3. Arbitrage explicite (ce qui est sacrifié / maintenu)	141
4. Interfaces activées (comment intervenir sans réduire).....	142
5. Décision de gouvernance	142
6. Clause de réversibilité.....	142
7. Formule synthétique (gouvernance architectonique).....	142
G.3 Ce que cet exemple rend visible	142
Conclusion de l'annexe	143
Annexe H — <i>Parcoursup</i> comme cas de gouvernance architectonique (exemple appliqué)	144
Statut de l'annexe.....	144
H.1 Configuration générale	144
H.2 Diagnostic interrégime (Triangle des déséquilibres)	144
H.3 Seuil critique 1 — Répondabilité (sens).....	145
H.4 Seuil critique 2 — Fiducie (communication)	145
H.5 Seuil critique 3 — Désassomption (métabolisme)	145
H.6 Résultat architectonique : performance locale, ingouvernabilité globale.....	146
H.7 Interfaces mobilisables (interventions minimales).....	146
Conclusion de l'annexe	146
Annexe I — L'ASI comme configuration-limite : seuil de non-gouvernabilité architectonique.....	147
Statut de l'annexe.....	147
I.1 Régime du sens : répondabilité et rupture d'endossement	147
I.2 Régime de la communication : saturation fiduciaire et perte d'orientation	147
I.3 Régime métabolique : dette entropique explosive et désassomption radicale	148
I.4 Verdict architectonique : fin des conditions de l'intelligence artefactuelle gouvernable.....	148
Conclusion de l'annexe	148
Annexe J — Tables de correspondances avec les trois essais initiaux	149
Partie II – Régime du sens	149
Partie III – Régime de la communication	155
Partie IV – Régime métabolique.....	159
Glossaire	167

Conclusion du glossaire.....	170
Bibliographie	170

Liste des figures et des encadrés

- Figure 1. Architectonique des trois régimes et superposition des interfaces
 Figure 2. Le sens comme effet de processus : cinq opérateurs de stabilisation
 Figure 3. L'IA comme système métabolique de flux : dynamique et dépendances matérielles
 Figure 4. Seuils entropiques : trajectoire de désassomption et points de bascule
 Figure 5. Schéma de synthèse : régimes, opérateurs, seuils et pathologies

*

- Encadré 1 (méthodologique) — L'entropie comme opérateur limite (et non totalisant)
 Encadré 2 (conceptuel) — Flèche temporelle et asymétrie du devenir
 Encadré 3 (analytique) — Désassomption : exemple d'un méga-déploiement infrastructurel
 Encadré 4 (conceptuel) — Irréversibilité et responsabilité : articulation conflictuelle
 Encadré 5 (analytique) — Répondabilité et désassomption : deux seuils asymétriques
 Encadré 6 (synoptique) — Table conceptuelle comparée des trois régimes
 Encadré 7 (empirique) — Seuil métabolique non anticipé et rattrapage institutionnel
-

Introduction générale

Penser l'IA en régimes : méthode, portée, limites

Pourquoi l'objet « IA » est éclaté

L'« intelligence artificielle » s'est imposée comme un objet central des discours contemporains, tout en demeurant conceptuellement instable. Elle est tour à tour décrite comme une technologie, une industrie, un ensemble de modèles, un risque systémique, un levier de productivité ou un problème éthique. Cette prolifération des angles d'approche ne tient pas seulement à la nouveauté ou à la complexité des dispositifs en cause ; elle signale plus profondément une difficulté à identifier ce qui constitue l'objet même que l'on prétend analyser, gouverner ou critiquer. L'IA apparaît ainsi comme un objet éclaté : ni unifié par une essence technique, ni réductible à une fonction sociale déterminée, ni stabilisé par un cadre normatif unique. [32][34]

Cette dispersion n'est pas accidentelle. Elle résulte du fait que ce que l'on appelle « IA » ne correspond pas à un artefact isolable, mais à une configuration instable de médiations hétérogènes : productions de sens, circulations symboliques, infrastructures matérielles, organisations du travail, dispositifs juridiques et normes institutionnelles. Ces dimensions ne s'ajoutent pas les unes aux autres comme des couches superposées ; elles opèrent simultanément, selon des logiques distinctes, parfois incompatibles, dont l'interaction conditionne les effets attribués à l'intelligence. C'est pourquoi toute tentative de saisir l'IA comme un objet unifié — qu'elle prenne la forme d'une théorie globale, d'un cadre éthique

général ou d'un régime juridique totalisant — tend à échouer ou à produire des effets de simplification trompeurs. [21]

Pourquoi l'unification échoue

Les approches dominantes de l'IA reposent implicitement sur un modèle d'unification. Celui-ci peut être technique (l'IA comme système computationnel), normatif (l'IA comme objet de principes éthiques universels), ou juridique (l'IA comme catégorie régulable par un cadre homogène). Dans tous les cas, l'unification suppose qu'il existe un noyau stable de l'objet, autour duquel pourraient s'articuler des dimensions secondaires. Or l'expérience contemporaine des systèmes d'IA montre exactement l'inverse : les crises associées à l'IA — crises du sens, de la responsabilité, de la soutenabilité, de la confiance — ne sont pas périphériques, mais structurelles. Elles ne surviennent pas à la marge d'un fonctionnement central, mais affectent directement ce qui est reconnu comme intelligence. [30][34]

L'échec de l'unification tient ainsi à une confusion méthodologique : on cherche à stabiliser l'objet avant d'avoir compris les conditions sous lesquelles il apparaît comme tel. On suppose une intelligence donnée, puis l'on tente de l'encadrer, de l'optimiser ou de la limiter. Ce livre ne cherche donc pas à corriger ces approches, mais à proposer un renversement de perspective : il ne part pas de l'intelligence comme propriété, mais de l'intelligence comme effet. Un effet toujours conditionnel, attribué, contestable, et réversible, qui dépend de la tenue conjointe de plusieurs régimes irréductibles. [19][21]

Présentation de l'architecture d'ensemble

Pour rendre intelligible cet effet, l'ouvrage adopte une méthode architectonique. Il ne s'agit pas de proposer une description exhaustive des dispositifs d'IA, ni d'en produire une typologie fonctionnelle, mais de penser les conditions de cohérence relative qui permettent à l'intelligence d'être reconnue, assumée et gouvernée. L'architecture ne cherche pas l'unité substantielle de l'objet ; elle analyse les contraintes, les seuils et les tensions par lesquels des dimensions hétérogènes parviennent — ou échouent — à tenir ensemble.

Cette architecture repose sur l'identification de trois régimes constitutifs de l'intelligence artefactuelle contemporaine :

- le régime du sens, qui concerne les conditions de stabilisation des interprétations et la responsabilité des productions au point d'usage ; [39]
- le régime de la communication, qui organise la circulation des énoncés, l'autorité distribuée, le crédit et les formes de mémoire ; [38]
- le régime métabolique, qui soutient matériellement l'ensemble à travers des flux d'énergie, de calcul, de données, d'organisations et de travail humain [40], et qui introduit des seuils de soutenabilité irréversibles.

Ces régimes sont irréductibles : aucun ne peut être traduit intégralement dans le langage d'un autre sans perte décisive. Ils sont également co-conditionnés : chacun dépend des autres pour fonctionner de manière effective. Mais cette co-conditionnalité ne fonde aucune hiérarchie stable. La relation entre régimes est latérale, conflictuelle, et traversée par des arbitrages permanents. L'intelligence artefactuelle n'émerge que dans l'intervalle étroit où ces régimes parviennent à se réguler réciproquement. [11][19]

Place du droit : institution de la contrainte, non fondement de l'intelligence

Dans ce cadre, il est essentiel de préciser le statut du droit. Le droit n'est pas ici pensé comme un régime de l'intelligence artificielle. Le considérer comme tel conduirait à une inflation analytique sans issue — il faudrait alors analyser le sens du droit, sa communication et son métabolisme — et réintroduirait une hiérarchie implicite que l'architectonique cherche précisément à éviter. Le droit ne constitue pas l'intelligence ; il n'en est ni le fondement ni le principe générateur.

Le droit intervient autrement : comme interface normative d'institution. [37] Parfois aussi, dans certaines périodes et certains espaces politiques, il devient fortement structurant : les régulations en vigueur reconfigurent ce qui est faisable, assignable et contestable, et peuvent peser durablement sur les architectures et les organisations. Mais ce caractère structurant ne suffit pas à en faire un régime : le droit ne constitue pas une dynamique irréductible de l'intelligence artificielle ; il institue, historiquement et politiquement, des prises sur des régimes déjà à l'œuvre, en opérant à l'intersection des régimes, en traduisant certains arbitrages en obligations, en procédures, en seuils opposables et en sanctions.

Il ne produit pas de sens à la place du régime du sens, ne garantit pas la fiducie communicationnelle, et ne résout pas les contraintes métaboliques. Il rend cependant certaines décisions irréversibles, certaines responsabilités assignables, et certains conflits juridiquement formulables. En ce sens, le droit n'unifie pas l'architectonique ; il en institue les contraintes, sans jamais pouvoir les totaliser.

Cette distinction est décisive pour éviter deux écueils symétriques : celui d'un juridisme naïf, qui ferait du droit la solution ultime aux pathologies de l'IA, et celui d'un technicisme désengagé, qui relèguerait le droit au rang d'ajustement tardif. Le droit n'est ni un supplément extérieur ni un méta-niveau souverain ; il est une interface parmi d'autres, fragile, située, et historiquement variable, par laquelle la gouvernabilité de l'intelligence artificielle devient partiellement pensable.

Le [RGPD](#) en donne un parfait exemple : il n'évalue pas la « qualité » du sens produit, mais institue des obligations et des droits (information, finalité, minimisation, contestation, recours) qui rendent opposables certains arbitrages entre circulation des données, endossement des usages et coûts organisationnels de conformité. Il fonctionne ainsi comme une interface, en imposant des seuils procéduraux, sans garantir à lui seul la fiducie ni la réponsabilité. [25]

Portée et limites de la démarche

L'ambition de cette introduction n'est pas de clore le débat, mais d'en déplacer le cadre. Penser l'IA en régimes ne fournit ni une définition définitive de l'intelligence, ni un modèle universel de gouvernance. Cela permet en revanche de rendre visibles les conditions sous lesquelles l'intelligence apparaît, se maintient ou se défait, et d'identifier les seuils à partir desquels les systèmes continuent de fonctionner tout en cessant d'être intelligibles, assumables ou gouvernables. [11]

Cette démarche assume ses limites. Elle est non totalisante, ouverte à l'apparition d'autres régimes, d'autres interfaces et d'autres configurations historiques. Elle ne prétend pas résoudre les conflits qui traversent l'intelligence artificielle, mais en préserver la visibilité. C'est à ce prix — en refusant l'unification prématurée et en acceptant la conflictualité comme

condition — que l'intelligence artefactuelle peut être pensée, non comme une promesse ou une menace abstraite, mais comme un effet architectonique toujours fragile, situé et réversible.

Organisation de l'ouvrage

L'architecture de cet ouvrage reflète le geste méthodologique qu'il défend. Elle ne procède ni par accumulation thématique, ni par hiérarchie disciplinaire, mais par déploiement progressif de l'architectonique de l'intelligence artefactuelle.

La **Partie I** établit les fondations méthodologiques : elle explicite la notion de régime et justifie le recours à une analyse architectonique non totalisante. **Les Parties II, III et IV** examinent successivement les trois régimes constitutifs — sens, communication, métabolisme — en respectant leur irréductibilité et leur co-conditionnement. **La Partie V** en propose une synthèse critique, en montrant que l'intelligence artefactuelle n'est ni une faculté ni une propriété, mais un effet conditionnel de leur co-appartenance tendue.

Les deux dernières parties déplacent l'analyse vers la question de la gouvernabilité : la **Partie VI** introduit les notions d'opérateurs et d'interfaces comme outils conceptuels de traduction entre analyse et action, sans prétendre à une méthode universelle ; la **Partie VII** explore enfin des configurations concrètes de gouvernance, en montrant comment ces outils peuvent être mobilisés pour instituer des seuils, rendre des arbitrages discutables et contenir les effets de désassomption*. [34][35][37]

Ce découpage n'a pas pour but de trancher la question, mais de rendre lisible un parcours : celui par lequel l'intelligence artefactuelle peut être pensée, puis gouvernée, sans jamais être réduite à une essence, un principe unique ou une procédure automatique.

PARTIE I — Principe d'architectonique : penser l'IA en régimes

1. Régimes, non couches

La Partie I explicite les fondements conceptuels de cette architectonique. [19]

1.1 Historicité des médiations : pensée, langage, représentation

Penser l'IA en régimes suppose de réinscrire les dispositifs contemporains dans une histoire longue des médiations. La pensée humaine n'a jamais été indépendante de ses supports matériels et symboliques : gestes, outils, parole, écriture, notation, dispositifs techniques ont toujours conditionné ce qui pouvait être pensé, dit, mémorisé et transmis. Il n'existe pas de pensée « pure », mais des formes de pensée historiquement situées, liées à des médiations spécifiques. [6][38]

Le langage constitue l'une de ces médiations fondamentales. Il ne se contente pas d'exprimer des contenus mentaux préexistants ; il structure les formes de l'argumentation, de la preuve,

* L'utilisation du terme « désassomption » dans ce texte désigne une situation où des effets continuent d'être produits (sens, décisions, infrastructures), mais où leurs coûts, leurs responsabilités ou leurs conséquences ne peuvent plus être endossés collectivement : ils sont déplacés, différés ou externalisés hors de toute chaîne explicite de reprise et d'assomption. Voir les **encadrés 3 et 5**.

de l'autorité et de la responsabilité. L'écriture, en particulier, introduit une transformation décisive : elle dissocie l'énoncé de la situation d'énonciation, rend possible la mémoire longue, l'institution de normes et la circulation à distance, tout en appelant des dispositifs de lecture et d'interprétation. [2][10]

La représentation, enfin, ne se limite pas aux images ou aux symboles visuels. Elle désigne la capacité à produire des formes substitutives — textes, chiffres, schémas, modèles — qui tiennent lieu du réel dans des opérations pratiques et décisionnelles. À chaque régime de représentation correspondent des modes spécifiques de validation, de contestation et de responsabilité.

Les médiations computationnelles contemporaines ne constituent donc pas une rupture absolue, mais un nouveau seuil dans cette trajectoire. Elles reconfigurent simultanément les conditions de la pensée, du langage et de la représentation, en introduisant des formes inédites de délégation, de conversion et d'automatisation. C'est cette reconfiguration conjointe qui justifie de parler de régimes, et non de simples outils. [5]

1.2 La persistance du modèle par couches

Les discours contemporains sur l'IA reposent très majoritairement sur un modèle implicite : celui des couches. Dans ce schéma, l'objet serait constitué d'un noyau technique — algorithmes, modèles, infrastructures computationnelles — sur lequel viendraient se greffer successivement des couches sociales, juridiques, éthiques ou politiques. Ce modèle présente un avantage apparent : il permet de préserver l'idée d'un objet central stable, tout en reconnaissant la pluralité des enjeux qui l'entourent.

Cependant, cette représentation est profondément trompeuse. Elle suggère que les dimensions dites « non techniques » pourraient être ajoutées ou retranchées sans affecter la nature même de l'objet. Or les crises contemporaines de l'IA — crises du sens, de la responsabilité, de la soutenabilité — ne sont pas des effets périphériques. Elles touchent le cœur même de ce qui est reconnu comme intelligence. Le modèle par couches empêche donc de comprendre pourquoi des systèmes peuvent fonctionner techniquement tout en cessant d'être intelligibles, gouvernables ou assumables. [21]

1.3 Ce que le modèle par couches rend invisible

Penser en couches revient à traiter certaines dimensions comme secondaires : la matérialité devient un contexte, la communication un canal, le sens un supplément interprétatif. Ce faisant, on rend invisibles les relations de dépendance réciproque entre ces dimensions. Le sens est supposé dériver de la performance ; la responsabilité être ajoutée après coup ; la soutenabilité être ajustée par optimisation.

Or cette hiérarchisation implicite ne correspond pas à la réalité des dispositifs contemporains. Les conditions matérielles influencent directement les formes de communication possibles ; les formats communicationnels transforment les conditions de stabilisation du sens ; et les exigences de réponsabilité modifient en retour les architectures techniques et organisationnelles. Le modèle par couches échoue parce qu'il suppose une asymétrie causale là où il existe en réalité une co-dépendance conflictuelle. [3][21]

1.4 Le concept de régime : définition générale

Le concept de régime est introduit pour rompre avec cette logique. Un régime désigne un mode de fonctionnement relativement autonome, caractérisé par :

- des médiations spécifiques,
- des critères propres de validité,
- des formes particulières de stabilité et de défaillance,
- des temporalités distinctes.

Un régime n'est ni un niveau d'analyse, ni un simple contexte. Il ne se superpose pas aux autres régimes comme une couche, mais opère selon sa logique propre, tout en étant structurellement dépendant de la présence des autres. Cette dépendance n'implique aucune hiérarchie stable : aucun régime ne fonde les autres, aucun ne peut les absorber sans les déformer. [11][19]

1.5 Irréductibilité, co-conditionnement, non-hiérarchie

Dire que les régimes sont irréductibles signifie qu'aucun ne peut être traduit intégralement dans le langage d'un autre sans perte décisive : le sens n'est pas une métrique de performance, la communication n'est pas un simple transport d'informations, et le métabolisme n'est pas un contexte technique neutre.

Cette irréductibilité n'est pas une thèse abstraite, mais un constat opératoire : toute tentative de réduction produit des pathologies identifiables — inflation discursive sans responsabilité, circulation sans autorité, optimisation sans soutenabilité. L'analyse en régimes vise précisément à rendre ces crises intelligibles. [8][30]

Toutefois, cette irréductibilité ne conduit pas à l'isolement. Les régimes sont co-conditionnés : chacun a besoin des autres pour fonctionner. Le sens doit circuler pour exister socialement ; la communication doit être soutenue matériellement ; le métabolisme doit être orienté et légitimé par le sens.

Pour autant, ce co-conditionnement ne fonde aucune hiérarchie stable. Il n'existe pas de régime premier. Toute tentative de hiérarchisation — par exemple la réduction du sens à la technique ou de la communication à l'optimisation — conduit à une neutralisation : les contraintes des autres régimes sont alors dissoutes plutôt qu'assumées.

Ainsi, la relation entre régimes est fondamentalement latérale et conflictuelle. C'est cette conflictualité maintenue, et non son effacement, qui rend possible une gouvernabilité minimale. [21][34]

1.6 Régimes et intelligence comme effet

Dans ce cadre, l'intelligence ne peut plus être pensée comme une faculté interne — qu'elle soit humaine ou machinique — ni comme une propriété émergente des systèmes techniques. Elle doit être comprise comme un effet conditionnel, produit lorsque certaines conditions sont réunies, et susceptible de disparaître lorsque ces conditions cessent de l'être. [20][21]

L'intelligence est conditionnelle parce qu'elle dépend de la tenue conjointe des trois régimes : suffisamment de sens pour orienter l'action, suffisamment de communication pour circuler

et faire autorité, suffisamment de métabolisme pour soutenir ces opérations dans le temps. Elle est attribuée parce qu'elle n'est jamais donnée comme un fait brut : elle est reconnue, contestée, retirée au gré des usages, des dispositifs et des arbitrages. [18]

Cette attribution n'est ni arbitraire ni purement discursive. Elle repose sur des conditions pratiques : possibilité de reprise, de contestation, de correction, d'assomption des effets. Lorsque ces conditions sont réunies, l'intelligence est reconnue comme telle ; lorsqu'elles ne le sont plus, les systèmes peuvent continuer à fonctionner tout en cessant d'être intelligents au sens fort. [4][34]

Cette conception permet de comprendre pourquoi l'intelligence artefactuelle est toujours locale, provisoire et réversible. Elle ne s'accumule pas avec la performance ; elle se maintient par l'équilibre instable des régimes. C'est cet équilibre que l'architectonique proposée dans cet ouvrage cherche à rendre visible, analysable et gouvernable. [11]

1.7 Portée méthodologique

Ce premier chapitre ne vise pas à substituer un vocabulaire à un autre, mais à établir une méthode d'analyse. Les chapitres suivants n'appliqueront pas ce cadre de manière illustrative ; ils exploreront, pour chaque régime, ses logiques propres, ses seuils critiques et ses formes de conflictualité. L'enjeu n'est pas de totaliser les régimes, mais de comprendre comment leur co-appartenance rend possible — ou impossible — l'intelligence artefactuelle. [21]

2. Schéma d'ensemble et règles de passage

2.1 Triangle des régimes et zones de stabilisation

L'architectonique proposée dans cet ouvrage peut être représentée de manière synthétique par un triangle, dont chacun des sommets correspond à l'un des trois régimes analysés : sens, communication, métabolisme. Cette représentation n'a pas pour fonction de figer une structure, mais de rendre visible une propriété essentielle de l'intelligence artefactuelle : elle ne réside ni dans un régime isolé, ni dans un centre substantiel, mais dans des zones de stabilisation produites par la co-appartenance contrainte des régimes. [19][21]

Chaque sommet du triangle désigne un pôle de contraintes spécifiques. Le régime du sens impose des exigences de réponsabilité et de stabilisation interprétative ; le régime de la communication impose des exigences de circulation, d'autorité et de crédibilité ; le régime métabolique impose des exigences de soutenabilité, de flux et de répartition des coûts. Ces exigences ne convergent pas spontanément. Elles entrent en tension dès lors que les dispositifs passent à l'échelle, accélèrent ou se généralisent. [28][8]

Les zones de stabilisation ne doivent donc pas être pensées comme un point d'équilibre idéal, mais comme des configurations locales et provisoires dans lesquelles les contraintes des régimes sont suffisamment compatibles pour que des productions soient reconnues comme intelligentes. Ces zones peuvent être institutionnelles (par exemple, des dispositifs professionnels fortement encadrés), techniques (architectures limitant explicitement certains usages), ou organisationnelles (chaînes de décision clairement endossées). Elles sont toujours situées, historiquement et socialement déterminées. [34]

Il n'existe pas de centre stable du triangle où l'intelligence serait garantie une fois pour toutes. L'intelligence artefactuelle apparaît et disparaît au gré des déplacements de ces zones de stabilisation, en fonction des arbitrages opérés entre régimes.

Pour rendre visible cette co-appartenance conflictuelle et la superposition des médiations qui la structurent, on peut représenter l'architectonique sous la forme suivante :

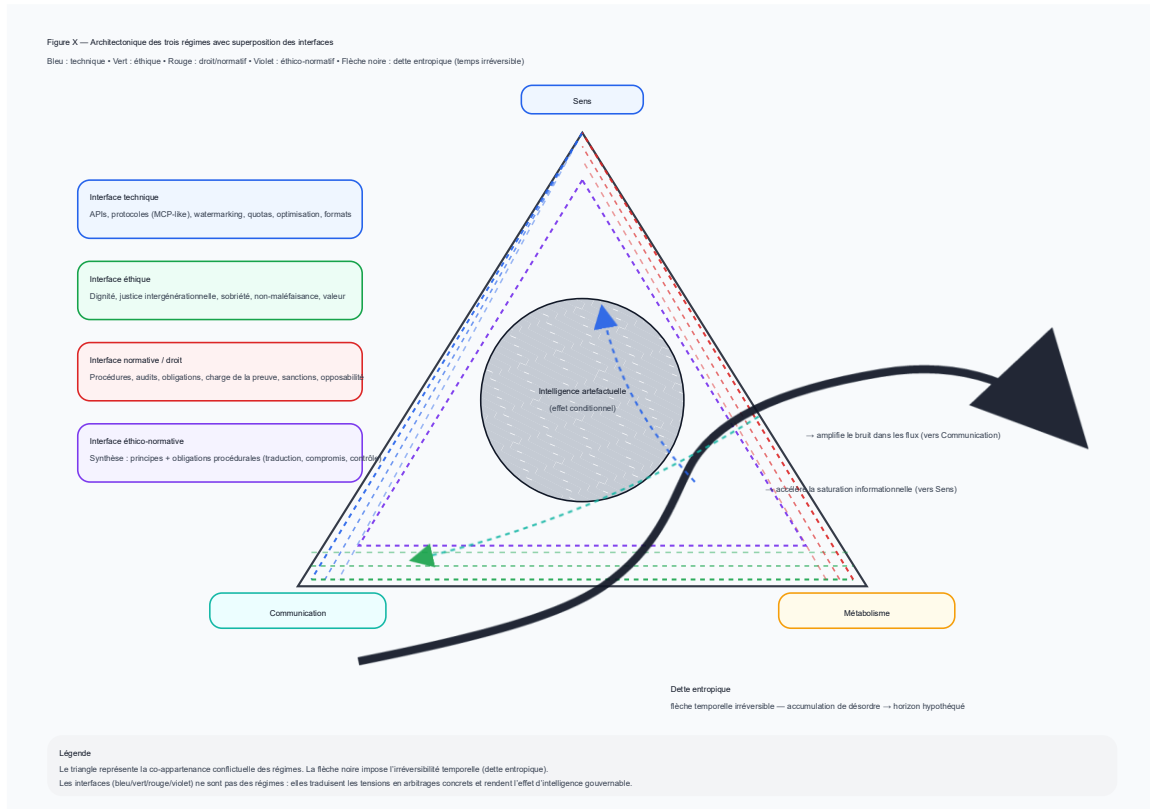


Figure 1. Architectonique des trois régimes avec superposition des interfaces

Ce schéma met en évidence que les zones de stabilisation ne résultent pas d'une harmonie spontanée, mais d'arbitrages concrets entre interfaces hétérogènes (technique, éthique, normative, éthico-normative). Ces interfaces ne sont pas des régimes autonomes, mais des traductions tendues qui tentent — souvent de manière précaire — de rendre compatibles les exigences des trois pôles. [31][37] La flèche noire qui traverse le diagramme souligne quant à elle l'irréversibilité temporelle de cette dynamique : l'accumulation de couches et de frottements produit une dette entropique qui sature progressivement les capacités de sens, de communication et de métabolisme. [8][11]

2.2 Seuils, frictions et conflits interrégimes

La co-appartenance des régimes n'est jamais lisse. Elle est traversée par des frictions, qui deviennent visibles à partir de certains seuils. Ces seuils ne sont pas seulement quantitatifs ; ils sont structurels. Ils marquent les points à partir desquels un régime commence à imposer sa logique aux autres, au risque de les neutraliser. [11][19]

On peut distinguer plusieurs types de seuils. Les seuils d'échelle, d'abord : ce qui est gouvernable dans un contexte limité cesse de l'être lorsqu'un dispositif est déployé

massivement. Les seuils de vitesse et de temporalité, ensuite : l'accélération de la production et de la circulation rend impraticables les dispositifs de vérification, de discussion et de responsabilité. Les seuils métaboliques, enfin : à partir d'un certain niveau de charge, les coûts matériels, humains et organisationnels excèdent ce qui peut être absorbé sans déplacement ou invisibilisation. [8][34]

Ces seuils donnent lieu à des conflits de régimes. Le régime de la communication peut exiger une fluidité incompatible avec les exigences de réponsabilité du sens. Le régime métabolique peut imposer des arbitrages de coûts qui rendent irréalistes les idéaux de transparence ou de responsabilité. Le régime du sens peut multiplier les exigences normatives sans tenir compte de leur coût métabolique.

Ces conflits ne sont pas des dysfonctionnements accidentels. Ils constituent la dynamique normale de l'architectonique. Chercher à les éliminer conduit à des solutions illusoires : soit par la réduction d'un régime aux autres, soit par la moralisation abstraite. À l'inverse, les rendre visibles permet de comprendre pourquoi certaines configurations tiennent — et pourquoi d'autres s'effondrent malgré leur performance apparente. [21]

2.3 Ce que gouverner l'IA veut dire

Dans ce cadre, gouverner l'IA ne peut signifier ni contrôler intégralement des systèmes, ni optimiser leurs performances, ni ajouter des normes après coup. Gouverner signifie ici intervenir sur les conditions de co-appartenance des régimes, c'est-à-dire sur les arbitrages qui rendent possibles — ou impossibles — les zones de stabilisation de l'intelligence artefactuelle. [34][35]*

Gouverner l'IA, c'est d'abord rendre visibles les médiations : identifier où se prennent les décisions, où s'exerce l'autorité, où se déplacent les coûts. C'est ensuite instituer des points d'endossement clairs, là où la responsabilité tend à se fragmenter ou à se dissoudre. C'est enfin accepter que toute gouvernance implique des pertes, des ralentissements et des renoncements : il n'existe pas de gouvernance sans conflit, ni de responsabilité sans coût. [21][8]

Cette conception s'oppose explicitement à deux illusions dominantes. La première est l'illusion techniciste, selon laquelle une meilleure ingénierie suffirait à résoudre les problèmes de sens, de communication et de soutenabilité. La seconde est l'illusion normative, selon laquelle des principes éthiques généraux pourraient être appliqués indépendamment des contraintes métaboliques et organisationnelles.

Gouverner l'IA, au sens proposé ici, revient à maintenir ouverts les conflits de régimes, à les rendre discutables plutôt qu'à les masquer. Il ne s'agit pas de garantir une intelligence parfaite, mais de préserver les conditions sous lesquelles une intelligence peut être reconnue, contestée et assumée. [18] C'est en ce sens que l'architectonique développée dans cette partie constitue non seulement un cadre analytique, mais une orientation pratique pour penser la gouvernabilité contemporaine de l'intelligence artefactuelle.

* La norme ISO/IEC 42001 constitue une tentative majeure d'institutionnalisation organisationnelle de la gouvernance de l'IA sous forme de système de management. Une lecture architectonique de sa fonction et de ses limites est proposée à l'**annexe E**. [35].

PARTIE II — Le régime du sens : répondre sans comprendre

Du sens comme contenu au sens comme condition d'endossement

La question du sens a longtemps été abordée, dans les sciences du langage comme dans la philosophie de l'esprit, à partir de catégories cognitives : compréhension, intention, signification, référence. Le sens était conçu comme un contenu à décoder, à interpréter ou à représenter adéquatement, que ce contenu soit situé dans un esprit humain ou inscrit dans un système symbolique. Les débats contemporains sur l'intelligence artificielle ont largement reconduit ce schéma, en se demandant si les systèmes « comprennent vraiment », s'ils possèdent une intention, ou s'ils accèdent à une forme authentique de signification. Or ces questions, pour légitimes qu'elles soient du point de vue théorique, se révèlent de plus en plus inopérantes face aux dispositifs contemporains, dont l'efficacité sociale ne dépend ni d'une compréhension interne, ni d'une intentionnalité attribuable. [2][30][39]

Cette partie propose un déplacement décisif : penser le sens non plus comme un contenu mental ou sémantique, mais comme une condition d'endossement. Ce qui est en jeu n'est pas la question de savoir si un système comprend ce qu'il produit, mais dans quelles conditions ses productions peuvent être reprises, assumées, contestées ou corrigées par des acteurs humains et institutionnels. Le sens cesse alors d'être une propriété interne pour devenir une propriété relationnelle et située : il n'existe que là où une réponse peut être donnée, là où une chaîne de responsabilité peut être activée, là où une interprétation peut être stabilisée suffisamment pour orienter l'action. [1][4]

Dans le contexte de l'intelligence artificielle, ce déplacement est crucial. Les systèmes contemporains produisent massivement des énoncés plausibles, contextualisés et efficaces, sans expérience vécue, sans intention propre et sans compréhension au sens fort. Pourtant, ces productions circulent, orientent des décisions et produisent des effets réels. La question du sens ne peut donc plus être posée en termes de vérité interne ou d'authenticité cognitive, mais en termes de répondabilité : qui peut répondre de ce qui est produit, à quel moment, selon quelles procédures, et avec quelles limites. [1][9]

Ce constat étant posé, l'objectif de cette partie est de reconstruire une théorie du sens adaptée à un régime de production discursive massive, automatisée et distribuée. Loin de vouloir disqualifier les approches classiques ou d'y substituer une théorie psychologisante, il s'agit d'analyser le sens comme une stabilisation opératoire des interprétations au point d'usage. À travers l'examen de ses conditions de possibilité, de ses seuils critiques et de ses pathologies propres, nous chercherons à identifier les leviers sous lesquels le sens peut encore être stabilisé, assumé et gouverné à l'ère des modèles. [39]

3. Déplacement de la question du sens

La reconfiguration du problème du sens esquissée dans l'introduction de cette partie appelle désormais une explicitation systématique. Il ne s'agit pas simplement d'ajouter une nouvelle perspective aux théories existantes, mais de déplacer le point d'application même de la question : quitter le terrain de la compréhension interne, de l'intention ou de la signification

mentale, pour interroger les conditions pratiques, institutionnelles et architecturales dans lesquelles des productions discursives peuvent faire sens. [2][4]

Ce déplacement est rendu nécessaire par l'émergence de systèmes capables de produire des réponses pertinentes sans accéder à ce que les traditions philosophiques reconnaissaient comme les conditions minimales de la compréhension. Le chapitre 3 examine ce glissement en partant de la distinction décisive entre *comprendre* et *répondre*, afin de montrer comment le sens se recompose à partir de la seule opérativité de la réponse. [30]

3.1 De comprendre à répondre

La philosophie du langage et de l'esprit a longtemps posé la compréhension comme condition de possibilité du sens. Comprendre un énoncé, c'était pouvoir en expliciter la signification, en saisir l'intention, en reconstruire les conditions de vérité ou en rapporter le contenu à une représentation mentale adéquate. Même lorsque le sens était conçu comme usage ou comme pratique, la compréhension demeurait l'horizon normatif implicite : parler avec sens supposait que quelqu'un, quelque part, comprenne ce qui était dit.

Or les dispositifs contemporains de production langagière mettent cette présupposition en crise. Les systèmes de génération automatique produisent des réponses syntaxiquement correctes, sémantiquement cohérentes et pragmatiquement ajustées, sans que rien ne permette de leur attribuer une compréhension au sens classique. Ils ne savent pas ce qu'ils disent, ne visent rien par ce qu'ils produisent, et ne peuvent pas rapporter leurs énoncés à une expérience ou à une intention. Pourtant, ces réponses fonctionnent : elles sont reprises, intégrées à des décisions, mobilisées dans des interactions sociales et institutionnelles. Le sens ne disparaît pas avec la compréhension ; il se reconfigure ailleurs.

C'est dans cet écart que se joue le déplacement proposé ici : substituer à la question « *le système comprend-il ?* » la demande « *à quelles conditions peut-on répondre de ce qui est produit ?* ». Répondre ne signifie pas seulement produire une sortie linguistique pertinente ; cela signifie être inscrit dans un régime où une réponse peut être interrogée, corrigée, assumée ou contestée. Là où la compréhension renvoie à un état interne supposé, la réponse engage une situation, une adresse et une possibilité de reprise. Le sens ne tient plus à ce qui est mentalement saisi, mais à ce qui peut être publiquement endossé.

Ce déplacement entraîne une conséquence théorique majeure : le sens cesse d'être conçu comme une propriété intrinsèque des énoncés ou des agents pour devenir une propriété relationnelle, distribuée entre des dispositifs techniques, des cadres institutionnels et des pratiques interprétatives. Un énoncé fait sens non pas parce qu'il est compris « en lui-même », mais parce qu'il s'inscrit dans une chaîne où quelqu'un peut être tenu pour répondant — même si ce quelqu'un n'est pas l'agent producteur immédiat de l'énoncé.

Ainsi, la réponse devient l'unité minimale du sens, non en tant qu'acte intentionnel, mais en tant qu'élément d'un circuit de répondabilité. Ce circuit peut être discontinu, délégué, asymétrique, mais il doit rester activable. Lorsque cette activation devient impossible — lorsqu'aucune instance ne peut répondre de la production, de ses conditions ou de ses effets — le sens ne s'effondre pas immédiatement, mais il entre dans un régime pathologique : celui de l'opacité opératoire et de l'irresponsabilité diffuse.

Penser le passage de la compréhension à la réponse, ce n'est donc pas renoncer à toute normativité du sens ; c'est en déplacer le foyer. La normativité ne porte plus sur l'adéquation d'un contenu à une intention ou à une vérité mentale, mais sur la possibilité de faire droit à une demande d'explication, de justification ou de correction. Le sens devient ce qui résiste à l'interrogation, ce qui peut être repris sans se dissoudre, ce qui demeure suffisamment stable pour orienter l'action tout en restant exposé à la contestation. [9]

3.2 Production massive sans intention ni expérience

Le déplacement du sens de la compréhension vers la réponse ne devient pleinement intelligible qu'à partir du moment où l'on prend la mesure du changement d'échelle introduit par les systèmes contemporains de production langagière. Ce qui est en jeu n'est pas seulement l'absence d'intention ou d'expérience vécue, déjà problématisée dans les débats classiques sur les machines, mais la capacité de produire, à très grande échelle, des énoncés localement pertinents, contextuellement ajustés et socialement opérants. La question du sens se trouve ainsi reconfigurée par un régime de production qui n'est plus événementiel, mais industriel, continu et distribué. [30]

Contrairement aux agents humains, ces systèmes ne parlent pas à partir d'un monde vécu. Ils ne perçoivent pas, ne se souviennent pas, n'anticipent pas au sens phénoménologique du terme. Leurs productions ne sont pas l'expression d'une expérience, mais le résultat d'une opération de calcul sur des régularités statistiques extraites de corpus massifs. Cette absence d'ancrage expérientiel n'empêche pourtant pas leurs énoncés d'entrer dans des circuits d'usage où ils sont traités comme porteurs de sens. Le paradoxe n'est donc pas que des systèmes « sans esprit » produisent du langage, mais que ce langage puisse fonctionner sans renvoyer à aucune source intentionnelle identifiable. [26][30]

La massification joue ici un rôle décisif. Là où les artefacts langagiers antérieurs produisaient des sorties limitées, spécialisées ou fortement encadrées, les systèmes actuels génèrent des réponses en flux continu, quasiment en temps réel, sur des domaines hétérogènes, à destination d'usages multiples et souvent imprévisibles. Le sens n'est plus stabilisé en amont, par l'intention de l'émetteur, mais en aval, par la sélection, la reprise et l'intégration des réponses dans des pratiques humaines. Ce renversement déplace la charge interprétative vers les points d'usage, tout en multipliant les situations où cette charge devient difficilement assumable. [2]

L'absence d'intention et d'expérience ne conduit donc pas à un déficit immédiat de sens, mais à une transformation de ses conditions de possibilité. Le sens n'est plus ce qui relie un énoncé à un sujet, mais ce qui permet à un énoncé de circuler sans sujet. Cette circulation, cependant, n'est pas neutre : elle produit des effets cognitifs, normatifs et politiques, précisément parce que les réponses générées sont plausibles, cohérentes et souvent convaincantes. Le régime du sens qui en résulte est un régime de crédibilité sans garantie, où la pertinence peut masquer l'absence de fondement répondable.

Cette situation introduit une dissymétrie nouvelle entre production et interprétation. Les systèmes produisent sans engager quoi que ce soit d'eux-mêmes, tandis que les acteurs humains — utilisateurs, institutions, organisations — sont amenés à assumer les effets de ces productions. La responsabilité ne disparaît pas ; elle se déplace et se fragmente. Plus la production est massive, plus il devient difficile d'identifier les points où une reprise critique

est possible, et plus le risque d'une naturalisation* du sens augmente : les réponses « font sens » parce qu'elles sont là, disponibles, répétées et intégrées aux routines décisionnelles. [28]

Il en résulte une tension structurante du régime contemporain du sens : d'un côté, une efficacité discursive inédite, fondée sur la continuité et la disponibilité des réponses ; de l'autre, une fragilisation des mécanismes de contrôle interprétatif, faute d'intention, d'expérience ou de sujet répondant du côté de la production. Cette tension ne peut être résolue en réintroduisant artificiellement une psychologie de la machine ou une fiction d'intentionnalité. Elle appelle au contraire une analyse des dispositifs, des architectures et des procédures capables de rendre cette production massive interprétable, contestable et gouvernable. [34]

C'est à partir de cette production sans intention ni expérience que le sens doit désormais être pensé comme une stabilisation opératoire, toujours précaire, toujours située, et dépendante des conditions dans lesquelles une réponse peut être arrêtée, validée ou remise en circulation. [39] La section suivante examinera précisément ce point : comment, dans un tel contexte, le sens peut encore se stabiliser suffisamment pour orienter l'action sans se confondre avec une simple plausibilité calculée.

3.3 Le sens comme stabilisation opératoire

Si le sens ne peut plus être fondé ni sur la compréhension interne, ni sur l'intention, ni sur l'expérience vécue, il devient nécessaire d'en reformuler positivement les conditions d'existence. Les sections précédentes ont montré que les productions discursives contemporaines fonctionnent sans sujet parlant au sens classique, tout en produisant des effets interprétatifs et pratiques bien réels. La question n'est donc plus de savoir *d'où* vient le sens, mais *comment* il tient. C'est ce déplacement qui conduit à penser le sens comme une stabilisation opératoire. [2]

Stabilisation ne veut pas dire fixation définitive de la signification, mais arrêt provisoire de l'interprétation, suffisant pour permettre l'action. Un énoncé fait sens lorsqu'il est interprété de manière assez convergente pour orienter une décision, justifier une action ou soutenir une coordination. Cette convergence n'est jamais garantie par la structure interne de l'énoncé lui-même ; elle est produite par un ensemble de médiations — contextuelles, institutionnelles, techniques — qui rendent une interprétation praticable plutôt qu'une autre. Le sens est ainsi moins un contenu qu'un état transitoire d'équilibre interprétatif. [21]

Le qualificatif d'« opératoire » est ici décisif. Le sens ne se mesure plus à sa conformité à une vérité sémantique ou à une intention originnaire, mais à sa capacité à fonctionner dans un dispositif donné. Un énoncé est porteur de sens dans la mesure où il peut être utilisé, repris, intégré et évalué dans une chaîne d'opérations. Cette opérativité n'implique ni justesse ni

* L'utilisation du terme « naturaliser / naturalisation » dans ce texte signifie que certaines productions artefactuelles ou institutionnelles tendent à être reçues comme allant de soi : non parce qu'elles sont fondées ou légitimes, mais parce qu'elles sont disponibles, répétées, intégrées aux routines et soustraites à la possibilité de reprise critique. Naturaliser un sens ne signifie pas le justifier, mais le rendre perceptivement et socialement indiscutable. Naturalisation désigne l'effet par lequel des médiations techniques ou sociales se déposent sous forme d'évidence : une construction devient perçue comme naturelle, précisément parce que ses conditions de production ne sont plus visibles ni contestables.

validité au sens fort ; elle désigne simplement le fait que l'énoncé peut entrer dans un circuit d'usage sans provoquer immédiatement de rupture interprétative. [4]

Dans le contexte des productions automatisées, cette stabilisation est largement externalisée. Ce ne sont plus les agents producteurs qui stabilisent le sens, mais les cadres dans lesquels les réponses sont insérées : interfaces, protocoles, normes organisationnelles, attentes professionnelles, routines cognitives. Le sens se constitue au point d'usage, là où une réponse est sélectionnée parmi d'autres, interprétée comme suffisante, puis mobilisée comme base d'action. Conséquence majeure : la stabilité du sens devient dépendante de la robustesse des dispositifs qui la produisent. [28]

Toutefois, cette stabilisation opératoire est structurellement fragile. Elle repose sur des mécanismes d'alignement interprétatif qui peuvent céder sous la pression de la massification, de la vitesse et de la diversité des contextes d'usage. Plus les réponses sont nombreuses, rapides et hétérogènes, plus il devient difficile de maintenir des seuils partagés d'acceptabilité interprétative. Le sens ne disparaît pas, mais il tend à se fragmenter, à se localiser excessivement, voire à se dissoudre dans une multiplicité de micro-stabilisations incompatibles entre elles.

C'est ici que se manifeste la dimension politique du sens comme stabilisation opératoire. Décider ce qui compte comme une interprétation suffisante, ce qui peut être tenu pour « faisant sens », ce qui doit être corrigé ou rejeté, relève de choix normatifs, souvent implicites. Dans les régimes discursifs traditionnels, ces choix étaient portés par des communautés interprétatives relativement stables. Dans les régimes contemporains, ils sont de plus en plus inscrits dans des architectures techniques et organisationnelles, qui distribuent inégalement la capacité de stabiliser ou de contester le sens. [18][28]

Penser le sens comme stabilisation opératoire, ce n'est donc pas renoncer à toute exigence critique ; c'est au contraire se donner les moyens d'analyser les conditions concrètes sous lesquelles le sens est produit, maintenu ou perdu. Cette approche permet de déplacer l'analyse vers les seuils, les procédures et les dispositifs qui rendent le sens praticable, sans supposer une compréhension préalable ou une intention fondatrice.

Ce déplacement prépare le pas suivant : si le sens n'est ni un contenu ni une propriété intrinsèque, mais le résultat toujours provisoire d'opérations de stabilisation, alors il devient pertinent de parler non plus seulement de sens, mais de *régime du sens*. [19] C'est ce régime — entendu comme l'ensemble des processus, des opérateurs et des architectures qui rendent la stabilisation possible — que le chapitre suivant se propose d'examiner.

4. Le sens comme régime

Le raisonnement développé dans ce chapitre s'appuie explicitement sur le cadre théorique et analytique élaboré dans le texte de référence déposé sur [Zenodo : Pour une théorie et une pratique du sens à l'ère des modèles](#). Il ne s'agit ni d'en proposer un commentaire linéaire, ni d'en répéter les thèses, mais d'en prolonger les lignes de force en les recontextualisant dans une réflexion systématique sur le sens à l'ère des productions discursives automatisées. Les concepts mobilisés — régime, opérativité, stabilisation, répondabilité — doivent ainsi être compris comme des outils analytiques issus de ce travail préalable, et réélaborés ici en vue d'une théorie du sens adaptée aux conditions contemporaines de production et de circulation des énoncés. [39]

Dans cette perspective, parler de « régime du sens » revient à déplacer définitivement l'analyse hors d'une ontologie du contenu ou d'une psychologie de la signification. Le sens n'est plus envisagé comme une substance à interpréter, mais comme un ensemble de processus hétérogènes, distribués et historiquement situés, qui rendent certaines interprétations possibles, praticables et socialement opérantes plutôt que d'autres. C'est ce déplacement, du contenu vers le processus, que la section suivante entreprend de préciser. [2][21]

4.1 Du contenu au processus

Penser le sens comme régime implique d'abandonner définitivement une conception substantialiste de la signification. Tant que le sens est appréhendé comme un contenu — mental, sémantique ou informationnel — l'analyse demeure prisonnière de la question de son adéquation : adéquation à une intention, à une référence, à un état du monde ou à une norme de vérité. Or ce cadre se révèle insuffisant dès lors que les productions discursives ne sont plus garanties par une instance de compréhension ou par une intention attribuable. Le déplacement proposé consiste donc à substituer à une ontologie du contenu une analyse des processus par lesquels le sens advient, circule et se stabilise. [2]

Ce déplacement ne nie pas l'existence de contenus interprétables ; il en suspend la primauté explicative. Ce qui devient central, ce ne sont plus les propriétés internes des énoncés, mais les opérations qui les rendent interprétables dans des contextes donnés. Le sens n'est pas ce qui est contenu dans un énoncé, mais ce qui se produit lorsqu'un énoncé est inséré dans un ensemble de pratiques, de dispositifs et d'attentes. Autrement dit, le sens n'est pas donné ; il est produit, maintenu et parfois dissous par des processus. [21]

Ces processus sont hétérogènes. Ils incluent des opérations techniques (sélection, classement, hiérarchisation des réponses), des opérations sociales (validation par l'usage, légitimation institutionnelle), et des opérations cognitives distribuées (attention, confiance, suspension du doute). Aucun de ces niveaux ne suffit à lui seul à expliquer le sens ; c'est leur articulation qui constitue un régime. Parler de régime du sens, c'est donc reconnaître que le sens résulte d'un agencement dynamique de conditions, plutôt que d'une propriété stable des signes.

Cette approche permet également de rendre compte de la variabilité du sens sans tomber dans un relativisme généralisé. Le sens varie parce que les processus qui le stabilisent varient : selon les contextes, les architectures techniques, les normes professionnelles ou les contraintes temporelles. Mais cette variabilité n'est pas arbitraire. Elle est structurée par des régularités procédurales, des seuils implicites d'acceptabilité et des routines interprétatives. Le régime du sens désigne précisément cet ensemble de régularités qui rendent certaines interprétations probables et d'autres marginales.

Dans les régimes discursifs contemporains, cette dimension processuelle devient visible parce qu'elle est partiellement automatisée. Les opérations qui contribuaient autrefois de manière diffuse à la stabilisation du sens — filtrage, reformulation, hiérarchisation — sont désormais explicitement prises en charge par des systèmes techniques. Le sens apparaît alors moins comme un donné que comme un effet : effet de chaînes de traitement, d'interfaces, de protocoles et de contraintes d'usage. Cette visibilité accrue ne simplifie pas l'analyse, elle en déplace les enjeux vers les lieux où ces processus sont conçus, paramétrés et gouvernés.

Enfin, penser le sens comme processus permet de reformuler la question normative sans revenir à une théorie de la compréhension. La normativité du sens ne réside plus dans la conformité d'un contenu à une signification correcte, mais dans la qualité des processus qui le produisent. Un régime du sens peut être plus ou moins robuste, plus ou moins contestable, apte à intégrer la critique et la correction. Ce sont ces différences de qualité processuelle — et non la présence ou l'absence d'une compréhension interne — qui deviennent décisives pour évaluer les effets sociaux et politiques des productions discursives automatisées. [34]

En ce sens, le passage du contenu au processus ne constitue pas une simple reformulation théorique ; il marque un changement de niveau d'analyse. Le sens n'est plus ce qu'il faut interpréter, mais ce qu'il faut organiser. Cette organisation — ses opérateurs, ses seuils et ses architectures — sera examinée dans les sections suivantes, afin de préciser ce qui fait tenir, ou au contraire vaciller, un régime du sens.

4.2 Opérateurs du régime du sens

Si le sens est conçu comme un effet de processus, alors il devient possible — et nécessaire — d'en identifier les opérateurs. Par « opérateurs du régime du sens », on n'entend pas des causes premières ni des principes abstraits, mais des mécanismes concrets qui contribuent à rendre certaines interprétations praticables, crédibles et stabilisées. Ces opérateurs n'agissent jamais isolément : ils forment un agencement dynamique dont les effets dépendent de leur articulation, de leur hiérarchisation et de leur inscription dans des dispositifs techniques et institutionnels. [39]

Un premier opérateur fondamental est la sélection. Dans un environnement de production discursive massive, le sens ne peut émerger qu'à condition que certaines réponses soient retenues plutôt que d'autres. Cette sélection peut être algorithmique (classement, filtrage, priorisation), organisationnelle (procédures, *workflows*) ou cognitive (attention, fatigue, heuristiques). Ce qui est sélectionné tend à faire sens non parce qu'intrinsèquement meilleur, mais parce que cela devient disponible, visible et utilisable. La sélection est ainsi un opérateur de rareté artificielle dans un régime d'abondance. [28]

Un second opérateur est la stabilisation temporelle. Pour qu'un énoncé fasse sens, il doit demeurer interprétable pendant une durée suffisante pour être intégré à une action ou à une décision. Or les régimes contemporains tendent à réduire drastiquement cette durée : les réponses sont consommées, remplacées, mises à jour en continu. La stabilisation temporelle peut alors être assurée par des mécanismes de gel (versions, archivage, traçabilité) ou, au contraire, être sacrifiée au profit de la fluidité. Le choix entre ces options n'est jamais neutre : il conditionne la possibilité même d'une reprise critique.

Un troisième opérateur central est la contextualisation. Le sens n'advient pas dans l'abstraction, mais dans un cadre d'usage déterminé. Interfaces, formats de présentation, métadonnées, indications de confiance ou de domaine jouent ici un rôle décisif. La contextualisation ne consiste pas à ajouter du sens à un contenu brut, mais à orienter l'interprétation en réduisant l'espace des lectures possibles. Elle agit comme un opérateur de canalisation interprétative, souvent invisible, mais extrêmement puissant.

À ces opérateurs s'ajoute la validation, entendue non comme certification de vérité, mais comme reconnaissance pratique de suffisance. Une réponse est validée lorsqu'elle est jugée « assez bonne » pour l'usage envisagé. Cette validation peut être explicite (revue humaine,

approbation institutionnelle) ou implicite (réutilisation, absence de contestation). Dans les régimes automatisés, la validation tend à devenir tacite, intégrée aux routines, ce qui renforce l'efficacité tout en affaiblissant la réflexivité. [9]

Enfin, un opérateur crucial du régime du sens est la délégation. Les acteurs humains délèguent aux dispositifs techniques non seulement la production des réponses, mais aussi une part de l'effort interprétatif : évaluation de la pertinence, reformulation, anticipation des attentes. Cette délégation n'élimine pas la responsabilité, mais la redistribue. Elle crée des zones grises où le sens fonctionne sans que les conditions de sa stabilisation soient pleinement assumées par une instance identifiable. [20][34]

Ces opérateurs ne sont ni bons ni mauvais en eux-mêmes. Leur effet dépend de la manière dont ils sont combinés et gouvernés. Un régime du sens devient problématique lorsque ces opérateurs convergent vers une maximisation de la fluidité et de la plausibilité au détriment de la réponsabilité. À l'inverse, un régime plus robuste est celui qui maintient des points de friction, des seuils explicites et des possibilités de reprise critique. [30]

Identifier les opérateurs du régime du sens, ce n'est donc pas seulement décrire un état de fait ; c'est ouvrir un espace d'intervention. En rendant visibles les mécanismes qui font tenir le sens, on rend également visibles les lieux où il peut être corrigé, réorienté ou rendu plus gouvernable. Cette analyse prépare le dernier déplacement du chapitre : montrer que ces opérateurs ne sont pas seulement des pratiques diffuses, mais qu'ils sont de plus en plus inscrits dans des architectures spécifiques, qui conditionnent en profondeur le régime du sens contemporain. [34]



Figure 2. Le sens comme effet de processus : cinq opérateurs de stabilisation

Le graphique ci-dessus ne représente pas le sens comme une chose fixe, une essence ou un contenu préexistant, mais plutôt comme un effet émergent,

quelque chose qui apparaît (et peut disparaître) quand cinq mécanismes spécifiques s'articulent correctement, cinq opérateurs qui travaillent ensemble pour stabiliser une interprétation au point qu'elle devienne praticable, crédible et suffisamment durable dans un contexte donné.

4.3 Le tournant architectural

L'identification des opérateurs du régime du sens conduit à une conclusion décisive : ces opérateurs ne relèvent plus principalement de pratiques interprétatives diffuses ou de conventions tacites, mais sont de plus en plus intégrés à des architectures techniques et organisationnelles explicites. Le sens ne se stabilise plus seulement à travers des usages partagés ou des traditions interprétatives ; il est désormais conditionné par des choix de conception, d'implémentation et de gouvernance. C'est ce déplacement que l'on peut qualifier de tournant architectural. [19][21]

Par « architecture », il ne faut pas entendre uniquement l'infrastructure technique des systèmes, mais l'ensemble des agencements qui organisent la production, la circulation et la reprise des réponses : interfaces, *pipelines* de traitement, mécanismes de traçabilité, procédures de validation, seuils d'intervention humaine. Ces architectures ne produisent pas directement du sens, mais elles déterminent ce qui peut en faire, à quelles conditions et avec quels effets. Elles fonctionnent comme des cadres de possibilité interprétative.

Avec ce tournant, la stabilisation du sens devient en grande partie une question de *design*. Décider où placer un seuil de validation, rendre visibles ou non les sources, permettre ou empêcher la contestation, offrir des mécanismes de correction ou de retour d'erreur — autant de choix architecturaux qui orientent la manière dont le sens est endossé. Là où la philosophie du langage interrogeait la signification des énoncés, l'analyse contemporaine doit interroger la configuration des systèmes qui les rendent opérants. [20]

Cette architecturalisation du sens modifie également la distribution du pouvoir interprétatif. Les acteurs capables de concevoir, de paramétrer et de maintenir ces architectures acquièrent une capacité accrue à orienter ce qui fait sens et ce qui ne le fait pas. À l'inverse, les utilisateurs finaux peuvent se trouver enfermés dans des régimes de sens où les conditions de stabilisation sont opaques, non négociables ou difficilement contestables. Le sens cesse alors d'être un espace de discussion pour devenir un effet de configuration. [27][28]

Cependant, le tournant architectural ne conduit pas nécessairement à une fermeture du sens. Il rend au contraire visibles les lieux où une intervention est possible. En explicitant les opérateurs du régime du sens dans des dispositifs concrets, il devient possible d'y introduire des points de friction, des mécanismes de reprise et des procédures de réponsabilité. L'architecture peut être conçue non comme un moyen de neutraliser la critique, mais comme un support de l'agentivité interprétative.

C'est précisément sur ce point que se joue la transition vers le chapitre suivant. Si le sens est désormais médiatisé par des architectures, alors la question centrale n'est plus seulement celle de la signification, mais celle de la réponsabilité : qui peut répondre de ce que ces architectures produisent, comment cette réponse peut-elle être exigée, et dans quelles conditions peut-elle être contestée. Le tournant architectural prépare ainsi le terrain d'une

reconfiguration de l'agentivité, non plus comme propriété d'un sujet parlant, mais comme capacité distribuée à endosser, corriger et gouverner le sens. [1][34]

Le chapitre 5 examinera cette reconfiguration à partir des notions de délégation interprétative, d'imputabilité et d'autorité, afin de comprendre comment, dans un régime architecturalisé du sens, la discussion peut encore avoir lieu — ou au contraire entrer en crise.

5. Répondabilité et agentivité interprétative

La notion de répondabilité constitue le point d'articulation central de l'analyse qui suit. Elle désigne la condition sous laquelle une production discursive peut être prise en charge, interrogée et éventuellement corrigée dans un cadre donné. Être répondable ne signifie pas être à l'origine intentionnelle d'un énoncé, ni en garantir la vérité, mais pouvoir en assumer publiquement les effets, les usages et les conséquences. La répondabilité ne relève donc ni d'une psychologie de l'agent, ni d'une propriété interne des systèmes, mais d'un agencement de rôles, de procédures et de dispositifs qui rendent possible une réponse au sens fort : une réponse à la question « *qui répond de cela ?* ». [1][4]

Dans les régimes discursifs traditionnels, la répondabilité était largement adossée à la figure du locuteur : celui qui parle peut être interpellé, mis en cause, sommé de s'expliquer. L'architecture contemporaine des productions langagières dissocie cette unité. Les énoncés circulent sans auteur identifiable, les décisions sont prises à partir de réponses produites ailleurs, et les effets se manifestent à distance. La répondabilité ne disparaît pas pour autant ; elle se déplace, se distribue et parfois se dilue entre plusieurs instances humaines et techniques. [14][18]

Ce déplacement impose de repenser l'agentivité interprétative. Si aucun sujet ne peut être tenu pour le porteur exclusif du sens, alors l'agentivité réside dans la capacité à reprendre, cadrer et transformer les interprétations à l'intérieur d'un régime donné. La question n'est plus de savoir qui parle, mais qui peut intervenir sur ce qui fait sens, à quel moment et selon quelles modalités. La répondabilité apparaît ainsi comme une propriété émergente des dispositifs, et non comme une qualité morale ou cognitive des agents. [18][21]

C'est à partir de ce cadre que le chapitre 5 analysera les formes contemporaines de délégation interprétative, les tensions entre imputabilité et responsabilité, et les transformations de l'autorité interprétative dans des régimes où le sens est produit sans compréhension mais doit néanmoins pouvoir être assumé.

5.1 Délégation interprétative

La délégation interprétative constitue l'un des mécanismes centraux par lesquels le régime contemporain du sens parvient à fonctionner malgré l'absence de compréhension et d'intention du côté de la production. Elle désigne le processus par lequel des acteurs humains confient à des dispositifs techniques — et, plus largement, à des architectures organisationnelles — une part croissante du travail d'interprétation : sélection des réponses pertinentes, reformulation des questions, hiérarchisation des options, et parfois anticipation des usages. Cette délégation ne se limite pas à un gain d'efficacité ; elle transforme en profondeur la structure même de l'agentivité interprétative. [20]

Dans les régimes discursifs classiques, interpréter impliquait un engagement direct : comprendre un énoncé, en évaluer la pertinence, en assumer les implications. La délégation contemporaine dissocie ces moments. Les systèmes produisent des réponses pré-interprétées, c'est-à-dire déjà ajustées à un contexte d'usage supposé, tandis que les acteurs humains interviennent en aval, souvent de manière minimale, pour entériner ou rejeter ces propositions. L'interprétation devient alors une opération distribuée, fragmentée et partiellement automatisée.

Cette distribution repose sur un pacte implicite : les dispositifs sont supposés produire des réponses suffisamment plausibles pour que l'effort critique puisse être réduit. La délégation interprétative fonctionne tant que cette plausibilité est maintenue et que les écarts restent invisibles ou socialement tolérables. Le sens se stabilise alors par défaut, non parce qu'il est activement endossé, mais parce qu'il n'est pas remis en question. La délégation agit ainsi comme un opérateur de désactivation critique. [30]

Toutefois, cette délégation n'annule pas la responsabilité interprétative ; elle la reconfigure. Les acteurs humains restent ceux qui utilisent, diffusent ou institutionnalisent les réponses produites. Mais leur capacité à répondre de ces usages dépend désormais de leur accès aux conditions de production, de leur compréhension des limites des dispositifs et de leur pouvoir d'intervention sur les architectures. Lorsque ces conditions ne sont pas réunies, la délégation devient une source de fragilisation du sens : les réponses circulent, mais personne ne peut véritablement en répondre. [34]

La délégation interprétative introduit également une asymétrie structurelle entre ceux qui conçoivent les dispositifs et ceux qui les utilisent. Les premiers disposent d'un pouvoir interprétatif indirect, inscrit dans les paramètres, les choix de *design* et les critères de performance. Les seconds sont confrontés à des réponses déjà cadrées, dont les présupposés sont souvent invisibles. Cette asymétrie contribue à déplacer l'autorité interprétative hors de l'espace public de la discussion, vers des lieux techniques et organisationnels difficilement accessibles. [28]

Enfin, la délégation interprétative produit un effet cumulatif. Plus les dispositifs sont utilisés, plus leurs réponses sont intégrées aux routines, et plus il devient coûteux, cognitivement et institutionnellement, de suspendre la délégation. Le sens se naturalise : ce qui est produit est pris pour ce qui va de soi. La délégation cesse alors d'être perçue comme un choix et devient une condition de fonctionnement ordinaire du régime du sens.

Analyser la délégation interprétative, ce n'est donc pas dénoncer un abandon de la pensée critique, mais mettre au jour les conditions sous lesquelles cette pensée peut encore être exercée. La question centrale devient celle de l'articulation entre délégation et répondabilité : comment déléguer sans dissoudre la capacité de répondre, comment automatiser sans rendre le sens irresponsable. C'est cette tension que la section suivante abordera en distinguant plus précisément imputabilité, répondabilité et responsabilité.

5.2 Imputabilité, répondabilité, responsabilité

L'analyse de la délégation interprétative impose de clarifier les catégories normatives mobilisées pour penser la prise en charge des productions discursives. Les notions d'imputabilité, de répondabilité et de responsabilité sont souvent utilisées de manière interchangeable, alors qu'elles renvoient à des régimes distincts d'attribution et

d'engagement. Dans le contexte des productions automatisées, leur distinction devient cruciale, car elle conditionne la possibilité même d'une gouvernance du sens. [34]

L'**imputabilité** désigne la capacité à attribuer un énoncé, une décision ou un effet à une instance identifiable. Elle opère selon une logique de traçabilité : qui a produit, validé ou déployé le dispositif à l'origine de la réponse ? Dans les régimes techniques contemporains, l'imputabilité est souvent partielle et fragmentée. Les chaînes de production sont longues, distribuées, et les responsabilités formelles se superposent sans se recouvrir. L'imputabilité permet de localiser des points de causalité, mais elle ne suffit pas à rendre compte du sens des productions ni de leurs usages. [35]

La **responsabilité**, quant à elle, relève d'un registre plus normatif et souvent juridique. Elle engage une obligation de répondre des conséquences, qu'elles aient été prévues ou non. La responsabilité suppose des cadres institutionnels capables de sanctionner, de réparer ou de réguler. Or, dans les régimes de production automatisée, la responsabilité tend à être soit diluée — en raison de la multiplicité des acteurs impliqués — soit déplacée vers les utilisateurs finaux, qui se voient contraints d'assumer des effets dont ils ne maîtrisent ni les causes ni les conditions. [25][37]

Quant à la **répondabilité**, telle qu'introduite en ouverture du chapitre, elle doit être située par contraste. Là où l'imputabilité organise une attribution (qui est impliqué dans la chaîne de production ?), et où la responsabilité engage un régime d'obligations et de conséquences (réparation, sanction, régulation), la répondabilité désigne la possibilité pratique d'une mise en question du sens : rendre explicitable ce qui a été produit, ses présupposés, ses limites et les conditions de son usage. [1]

Cette distinction éclaire une configuration désormais fréquente : des dispositifs peuvent être imputables — on peut remonter à une organisation, un service, un système — sans être pour autant répondables, c'est-à-dire sans que cette attribution ouvre un espace effectif de discussion, de justification ou de correction. Dans ce cas, l'imputabilité tend à fonctionner comme clôture administrative (« *c'est ce système-là* ») plutôt que comme ouverture discursive (« *de quoi cela répond-il ?* »).

La répondabilité joue alors le rôle d'opérateur critique : elle ne garantit pas la vérité, mais maintient la contestabilité des énoncés au moment même où ils orientent l'action. Elle déplace l'enjeu d'un traitement a posteriori des dommages vers une capacité d'interruption et de reprise en cours d'usage, condition minimale pour éviter que le sens produit ne s'impose indifféremment comme une évidence.

Enfin, distinguer ces trois notions permet de déplacer le débat éthique et politique. La question n'est plus seulement de savoir qui est responsable en cas de préjudice, mais comment rendre les systèmes répondables au quotidien. Cette perspective ouvre un espace d'intervention qui n'est ni strictement juridique ni purement moral, mais architectural et procédural. [19][34] Elle prépare ainsi l'analyse de la section suivante, consacrée à l'autorité interprétative et à la crise contemporaine de la discussion.

5.3 Autorité interprétative et crise de la discussion

L'ensemble des déplacements analysés dans ce chapitre converge vers une transformation profonde de l'autorité interprétative. Dans les régimes discursifs classiques, cette autorité

était liée, de manière certes contestable mais relativement stable, à des figures identifiables : locuteurs, experts, institutions savantes ou instances politiques. La possibilité de la discussion reposait sur la reconnaissance de ces figures comme répondantes possibles du sens, c'est-à-dire comme susceptibles d'être interrogées, contredites ou mises en demeure de s'expliquer. La crise contemporaine ne tient pas à la disparition de l'autorité en tant que telle, mais à sa reconfiguration silencieuse. [9][28]

Dans un régime architecturalisé du sens, l'autorité interprétative se déplace vers les dispositifs qui organisent la production et la circulation des réponses. Ce déplacement n'implique pas nécessairement une intention de domination ou de contrôle ; il résulte de la convergence entre délégation interprétative, massification des énoncés et automatisation des opérations de stabilisation. Ce qui fait autorité n'est plus ce qui peut être argumenté, mais ce qui est disponible, fluide et intégré aux routines décisionnelles. L'autorité cesse d'être discursive pour devenir fonctionnelle. [27]

Cette transformation a un effet direct sur la possibilité même de la discussion. Discuter suppose non seulement des désaccords, mais des prises possibles sur le sens : des points où l'interprétation peut être suspendue, reformulée ou contestée. Or, lorsque les réponses sont produites sans auteur identifiable et stabilisées par des architectures opaques, ces points d'appui tendent à disparaître. La discussion ne s'éteint pas ; elle se déplace vers des marges, souvent déconnectées des lieux où les décisions sont effectivement prises.

La crise de la discussion ne se manifeste donc pas par un excès de conflits interprétatifs, mais par leur neutralisation en amont. Les réponses « font autorité » non parce qu'elles sont reconnues comme justes, mais parce qu'elles sont difficilement contestables : trop rapides, trop intégrées, trop techniques pour être suspendues. Le débat se trouve court-circuité par la fluidité même du régime du sens. Cette neutralisation est d'autant plus efficace que rarement perçue comme telle, elle se présente sous les traits de l'évidence, de l'optimisation ou de la rationalité procédurale.

Une crise de la discussion qui ne relève pas seulement du régime du sens : elle est indissociable du régime de la communication, c'est-à-dire des conditions fiduciaires de circulation, d'autorité en chaîne et de reprise. La responsabilité se défait autant par défaut d'endossement local que par saturation des médiations qui rendent la contestation impraticable. [9][27]

Dans ce contexte, la responsabilité apparaît comme la condition minimale de possibilité d'une autorité non autoritaire. Une autorité interprétative qui ne peut pas être interrogée cesse d'être une autorité au sens fort pour devenir un simple fait de pouvoir. Rendre le sens répondable, ce n'est pas abolir l'autorité, mais la réinscrire dans un espace de justification et de contestation. Sans cette réinscription, la discussion devient purement expressive, sans prise sur les processus qui produisent effectivement le sens. [18]

La crise de la discussion est ainsi indissociable d'une crise de l'agentivité interprétative. Lorsque les acteurs ne disposent plus des moyens d'intervenir sur les conditions de stabilisation du sens, leur capacité à discuter se réduit à des commentaires périphériques. La question n'est donc pas seulement démocratique ou épistémique, elle est architecturale. Elle concerne les dispositifs concrets qui permettent — ou empêchent — que la discussion ait des effets.

Ce constat ouvre directement sur le chapitre suivant. Si l'autorité interprétative est désormais inscrite dans des architectures, alors la tâche critique ne consiste plus seulement à analyser des discours, mais à concevoir des dispositifs capables de rendre la réponsabilité praticable. Le chapitre 6 s'attachera à ces architectures du sens, en examinant les seuils, les littératies et les écologies nécessaires pour que le sens puisse encore être assumé, discuté et gouverné.

6. Architectures du sens : rendre la réponsabilité praticable

Les chapitres précédents ont montré que la crise contemporaine du sens ne tient ni à une perte de signification, ni à un déficit de rationalité, mais à une fragilisation des conditions de réponsabilité. Lorsque le sens est produit sans compréhension et stabilisé par des architectures opaques, la question n'est plus seulement théorique ou normative : elle devient une question de conception. Si la discussion doit encore avoir une prise sur ce qui fait sens, alors les régimes discursifs doivent être pensés en termes de dispositifs concrets, capables d'organiser la production, la circulation et la reprise des interprétations. [34][39]

Ce chapitre propose ainsi un déplacement résolument pragmatique. Il ne s'agit pas d'identifier un modèle idéal de gouvernance du sens, mais d'examiner les architectures à travers lesquelles la réponsabilité peut être rendue praticable, sans supposer une compréhension préalable ni une maîtrise totale des systèmes. Les sections qui suivent analyseront les seuils, les compétences et les écologies nécessaires pour que le sens puisse être stabilisé sans être soustrait à la contestation. [20]

6.1 Dispositifs à seuils multiples

Rendre la réponsabilité praticable dans un régime de production discursive massive suppose de rompre avec une conception binaire de la validation et de la responsabilité. Les architectures traditionnelles reposaient souvent sur des seuils uniques : un auteur identifié, une instance de validation centrale, une décision finale clairement attribuable. Or ces modèles deviennent inopérants dès lors que les réponses sont produites en continu, à grande échelle, grande vitesse, et intégrées à des usages hétérogènes. La réponsabilité ne peut plus être pensée comme un point fixe ; elle doit être distribuée à travers une pluralité de seuils. [34][35]

Par « dispositifs à seuils multiples », on désigne des architectures dans lesquelles la stabilisation du sens s'opère par étapes, selon des niveaux différenciés d'engagement interprétatif. Chaque seuil correspond à une intensité particulière de prise en charge : exposition minimale à l'usage, validation contextuelle, révision critique, voire suspension ou retrait. Le sens n'est pas immédiatement arrêté ; il est progressivement endossé, avec la possibilité de revenir en arrière à mesure que les enjeux augmentent.

Un premier seuil est celui de la plausibilité opératoire. À ce niveau, les réponses sont jugées acceptables pour un usage exploratoire, informel ou réversible. La réponsabilité y est faible, mais assumée comme telle : le dispositif rend visible le caractère provisoire et non garanti des interprétations proposées. Ce seuil permet de maintenir la fluidité sans naturaliser le sens.

Un second seuil concerne la validation contextuelle. Lorsque les réponses commencent à orienter des décisions situées — professionnelles, organisationnelles, éducatives — un

premier cadre de justification devient nécessaire. Le dispositif doit alors permettre de retracer les conditions générales de production, d'indiquer les limites d'usage et d'offrir des points d'intervention humaine. La responsabilité n'est pas encore pleine, mais elle devient activable.

Un troisième seuil correspond à l'engagement institutionnel. À ce niveau, les réponses produites sont intégrées à des procédures formelles, des politiques ou des décisions à effets durables. La responsabilité exige alors des mécanismes explicites de contrôle, de documentation et de correction. Ce seuil n'implique pas une transparence totale des systèmes, mais une capacité effective à suspendre, auditer ou réviser les interprétations stabilisées. [25][37]

Ce qui distingue un dispositif à seuils multiples d'un régime rigide, ce n'est pas seulement la pluralité des niveaux, mais la possibilité de circulation entre eux. Un sens stabilisé à un seuil faible doit pouvoir être réinterrogé à un seuil plus élevé lorsque les conséquences changent. Inversement, un excès de validation précoce rigidifie le régime du sens et neutralise la discussion avant qu'elle ne soit nécessaire. [11]

Ces architectures introduisent également une temporalité du sens. La responsabilité n'est plus un état, mais une trajectoire : elle augmente ou se renforce à mesure que le sens est mobilisé. Cette temporalité permet de concilier efficacité et critique, en évitant à la fois la paralysie par excès de contrôle et l'irresponsabilité par défaut de prise en charge.

Enfin, les dispositifs à seuils multiples redistribuent l'agentivité interprétative. Ils ne cherchent pas à tout confier à une instance centrale, mais à organiser la coopération entre systèmes, utilisateurs et institutions. La responsabilité devient alors une propriété émergente de l'architecture elle-même, plutôt qu'une exigence morale imposée aux seuls acteurs humains. [18][21]

Cette approche ne garantit pas l'absence d'erreurs ni de conflits interprétatifs. Elle vise plus modestement à rendre ces conflits visibles, situables et traitables. En ce sens, les dispositifs à seuils multiples constituent une condition minimale pour qu'un régime du sens automatisé demeure gouvernable.

La section suivante examinera un autre levier essentiel de cette gouvernabilité : les littératies du sens, c'est-à-dire les compétences nécessaires pour reconnaître les seuils, activer la responsabilité et intervenir de manière informée dans ces architectures.

6.2 Littératies du sens à l'ère des modèles

Les architectures du sens, même conçues selon des dispositifs à seuils multiples, ne suffisent pas à garantir la responsabilité si elles ne sont pas accompagnées de formes adéquates de littératie. Par littératies du sens, on n'entend pas un simple ensemble de compétences techniques ou informationnelles, mais la capacité pratique à reconnaître les conditions dans lesquelles une réponse fait sens, à identifier les seuils d'engagement interprétatif, et à intervenir sur les processus de stabilisation lorsque cela est nécessaire. À l'ère des modèles, cette littératie devient une condition centrale de l'agentivité interprétative. [10][13]

Les régimes discursifs traditionnels supposaient une littératie principalement sémantique et argumentative : comprendre un texte, évaluer une thèse, répondre par des raisons. Les

régimes contemporains exigent une littératie supplémentaire, de nature procédurale et situationnelle. Il s'agit moins de juger la signification d'un énoncé que d'évaluer son statut : à quel seuil se situe-t-il ? Dans quel cadre peut-il être utilisé ? Quelles garanties lui sont associées, et lesquelles font défaut ? Cette compétence ne relève pas de la maîtrise des modèles, mais de la compréhension de leurs conditions d'usage. [34]

Une première dimension de cette littératie est la capacité de désambiguïsation des statuts du sens. Les réponses produites par les modèles tendent à se présenter sous une forme homogène, quel que soit leur degré de fiabilité ou d'engagement. La littératie du sens consiste à ne pas confondre plausibilité, validité et autorité, et à reconnaître les indices — explicites ou implicites — qui signalent le niveau de répondabilité requis. Sans cette capacité, les dispositifs à seuils multiples risquent d'être perçus comme arbitraires ou inutiles.

Une seconde dimension est la compétence d'intervention interprétative. Être littéré au sens, ce n'est pas seulement savoir évaluer une réponse, mais savoir quand et comment la suspendre, la reformuler ou en demander justification. Cette compétence est profondément liée à la conception des interfaces et des procédures : elle suppose que des points d'intervention soient effectivement accessibles, et que leur usage soit légitime. Là où les architectures ferment ces possibilités, la littératie se transforme en impuissance. [4]

Une troisième dimension concerne la conscience des effets cumulatifs du sens. Les réponses produites par les modèles ne font pas sens isolément ; elles s'inscrivent dans des séries, des habitudes et des routines. La littératie du sens implique la capacité à percevoir ces dynamiques de sédimentation : ce qui est accepté aujourd'hui comme une aide ponctuelle peut devenir demain une norme implicite. Reconnaître ces glissements est essentiel pour maintenir la répondabilité dans la durée.

Enfin, les littératies du sens ne peuvent être pensées uniquement à l'échelle individuelle. Elles sont aussi collectives et institutionnelles. Organisations, professions et communautés doivent développer des pratiques partagées pour qualifier, discuter et gouverner les usages des réponses automatisées. Sans ces cadres collectifs, la charge interprétative retombe sur des individus isolés, sommés de répondre de décisions qu'ils ne contrôlent pas pleinement. [28]

À l'ère des modèles, la littératie du sens n'est donc ni un supplément éducatif ni une compétence accessoire. Elle constitue l'une des conditions de possibilité d'un régime du sens répondable. Sans elle, même les architectures les mieux conçues risquent de produire un sens efficace mais irresponsable. La section suivante élargira encore la focale en proposant une approche écologique du sens, attentive aux interactions entre dispositifs, acteurs et environnements.

6.3 Écologie du sens répondable

Penser la répondabilité du sens à l'ère des productions discursives automatisées impose de dépasser une approche strictement architecturale ou normative. Les dispositifs à seuils multiples et les littératies du sens ne deviennent pleinement opérants qu'en les envisageant comme composantes d'une écologie plus large, dans laquelle interagissent des acteurs, des institutions, des technologies et des environnements symboliques. Parler d'écologie du sens répondable, c'est reconnaître que le sens ne se stabilise jamais isolément, mais toujours dans un tissu de relations, de dépendances et de rétroactions. [11][21]

Cette écologie est caractérisée par une pluralité d'échelles. À l'échelle micro, le sens se joue dans des situations d'usage concrètes : une réponse est acceptée, contestée ou reformulée. À l'échelle méso, il se stabilise dans des organisations, des professions et des communautés interprétatives qui développent des routines, des normes et des seuils partagés. À l'échelle macro, il est influencé par des infrastructures techniques, des régimes économiques et des cadres réglementaires qui orientent ce qui peut circuler, être rendu visible ou être interrogé. La réponsabilité dépend de la cohérence — ou de la dissonance — entre ces échelles. [23][27][37]

Une écologie du sens réponsable se distingue d'un régime purement extractif du sens, où les réponses sont produites et consommées sans souci de leurs conditions d'endossement. Elle suppose au contraire une attention constante aux effets cumulés de la production discursive : saturation cognitive, naturalisation des interprétations, érosion des capacités critiques. Ces effets ne sont pas des dérives accidentelles, mais des conséquences systémiques d'un régime de fluidité maximale. Les corriger exige non pas des ajustements ponctuels, mais une reconfiguration des relations entre production, usage et critique. [29]

Dans cette perspective, la réponsabilité apparaît comme un principe d'équilibre écologique. Trop faible, elle conduit à un sens efficace mais irresponsable ; trop exigeante, elle risque de paralyser l'action par un excès de justification. L'enjeu n'est donc pas de maximiser la réponsabilité, mais de l'ajuster aux contextes, aux enjeux et aux conséquences. Une écologie du sens réponsable est une écologie des seuils, des temporalités et des responsabilités distribuées.

Cette approche permet également de reformuler la question politique du sens. Gouverner le sens ne signifie pas en contrôler les contenus, mais en organiser les conditions de circulation, de contestation et de révision. Dans une écologie réponsable, le conflit interprétatif n'est pas un dysfonctionnement à éliminer, mais un indicateur de vitalité. Ce qui importe, ce n'est pas l'absence de désaccord, mais l'existence de dispositifs capables de le rendre traitable. [21][28]

Enfin, penser le sens en termes écologiques permet de sortir d'une opposition stérile entre autonomie humaine et automatisation technique. Le sens n'est ni entièrement produit par les machines, ni entièrement maîtrisé par les humains. Il émerge de leurs interactions, médiées par des architectures, des pratiques et des normes. La question décisive n'est donc pas de savoir si les systèmes comprennent, mais si les environnements dans lesquels ils opèrent permettent encore aux humains de répondre de ce qui fait sens. [1][39]

Cette Partie II a proposé un déplacement théorique : du sens comme contenu au sens comme régime, de la compréhension à la réponsabilité, de l'énoncé isolé à l'écologie des pratiques interprétatives. Ce déplacement ne clôt pas la réflexion ; il en définit les conditions. À l'ère des modèles, le sens ne peut plus être garanti par une intériorité cognitive, mais il peut encore être gouverné — à condition d'en reconnaître la fragilité, la distribution et la dépendance à des architectures que nous avons collectivement la responsabilité de concevoir. [39]

Afin de rendre explicite l'ancrage de cette reconstruction dans le texte de référence, une table de correspondances figure à l'**annexe J**. Elle met en regard les arguments développés dans cette Partie II et les sections du [texte source](#) dont ils procèdent ou qu'ils réélaborent. Cette annexe n'a pas pour fonction de justifier rétrospectivement les choix conceptuels, mais de

documenter le travail de déplacement, de recomposition et de systématisation opéré ici. Des annexes analogues accompagneront les Parties III (Communication) et IV (Métabolisme), selon le même principe.

PARTIE III — Le régime de la communication : circuler sans origine stable

De la transmission à la circulation algorithmique

Les théories classiques du langage et de la communication ont longtemps reposé sur un modèle de la transmission. Qu'il s'agisse de la parole, de l'écriture ou des médias de masse, la communication était pensée comme le passage d'un message d'un émetteur vers un récepteur, à travers un canal, selon des codes partagés. Même lorsque ce modèle a été critiqué — notamment pour son simplisme ou son ignorance des contextes —, il a continué de structurer implicitement les manières de penser l'autorité, la responsabilité et la validité des énoncés : une parole faisait sens parce qu'elle pouvait être rapportée à une origine identifiable, à une intention, à une situation d'énonciation. [4][9][38]

Cette partie soutient que ce modèle est désormais insuffisant pour comprendre les formes contemporaines de communication dans lesquelles s'inscrit l'intelligence artefactuelle. Les productions linguistiques n'y sont plus principalement transmises, elles circulent. Elles sont générées, reprises, transformées, recombinaées, classées et rediffusées dans des environnements algorithmisés où l'origine des énoncés est instable, fragmentée ou indifférente. La communication devient un milieu autonome, structuré par des logiques de convertibilité, de visibilité et de répétition, plutôt que par la fidélité à une source ou la cohérence d'une intention initiale. [27][28]

Dans ce régime de la circulation, l'autorité ne disparaît pas, elle change de nature. Elle ne repose plus principalement sur la position de l'énonciateur, mais sur des chaînes de validation distribuées, sur des effets de cohérence statistique, de fréquence, de format et de crédibilité accumulée. La mémoire elle-même se transforme : elle n'est plus garantie par la stabilité des supports ou des institutions, mais par la persistance computationnelle et la recomposition permanente des traces. Cette partie analysera ainsi le régime de la communication comme un régime autonome, irréductible à celui du sens, qui conditionne profondément les formes contemporaines de validité, de confiance et de responsabilité — et qui, lorsqu'il se déséquilibre, produit des pathologies spécifiques de circulation sans orientation. [28][29]

Pour comprendre comment la communication a pu devenir un régime autonome de circulation sans origine stable, il est nécessaire de revenir sur les transformations historiques des régimes du langage, non pour en retracer une évolution linéaire, mais pour identifier les seuils à partir desquels l'autorité, la mémoire et la responsabilité des énoncés ont cessé d'être indexées à une situation d'énonciation identifiable. [38]

7. Généalogie des régimes du langage

L'hypothèse selon laquelle la communication contemporaine constitue un régime autonome de circulation ne peut être soutenue qu'à condition de rompre avec une conception continue ou évolutionniste des formes du langage. Il ne s'agit pas de décrire une progression linéaire

allant de l'oralité à l'algorithmisation, mais d'identifier des régimes hétérogènes, définis par des modes spécifiques d'articulation entre énonciation, mémoire, autorité et responsabilité. [10][13][38]

Dans cette perspective, un régime du langage n'est ni un simple médium ni une technologie isolée, mais une configuration opératoire dans laquelle les énoncés acquièrent leur capacité à circuler, à persister et à produire des effets de validité. [5]

La généalogie proposée ici vise donc à dégager les seuils de désindexation à partir desquels l'autorité des énoncés cesse progressivement d'être rapportable à une situation d'énonciation identifiable, préparant l'émergence d'un régime communicationnel fondé non sur la transmission, mais sur la circulation. [27]

7.1 Oralité primaire et imputabilité située

Dans le régime de l'oralité primaire, la parole ne constitue pas un objet circulant, mais un acte localisé, inséparable de la situation qui le rend possible. L'énoncé n'y possède aucune autonomie matérielle ou symbolique : il n'existe qu'en tant qu'événement éphémère, produit dans une coprésence sociale et immédiatement exposé à l'évaluation collective.

L'autorité de la parole ne repose pas sur sa stabilité, mais sur son ancrage relationnel. Elle est fonction de la position sociale de l'énonciateur, de la reconnaissance de sa voix et de la mémoire partagée des interactions antérieures. La responsabilité est dès lors strictement située : elle n'est ni déléguée ni différable, mais directement imputable à un corps parlant.

La mémoire, dans ce régime, ne relève pas d'une conservation des traces, mais d'une réactivation continue. Elle est fondamentalement performative : ce qui n'est pas repris, répété ou ritualisé cesse d'exister comme énoncé opérant. L'oubli n'est pas une défaillance technique, mais une condition structurelle de la parole orale. [13]

Ce régime se caractérise ainsi par une coïncidence forte entre autorité, mémoire et responsabilité, rendue possible par la non-circulation des énoncés hors de leur contexte. La parole n'est pas conçue pour circuler, mais pour agir localement. C'est précisément cette non-circulabilité qui garantit son imputabilité.

Dans ce cadre, la responsabilité est directe et non déléguée. Une parole engage celui qui la profère, non seulement sur le plan symbolique, mais souvent sur le plan juridique ou rituel. La possibilité de la *parrhêsia* — au sens d'un dire-vrai risqué — est indissociable de cette imputabilité située : parler, c'est se rendre visible et vulnérable. La parole ne circule pas sans origine ; elle s'éteint avec celui qui la porte, sauf à être reprise dans un autre événement d'énonciation. [14]

7.2 Écriture, mémoire et prescriptions de lecture

L'introduction de l'écriture constitue un premier seuil de transformation décisif, dans la mesure où elle rend possible la dissociation entre l'énoncé et la situation de son énonciation. La parole devient inscriptible, stockable, transportable. Elle acquiert une persistance matérielle qui excède la mémoire des sujets. [10][13]

Cependant, cette autonomisation est immédiatement compensée par la mise en place de dispositifs de recontextualisation. Les régimes de l'écrit ne laissent pas les textes circuler librement : ils les insèrent dans des chaînes d'autorité (auteur, canon, institution, tradition)

et dans des cadres interprétatifs normés. La circulation reste subordonnée à la fidélité à une origine supposée, qu'il s'agisse d'une intention, d'un sens premier ou d'une autorité fondatrice. [14]

La mémoire écrite est donc une mémoire prescriptive. Elle conserve, mais prescrit également les conditions de lecture, d'interprétation et de transmission. L'oubli ne disparaît pas ; il est administré par des institutions qui décident de ce qui mérite d'être conservé, copié ou commenté. [10]

Dans ce régime, l'autorité se déplace sans se dissoudre. Elle n'est plus incarnée dans un corps présent, mais dans des figures stabilisées (l'auteur, le texte, la loi, l'archive). La responsabilité devient différée et distribuée, mais elle demeure indexée à une origine identifiable, même lorsque celle-ci est lointaine ou abstraite.

L'écriture inaugure ainsi une première forme de circulation, mais une circulation contrôlée, subordonnée à des régimes de légitimation et à une économie de la rareté.

7.3 Littératies et seuils de transformation

Les transformations ultérieures des régimes du langage ne peuvent être comprises qu'à travers la notion de littératie, entendue non comme simple compétence technique, mais comme ensemble de pratiques socialement et techniquement configurées permettant de produire et d'interpréter des énoncés. [10][13]

Chaque seuil de littératie redéfinit les conditions de circulation, sans pour autant abolir immédiatement l'indexation à l'origine. L'imprimerie, la presse, puis les médias audiovisuels étendent massivement la portée des énoncés, tout en maintenant des centres de production relativement identifiables. La circulation s'intensifie, mais elle demeure structurée par des asymétries énonciatives stables.

Le seuil contemporain, en revanche, introduit une rupture qualitative. Les littératies computationnelles ne se contentent pas d'augmenter la vitesse ou la quantité des échanges : elles transforment les énoncés en objets opératoires, soumis à des processus automatiques de conversion, de classement, de génération et de redistribution. Les énoncés ne sont plus principalement transmis ; ils sont mis en circulation dans un milieu algorithmique qui en conditionne la visibilité et la valeur. [27][28]

À partir de ce point, l'origine cesse d'être le principe organisateur du langage. L'autorité, la mémoire et la responsabilité se trouvent reconfigurées par des logiques de persistance computationnelle, de cohérence statistique et de crédit accumulé. Ce déplacement ne marque pas la disparition du sens, mais l'émergence d'un régime communicationnel autonome, irréductible aux modèles antérieurs de la transmission.

8. Convertibilité et textoralité

Le passage d'un régime de la transmission à un régime de la circulation ne s'opère pas uniquement par l'introduction de nouveaux supports techniques ou par l'accélération des flux communicationnels. Il repose sur une transformation plus profonde : celle du statut opératoire des énoncés. Dans le régime contemporain, les productions linguistiques ne sont plus principalement définies par leur contenu sémantique ou par leur intention énonciative,

mais par leur convertibilité, c'est-à-dire par leur capacité à être transformées, recombinaées et redéployées dans des environnements computationnels hétérogènes. [5][27]

La notion de convertibilité permet ainsi de penser la communication non comme une médiation entre sujets, mais comme un milieu de circulation autonome, au sein duquel les énoncés acquièrent une efficacité indépendamment de toute origine stable. C'est dans ce contexte qu'émerge ce que l'on peut nommer une *textoralité* : une forme hybride de langage, ni strictement orale ni strictement écrite, mais structurée par des opérations algorithmiques de traitement et de génération. [13]

8.1 De la médiation à la convertibilité généralisée

Dans les régimes classiques de la communication, la médiation désigne le passage d'un message d'un point à un autre à travers un dispositif technique ou institutionnel. Même lorsque cette médiation est complexe, elle reste subordonnée à une logique de transmission : quelque chose doit être préservé, traduit ou interprété à partir d'une source.

La convertibilité généralisée rompt avec cette logique. Un énoncé n'est plus conçu comme une unité à transmettre, mais comme un ensemble de traits exploitables, susceptibles d'être transformés selon des règles formelles. Texte, image, voix, métadonnées et contextes d'usage sont intégrés dans des chaînes de traitement où leur valeur dépend moins de leur signification que de leur compatibilité opérationnelle avec d'autres systèmes. [5]

Dans ce cadre, la communication cesse d'être orientée vers un destinataire déterminé. Elle devient un processus de mise à disposition des énoncés pour des usages futurs indéterminés. La circulation n'est plus secondaire par rapport à la production du sens ; elle en devient la condition première. Ce qui compte n'est pas ce qui est dit, mais ce qui peut être repris, recalculé, recontextualisé. [27][28]

La convertibilité ne désigne donc pas une simple traduction entre formats, mais une transformation du langage en ressource computationnelle, intégrée à des architectures techniques qui déterminent en amont les conditions de visibilité, de pertinence et de crédibilité.

8.2 Textoralité opérationnelle

La généralisation de la convertibilité produit une forme de langage spécifique, que l'on peut qualifier de *textoralité*. Celle-ci ne se réduit ni à l'oralité médiatisée ni à l'écriture numérisée. Elle se définit par son fonctionnement opératoire : les énoncés textoraux sont conçus pour être lus, mais aussi pour être dits, traités, indexés, générés et évalués par des systèmes automatisés. [13][27][38]

Dans ce régime, la distinction entre production et réception se brouille. Un même énoncé peut être simultanément :

- une réponse à une sollicitation,
- une donnée d'entraînement,
- un fragment réutilisable,
- un signal de crédibilité,

- ou un élément de classement.

La textoralité est ainsi caractérisée par une polyfonctionnalité structurelle. Elle ne vise pas prioritairement la compréhension humaine, mais l'intégration dans des chaînes de circulation où la lisibilité n'est qu'un paramètre parmi d'autres. La cohérence y est moins narrative que statistique ; la continuité y est moins logique que computationnelle. [26]

Ce régime transforme profondément la temporalité du langage. Les énoncés ne sont plus ancrés dans un présent d'énonciation, mais dans un temps de la persistance, où ils peuvent être réactivés, recombinaison ou réévalués longtemps après leur production initiale. La mémoire n'est plus une conservation fidèle, mais une recombinaison permanente des traces. [28]

8.3 Pragmatique de la parole machinique

La textoralité atteint son point de tension maximal avec l'émergence de la parole machinique. Celle-ci ne se contente pas d'automatiser la diffusion des énoncés : elle en automatise la production même, selon des régularités statistiques et des modèles de probabilité qui court-circuitent toute intentionnalité au sens classique. [26][30]

Du point de vue de ses effets de présence et de réception, cette parole présente un paradoxe constitutif. Elle est immédiatement reconnaissable comme langage — syntaxiquement correcte, pragmatiquement ajustée, contextuellement pertinente — tout en étant dépourvue de toute expérience vécue, de toute situation d'énonciation au sens fort. Elle parle sans avoir parlé, répond sans avoir entendu, produit du sens sans le viser. [2]

Ce paradoxe ne relève pas d'un déficit, mais d'un changement de régime. La parole machinique ne s'inscrit pas dans une relation intersubjective, elle opère dans un milieu de circulation où la validité des énoncés dépend de leur conformité à des distributions, à des formats et à des attentes probabilisées. L'autorité n'y est pas incarnée : elle devient un effet de chaîne, souvent calculé. [28]

La textoralité machinique rend ainsi manifeste l'autonomie du [régime communicationnel contemporain](#). Elle révèle un langage qui ne renvoie plus prioritairement au monde, ni même à des sujets parlants, mais à un espace de circulation auto-référentiel, structuré par des critères de performance, de crédibilité et de compatibilité. [28][29]

À ce stade, la question n'est plus de savoir comment le langage signifie, mais comment il acquiert du crédit, comment il persiste, et comment il engage — ou non — une responsabilité. [34]

Le chapitre 9 – *Autorité, mémoire et crédit* analysera précisément cette reconfiguration, en montrant comment la communication algorithmique produit de nouvelles formes d'autorité distribuée et de mémoire computationnelle, ainsi que les tensions épistémiques et fiduciaires qui en résultent.

9. Autorité, mémoire et crédit

L'autonomisation du régime communicationnel ne se manifeste pas seulement par une transformation des formes du langage, mais par une reconfiguration profonde des conditions de l'autorité, de la mémoire et de la responsabilité. Dans le régime de la circulation algorithmique, ces dimensions ne disparaissent pas ; elles cessent toutefois d'être indexées à

des sujets, à des intentions ou à des institutions stables. Elles se redistribuent au sein de chaînes computationnelles où la validité des énoncés dépend de leur persistance, de leur visibilité et de leur capacité à accumuler du crédit. [27][28]

Ce déplacement marque une rupture décisive avec les régimes antérieurs. Là où l'autorité reposait sur la position de l'énonciateur ou sur la stabilité des supports, elle devient désormais un effet émergent de la circulation elle-même. La mémoire n'est plus un lieu de conservation, mais un processus dynamique de réactivation et d'effacement. Le crédit, enfin, remplace progressivement la vérité comme principe organisateur de la confiance communicationnelle. [28][29]

9.1 Mémoire computationnelle : persistance et oubli

Dans le régime communicationnel contemporain, la mémoire ne peut plus être pensée sur le modèle de l'archive classique. Elle n'est ni un stock passif de traces ni une garantie de fidélité au passé. Elle prend la forme d'une mémoire computationnelle, définie par des opérations de stockage, d'indexation, de pondération et de réactivation automatisées. [10][28]

La persistance des énoncés n'y est jamais absolue. Elle dépend de paramètres techniques (formats, protocoles, architectures), économiques (coûts de stockage, valeur des données) et algorithmiques (classement, recommandation, filtrage). Un énoncé persiste non parce qu'il est vrai ou légitime, mais parce qu'il reste opérable dans des chaînes de traitement. [27]

Inversement, l'oubli n'est plus principalement le résultat d'une dégradation matérielle ou d'un effacement intentionnel. Il devient une fonction systémique, produite par l'obsolescence des formats, la dépriorisation algorithmique ou la perte de compatibilité. Ce qui est oublié n'est pas ce qui est supprimé, mais ce qui cesse d'être recalculé.

Cette mémoire computationnelle ne stabilise pas le passé ; elle le recompose en permanence. Les énoncés peuvent être réactivés hors de tout contexte d'origine, fragmentés, agrégés à d'autres traces, et dotés de nouvelles valeurs d'usage. La mémoire cesse ainsi d'être un garant de continuité symbolique pour devenir un opérateur de circulation différentielle. [28]

9.2 Autorité en chaîne et asymétrie structurelle

Dans ce contexte, l'autorité ne peut plus être attribuée à une source identifiable. Elle ne disparaît pas pour autant : elle se déplace vers des chaînes de validation distribuées, où chaque opération — classement, recommandation, reformulation, génération — contribue à produire des effets de crédibilité.

L'autorité devient ainsi cumulative et différentielle. Elle ne repose plus sur un acte fondateur, mais sur une série de confirmations partielles, souvent opaques, qui renforcent la visibilité et la légitimité apparente des énoncés. Ce processus produit une autorité sans auteur, mais non sans structure.

Cette autorité en chaîne est profondément asymétrique. Les conditions de circulation ne sont pas équitablement distribuées : certains formats, certaines régularités discursives, certains styles énonciatifs sont systématiquement favorisés. L'autorité ne dépend plus de ce qui est dit, mais de la conformité aux critères implicites du milieu communicationnel.

Il en résulte une dissociation croissante entre autorité et responsabilité. Là où les régimes antérieurs permettaient, au moins en principe, de remonter des énoncés vers des instances

imputables, le régime communicationnel contemporain dilue cette possibilité dans une multiplicité d'opérations intermédiaires. La responsabilité devient structurellement difficile à assigner, non par absence de causes, mais par excès de médiations.

9.3 De la crise épistémique à la crise fiduciaire

Cette transformation est souvent décrite en termes de crise de la vérité ou de crise épistémique. Une telle description demeure toutefois insuffisante. Le problème central n'est pas seulement la production de faux énoncés, mais la reconfiguration des conditions de la confiance.

Dans le régime de la circulation algorithmique, la confiance ne repose plus prioritairement sur des critères de véracité, mais sur des indicateurs de crédit : fréquence d'apparition, cohérence statistique, reconnaissance implicite, conformité aux attentes. Un énoncé est crédible non parce qu'il est fondé, mais parce qu'il circule bien. [29]

Il s'agit donc moins d'une crise épistémique que d'une crise fiduciaire. Les sujets continuent d'accorder leur confiance, mais ils la placent dans des structures de circulation dont les logiques leur échappent largement. La communication ne cesse pas de fonctionner, elle fonctionne sans orientation normative stable.

Cette crise fiduciaire n'est pas un accident, mais une conséquence structurelle de l'autonomie du régime communicationnel. Lorsque la circulation devient le principe premier, la validité se mesure à la capacité d'un énoncé à s'inscrire durablement dans le flux, et non à sa correspondance avec un monde ou à l'engagement d'un sujet.

Il reste toutefois à analyser les figures-limites produites par ce régime : là où la circulation sans origine stable atteint ses points de tension maximale, et où la question de la responsabilité revient sous des formes paradoxales.

10. Figures-limites du régime de la communication

Les régimes du langage ne se manifestent jamais de manière pure. Ils se donnent à voir à travers des figures-limites, c'est-à-dire des configurations dans lesquelles leurs principes opératoires apparaissent avec une netteté critique, souvent au prix de déséquilibres ou de paradoxes. [11]

Le régime communicationnel contemporain ne fait pas exception : son autonomie, fondée sur la circulation, la convertibilité et la persistance computationnelle, produit des formes de langage qui mettent à l'épreuve les catégories classiques de l'oralité, de l'écriture, de l'autorité et de la responsabilité. [27][28]

Les figures analysées ici ne constituent pas des anomalies marginales, mais des points de condensation où se rendent visibles les logiques profondes du régime. Elles permettent de penser ce que devient le langage lorsque l'origine cesse d'être structurante, sans pour autant disparaître complètement.

10.1 Oralité tertiaire et transformation de la parrhêsia

L'oralité tertiaire ne désigne ni un simple retour de la parole ni une oralité médiatisée au sens classique. Elle correspond à une forme de parole produite, diffusée et évaluée dans un milieu

algorithmique, où la voix — humaine ou synthétique — circule indépendamment de toute situation de coprésence. [13][27]

Dans ce régime, la parole retrouve certaines propriétés de l'oralité (immédiateté, fluidité, adressage apparent), tout en étant radicalement désancrée de l'imputabilité située qui caractérisait l'oralité primaire. La parole peut être rejouée, redécoupée, synthétisée, générée sans que le risque associé à l'acte de parler ne soit assumé par un sujet identifiable. [13][28]

Cette transformation affecte directement la *parrhêsia*. Le dire-vrai ne disparaît pas, mais il se désindexe du courage de celui qui parle. La parole peut produire des effets de dévoilement ou de transgression sans exposition réelle de l'énonciateur. Le risque est transféré du sujet vers le système de circulation lui-même, où il se dilue. [14]

L'oralité tertiaire rend ainsi possible une parole sans vulnérabilité, une franchise sans corps, et donc une reconfiguration profonde du rapport entre vérité, engagement et responsabilité.

10.2 Palimptextualité et circulation sans origine stable

La circulation algorithmique des énoncés produit une structure palimpsestique généralisée que l'on peut qualifier de *palimptextualité*. Chaque énoncé est potentiellement le résultat d'une superposition de traces, de reformulations, de citations implicites et de recombinaisons statistiques. [10][28][38]

Dans ce régime, il n'existe plus de texte originaire au sens fort. Les énoncés ne cessent d'être réécrits sans être explicitement effacés. L'origine n'est pas supprimée ; elle devient indifférente. Ce qui importe n'est pas d'où vient un énoncé, mais comment il s'inscrit dans des chaînes de circulation actives. [12][28]

La palimptextualité produit une instabilité structurelle du sens. Les énoncés conservent une apparence de cohérence locale, mais cette cohérence est sans mémoire unifiée. Elle est le résultat momentané d'une convergence de traces hétérogènes, dont aucune ne peut prétendre à une autorité fondatrice.

Cette figure-limite révèle un régime où la communication fonctionne sans point d'orientation symbolique stable. Le langage ne s'organise plus autour de textes, d'auteurs ou de traditions, mais autour de flux de reformulation continus.

Exemple minimal : une page Wikipédia ou une documentation collaborative, dont l'autorité résulte d'un processus de réécritures, de discussions et de versions plutôt que d'une origine stable. [17]*

10.3 Textoralité : milieu hybride de la circulation

La textoralité apparaît ici non plus seulement comme une forme de langage, mais comme un milieu. Elle ne relève ni de l'oral ni de l'écrit, mais d'un espace hybride où les énoncés sont conçus pour circuler, être traités et produire des effets indépendamment de leur support ou de leur modalité. [5][27][38]

* Ward Cunningham est l'inventeur du concept de Wiki (le *WikiWikiWeb*), source technique d'une rupture majeure : le passage d'un document écrit par un auteur unique à un système de gestion de connaissances par itérations.

Dans ce milieu, les distinctions traditionnelles — parler/écrire, lire/écouter, produire/interpréter — perdent leur pertinence analytique. Les énoncés sont immédiatement intégrés dans des chaînes d'opérations qui les transforment en ressources exploitables. La communication ne relie plus des sujets ; elle configure des environnements. [5][21]

Cette hybridation radicale rend le langage partiellement opaque à ses propres usagers. Les sujets continuent de lire, d'écouter et de parler, mais les effets de leurs énoncés sont déterminés par des structures de circulation qui excèdent leur compréhension et leur contrôle. [28]

La textoralité comme milieu constitue ainsi le cadre matériel et symbolique dans lequel s'installe durablement l'autonomie du régime communicationnel.

Exemple minimal : un échange avec un assistant conversationnel, où la stabilité de l'écrit (archivable, citables) se combine à la dynamique de l'oral (tours de parole, ajustements contextuels, relances). [26][30]

10.4 Textauthorité et responsabilité requalifiée

Face à la dilution de l'origine et à la fragmentation de l'autorité, la question de la responsabilité ne peut plus être posée dans les termes classiques de l'auteur souverain. De même qu'elle ne peut pas être abandonnée purement et simplement. C'est dans cette tension qu'émerge la figure de la *textauthorité*. [7][14][38]

La textauthorité ne renvoie pas à un retour de l'auteur, mais à une requalification de l'imputabilité. Elle désigne la possibilité d'assumer une responsabilité non sur l'origine d'un énoncé, mais sur ses conditions de circulation, de mise en forme et d'usage. Être responsable ne signifie plus être la source première, mais intervenir dans la configuration du milieu communicationnel. [28][34]

Cette responsabilité est nécessairement partielle, distribuée et asymétrique. Elle ne repose pas sur la maîtrise totale des effets, mais sur la reconnaissance d'un pouvoir d'influence sur les chaînes de circulation. La textauthorité n'est donc pas une position stable, mais une fonction critique, toujours à reconquérir. [21]

Exemple minimal : un texte généré puis « signé » et relayé comme s'il provenait d'un auteur identifiable, alors que l'origine effective est distribuée entre prompts, modèles, réécritures et dispositifs de publication.

Cette figure-limite ouvre une perspective normative a minima : non pas un retour à un régime du sens ou de la vérité fondé sur l'origine, mais une éthique de la circulation, attentive aux conditions de visibilité, de persistance et de crédit des énoncés. [28][29]

La Partie III a montré que la communication contemporaine ne peut plus être pensée comme transmission, mais comme circulation sans origine stable, qui constitue un régime autonome, doté de ses propres formes d'autorité, de mémoire et de pathologies.

Les figures-limites analysées dans ce chapitre n'en marquent pas l'échec, mais la zone critique : là où la circulation, livrée à elle-même, met en tension les catégories héritées de la

responsabilité, du sens et de la confiance — et où se dessinent, en creux, les conditions d'une requalification possible de l'agir communicationnel.

PARTIE IV — Le régime métabolique : soutenir sans assumer

Pendant longtemps, les dimensions matérielles de l'IA ont été pensées sous le registre du coût : coût de calcul, coût énergétique, coût d'infrastructure, coût humain : autant d'éléments relégués à l'arrière-plan des analyses, traités comme des variables d'optimisation ou des contraintes contingentes. Dans cette perspective, le métabolisme de l'IA apparaît comme un support neutre, ajustable à la marge, dont les effets pourraient être compensés par des gains de performance, des innovations techniques ou des améliorations d'efficacité. Le coût est alors conçu comme réversible, substituable, et fondamentalement secondaire par rapport aux régimes du sens et de la communication. [32][40]

Or cette manière de voir est devenue intenable. À mesure que les dispositifs d'IA se sont intensifiés, étendus et massifiés, ce qui était perçu comme un coût discret s'est transformé en accumulation durable : infrastructures lourdes, dissipation énergétique massive, extraction matérielle irréversible, charge organisationnelle et humaine croissante. Ces effets ne se contentent plus d'accompagner les systèmes ; ils s'inscrivent dans le temps, se déposent dans les territoires, les corps et les institutions, et contraignent les possibilités futures d'action. Ce déplacement impose un changement de registre : il ne s'agit plus de coûts à absorber, mais de dettes qui engagent le devenir même des régimes techniques et sociaux.

Parler de dette irréversible ne signifie pas seulement que les coûts augmentent. Cela signifie qu'ils ne peuvent pas être annulés sans produire davantage de désordre, et qu'ils introduisent une direction temporelle qui traverse l'ensemble de l'architectonique. Le métabolisme n'est pas un contexte. C'est **le régime de l'irréversibilité entropique**. Par conséquent, c'est dans le régime métabolique — et dans l'introduction de la dette entropique située comme opérateur central — que l'architectonique rencontre son point de contrainte décisif. C'est cette transformation — du coût invisible à la dette irréversible — qui prépare l'analyse du chapitre suivant. [8][11]

Dès lors, penser l'IA comme régime métabolique revient à décrire non un « contexte » des systèmes, mais la condition matérielle et organisationnelle continue qui rend possibles leurs effets : une économie de flux, d'entretien et de renouvellement dont la dynamique propre impose des seuils, des conflits de répartition, et une contrainte temporelle irréversible. [23]

11. L'IA comme machine entropique

11.1 L'IA comme système de flux

Penser l'IA comme système métabolique implique de rompre avec une représentation statique ou purement fonctionnelle des dispositifs techniques. Les systèmes contemporains dits d'« intelligence artificielle » ne sont ni des objets isolés, ni des algorithmes autonomes, ni même des infrastructures closes. Ils sont des configurations dynamiques de flux, continuellement traversées, alimentées et soutenues par des circulations hétérogènes.

Ces flux ne constituent pas un simple environnement externe du système : ils en sont la condition de fonctionnement permanente. Une IA ne « consomme » pas des ressources ponctuellement ; elle dépend d'un régime continu de circulation, d'entretien, de renouvellement et d'absorption. C'est en ce sens précis qu'il est légitime de parler de métabolisme. [3][19]

L'analyse suivante distingue quatre grandes catégories de flux, analytiquement séparables mais empiriquement intriquées :

- a) Flux de données : extraction, annotation, maintenance
- b) Flux énergétiques et computationnels : alimentation, calcul, dissipation
- c) Flux de travail humain : supervision, correction, modération, maintenance sociale
- d) Flux organisationnels et institutionnels : gouvernance interne, chaînes de décision, gestion de l'opacité et des arbitrages

Cette typologie n'implique aucune équivalence : selon les systèmes, les échelles et les usages, certains flux deviennent dominants et d'autres secondaires. Elle vise à cartographier des conditions de possibilité, non à stabiliser une métrique unique. [34]

a) Flux de données : extraction, annotation, maintenance

Le premier flux constitutif est celui des données. Contrairement à une représentation courante qui traite les données comme un stock disponible, le régime contemporain de l'IA repose sur des flux continus d'extraction, de transformation et de maintenance. [29][32]

Les données doivent être :

- collectées (souvent par extraction massive, *scraping*, captation indirecte),
- sélectionnées, nettoyées, filtrées,
- annotées, classifiées, corrigées,
- mises à jour, éliminées ou réentraînées.

Ce flux est à la fois technique et social. Il mobilise des infrastructures numériques, mais aussi des chaînes de travail humain largement invisibilisées (annotation, modération, correction, évaluation). [40] Il est également asymétrique : les données sont souvent extraites dans certains contextes sociaux ou géographiques, puis valorisées ailleurs.

Surtout, ce flux n'est pas neutre : chaque opération de sélection, de compression ou de filtrage produit une perte irréversible d'information, une transformation des contextes d'énonciation et une orientation implicite des usages futurs. Le flux de données n'est donc pas seulement un input : il est déjà une opération métabolique à coût informationnel. [8][10]

Cette perte de contexte constitue le versant métabolique de la convertibilité : ce qui rend l'énoncé circulaire augmente simultanément l'entropie* informationnelle qui fragilise l'orientation interprétative.

* [Grandeur physique](#) qui quantifie le désordre, l'incertitude ou la dispersion de l'énergie dans un système.

b) Flux énergétiques et computationnels

Le second flux fondamental est celui de l'énergie et du calcul. L'IA contemporaine repose sur une intensification sans précédent des opérations computationnelles, tant à l'entraînement qu'à l'inférence.

Ces opérations impliquent :

- une consommation massive d'électricité,
- des infrastructures de refroidissement,
- des dispositifs matériels spécialisés (GPU, TPU, ASIC),
- une dissipation thermique continue.

Ce flux est doublement contraint :

1. par des limites physiques (capacité de refroidissement, disponibilité énergétique),
2. par des contraintes territoriales et politiques (accès à l'énergie, conflits d'usage, régulation environnementale). [23]

Il est également structurellement irréversible : l'énergie dissipée sous forme de chaleur ne peut être récupérée sans coût supplémentaire. Chaque opération computationnelle contribue ainsi à une accumulation de dissipation qui excède toute logique de simple optimisation locale. [8][11]

C'est ce flux qui fonde, ultérieurement, la formalisation de la dette entropique située.

c) Flux de travail humain

Un troisième flux, longtemps invisibilisé, est celui du travail humain. Contrairement au récit d'une automatisation intégrale, les systèmes d'IA contemporains dépendent d'un travail continu, fragmenté et distribué.

Ce travail inclut notamment :

- l'annotation et la labellisation des données,
- la modération de contenus,
- le [RLHF](#) et autres dispositifs d'ajustement,
- la supervision, la correction d'erreurs, la gestion des incidents,
- la maintenance cognitive et organisationnelle.

Ce flux est caractérisé par :

- une externalisation sociale (sous-traitance, précarisation),
- une invisibilisation institutionnelle,
- une non-reconnaissance métabolique (ce travail est rarement comptabilisé comme coût structurel). [32]

Or ce travail n'est pas périphérique : il est une condition de stabilité minimale des systèmes. Lorsqu'il sature ou se dégrade, les performances chutent, outre la possibilité d'endosser les productions. [34][40]

d) Flux organisationnels et institutionnels

Enfin, les systèmes d'IA sont traversés par des flux organisationnels : décisions, normes, procédures, arbitrages, responsabilités.

Ces flux concernent :

- la coordination entre équipes techniques, juridiques, commerciales,
- la gestion des risques,
- l'adaptation aux cadres réglementaires,
- la prise en charge des incidents et des contestations.

Ils sont souvent sous-estimés parce qu'ils ne sont ni matériels ni immédiatement mesurables. Ils constituent pourtant un métabolisme institutionnel à part entière. Lorsque ces flux se rigidifient, se fragmentent ou se saturent, la gouvernabilité du système s'effondre, indépendamment de sa performance technique. [34][37]

Conclusion intermédiaire

Ces quatre types de flux montrent que l'IA n'est pas un objet, mais un processus soutenu, et que ce soutien a un coût structurel. Avant même de parler d'entropie, de seuils ou de désassomption, il est donc nécessaire de reconnaître que l'intelligence artefactuelle repose sur une circulation continue de ressources hétérogènes, dont aucune n'est indéfiniment extensible. [8][11]

En fait, le concept de dette entropique doit être construit comme une synthèse contrôlée par la description fine des flux, et non comme un principe général qui réinterprète tout depuis un point de vue extérieur.

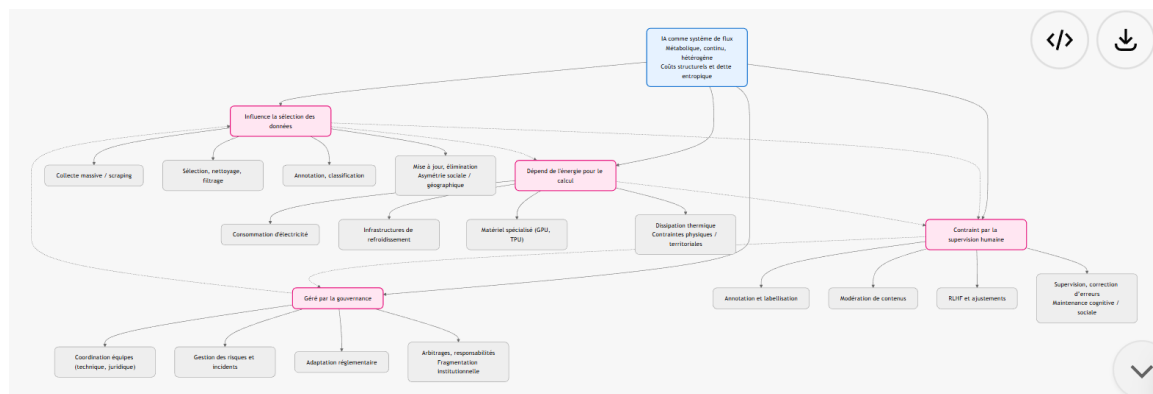


Figure 3. L'IA comme un système métabolique dynamique de flux

L'IA comme système de flux est au plan supérieur ; les quatre flux sont les branches principales ; les composants sont les sous-nœuds grisés, et les flèches indiquent les directions de dépendance (même si, en réalité, tout est interconnecté ; le réseau n'est simplifié qu'à des fins de clarté).

11.2 L'extractivisme computationnel différé

Le terme d'extractivisme est employé ici dans un sens opératoire : il désigne un régime de captation asymétrique de conditions de possibilité ; aucune analogie morale avec l'extraction minière.

L'analyse des flux qui précède permet de dépasser une compréhension purement technique ou fonctionnelle du métabolisme de l'IA. Elle rend possible une qualification plus précise du type de relation que ces systèmes entretiennent avec leurs conditions matérielles, sociales et temporelles de possibilité. Cette relation peut être désignée comme une forme spécifique d'extractivisme, qu'il convient toutefois de distinguer des modèles classiques de l'extraction.

L'IA contemporaine ne repose pas sur une extraction immédiate, localisée et visible de ressources, mais sur un extractivisme computationnel différé : dans l'espace, dans le temps et dans la structure même de l'imputation des coûts. Cette différenciation est décisive pour comprendre pourquoi les dispositifs peuvent apparaître légers, immatériels ou « propres » au point d'usage, tout en reposant sur des chaînes métaboliques lourdes et asymétriques. [23][32]

a) De l'extractivisme classique à l'extractivisme computationnel

Dans sa forme classique, l'extractivisme se caractérise par plusieurs traits bien identifiés : la localisation géographique de l'extraction, la visibilité des infrastructures, l'imputation relativement directe des coûts environnementaux et sociaux, et la temporalité relativement courte entre l'acte d'extraction et ses effets.

L'extractivisme computationnel, tel qu'il s'est constitué avec les infrastructures numériques globales, rompt avec cette configuration. Il ne se fonde pas sur l'extraction d'une ressource unique, mais sur la captation simultanée de flux hétérogènes : données, énergie, travail humain, attention, infrastructures territoriales. Cette captation ne se présente pas comme une extraction au sens intuitif du terme, car elle est médiatisée par des dispositifs techniques, juridiques et organisationnels qui en fragmentent la perception. [27][29]

Ainsi, l'IA ne « prélève » pas directement une matière identifiable ; elle convertit des environnements sociaux, informationnels et énergétiques en intrants computationnels, souvent sans rupture perceptible pour les acteurs situés en aval de la chaîne.

b) Différenciation spatiale : extraire ici, valoriser ailleurs

La première dimension du différé est spatiale. Les flux nécessaires au fonctionnement des systèmes d'IA sont rarement co-localisés avec les lieux où leurs bénéfices économiques, symboliques ou décisionnels sont produits.

Les données sont collectées à l'échelle globale, souvent dans des contextes où les capacités de contestation, de négociation ou de régulation sont faibles. L'énergie est mobilisée dans des territoires spécifiques, soumis à des contraintes hydriques, climatiques ou politiques propres. Le travail humain est externalisé vers des populations précarisées, invisibilisées dans les récits dominants de l'innovation. [40]

En revanche, la valeur produite — financière, stratégique, cognitive — est concentrée auprès d'un nombre restreint d'acteurs et de zones géopolitiques. Cette dissociation structurelle

entre lieux d'extraction et lieux de valorisation constitue une asymétrie métabolique centrale du régime contemporain de l'IA. [23][32]

c) Différenciation temporelle : repousser les coûts

La seconde dimension du différé est temporelle. Une part significative des coûts métaboliques de l'IA n'est ni immédiate ni intégralement actualisée au moment de l'usage.

Les effets environnementaux, la dégradation des infrastructures, l'épuisement des ressources, la saturation organisationnelle ou l'usure du travail humain se manifestent souvent après coup, selon des temporalités longues, discontinues ou non linéaires. Cette temporalité différée permet aux systèmes de fonctionner dans un régime de relative impunité métabolique, tant que les seuils critiques ne sont pas franchis.

Cette dissociation temporelle rend également la responsabilité difficile à instituer. Les décisions qui engagent des coûts lourds sont prises dans un présent organisationnel court, tandis que leurs conséquences se déploient dans un futur incertain, souvent hors du champ de visibilité des acteurs décisionnels initiaux. [8][34]

d) Différenciation structurelle : invisibiliser la chaîne

Enfin, l'extractivisme computationnel est différé au sens structurel. Les dispositifs techniques fragmentent la chaîne métabolique en une multitude d'opérations spécialisées, contractualisées, automatisées ou externalisées.

Cette fragmentation produit un effet d'invisibilisation systémique [29][34] :

- les coûts sont répartis entre de multiples acteurs,
- les responsabilités sont diluées,
- les seuils sont difficiles à percevoir localement.

Il devient alors possible de soutenir, de bonne foi, que chaque composant du système est optimisé, responsable ou conforme, tout en participant collectivement à une dynamique extractive globale. L'extractivisme computationnel n'est pas une dérive morale ou un excès ponctuel : il est structurellement inscrit dans l'architecture des systèmes à grande échelle.

e) De l'extractivisme au métabolisme de capture

Ce régime d'extractivisme différé permet l'émergence de formes de capture métabolique. Certains acteurs acquièrent la capacité de capter une part disproportionnée des flux énergétiques, informationnels et humains, non seulement pour produire des services, mais pour structurer durablement les conditions d'accès à ces flux.

Cette capture renforce des asymétries de pouvoir qui ne sont plus seulement économiques ou informationnelles, mais proprement métaboliques. [23][24] Elle confère aux infrastructures d'IA un rôle stratégique dans la redistribution des capacités d'action, de décision et de contrainte à l'échelle globale.

Conclusion intermédiaire

L'extractivisme computationnel différé permet de comprendre pourquoi le régime métabolique de l'IA ne peut être réduit à une question de coûts techniques ou d'optimisation énergétique. Il révèle une logique de captation asymétrique, spatialement et temporellement

décalée, qui rend les seuils de soutenabilité difficiles à identifier et les conflits de répartition difficiles à politiser*.

Cette analyse prépare l'introduction de la notion de dette entropique située, qui permettra, au chapitre suivant, d'unifier ces différentes formes de dégradation sous un opérateur commun, sans effacer leur hétérogénéité ni leur conflictualité.

11.3 Accumulation, invisibilisation et externalisation

L'un des traits structurants du régime métabolique de l'intelligence artificielle tient au fait que ses coûts ne se manifestent ni immédiatement, ni là où les effets d'intelligence sont produits et valorisés. Contrairement à d'autres formes de production technique, les dispositifs contemporains d'IA reposent sur des mécanismes d'accumulation différée, d'invisibilisation systématique et d'externalisation spatiale, temporelle et sociale des charges qu'ils génèrent. Ces trois mécanismes ne sont pas contingents, ils constituent le mode de fonctionnement ordinaire du régime métabolique à l'ère des modèles à grande échelle. [23][29]

Accumulation : une croissance non linéaire des passifs

L'accumulation métabolique ne se réduit pas à une addition progressive de coûts. Elle procède par effets cumulatifs et boucles de rétroaction. Chaque extension fonctionnelle — augmentation du nombre de paramètres, allongement des contextes, multiplication des agents, généralisation des usages — accroît non seulement les flux immédiats (énergie, calcul, données), mais aussi les coûts futurs de maintenance, de refroidissement, de supervision humaine et de renouvellement des infrastructures. À mesure que les systèmes gagnent en complexité, la part de ressources consacrée à la simple conservation de leurs performances augmente plus vite que celle consacrée à leur amélioration.

Cette dynamique produit une forme de croissance entropique : plus le système est performant localement, plus il devient coûteux à maintenir globalement. L'accumulation n'est donc pas seulement quantitative, elle est structurelle. Elle modifie les conditions mêmes de soutenabilité des régimes du sens et de la communication, en introduisant des contraintes qui ne peuvent être levées par de simples optimisations marginales. [8][11]

Invisibilisation : rendre les coûts non pertinents pour la décision

Pour que cette accumulation puisse se poursuivre, elle doit être rendue invisible au point de décision. L'invisibilisation n'est pas un défaut d'information, mais une condition de possibilité du régime. Les interfaces d'usage, les modèles économiques et les métriques dominantes (latence, précision, satisfaction utilisateur, coût unitaire apparent) sont construits de manière à découpler l'expérience du sens et de la communication de leurs conditions métaboliques réelles. [27][29]

Ainsi, les coûts énergétiques sont agrégés, mutualisés ou dilués dans des contrats globaux ; le travail humain est fragmenté, sous-traité ou rendu statistiquement négligeable ; l'usure matérielle est masquée par des cycles rapides de renouvellement ; la pollution

* L'utilisation du terme « politiser / politisation » dans ce texte signifie que des tensions ou des coûts ne doivent pas être traités comme de simples contraintes techniques, économiques ou naturelles, mais comme des arbitrages collectifs explicitement discutables. Politiser, ici, désigne l'acte de rendre visibles des conflits de répartition pour les réintroduire dans l'espace de la décision publique, de la contestation et de la responsabilité institutionnelle.

informationnelle est absorbée par des procédures de nettoyage dont le coût n'est jamais imputé aux usages qui la produisent. Cette invisibilisation transforme des contraintes matérielles effectives en arrière-plans supposés neutres, ce qui empêche leur prise en compte dans les arbitrages de conception et de gouvernance.

Externalisation : déplacer plutôt que résoudre

L'invisibilisation rend possible l'externalisation. Les coûts du régime métabolique sont déplacés hors du champ de visibilité et de responsabilité des acteurs qui bénéficient des effets d'intelligence. Cette externalisation opère selon trois axes principaux.

Sur le plan spatial, les charges sont transférées vers des territoires périphériques : extraction minière dans des régions faiblement protégées, implantation de centres de données dans des zones à énergie bon marché ou à régulation environnementale plus souple, sous-traitance du travail humain vers des marchés précaires. Sur le plan temporel, les coûts sont reportés vers l'avenir : obsolescence différée, pollution cumulative, dépendances techniques dont les effets ne se manifestent qu'à moyen ou long terme. Sur le plan social, enfin, l'externalisation repose sur une dissociation entre ceux qui décident, ceux qui bénéficient et ceux qui supportent effectivement les charges — qu'il s'agisse de travailleurs invisibilisés, de communautés locales ou de générations futures.

Cette triple externalisation n'élimine pas les coûts, elle les redistribue de manière asymétrique. Elle permet au régime métabolique de soutenir provisoirement des niveaux élevés de production de sens et de circulation communicationnelle, tout en creusant un passif qui finit par se manifester sous forme de seuils critiques, de conflits de répartition ou de désassomption. [23][32]

Conséquence architectonique

Accumulation, invisibilisation et externalisation ne sont pas des dérives accidentelles ; elles constituent le cœur du fonctionnement métabolique contemporain. Elles expliquent pourquoi des systèmes peuvent paraître parfaitement viables à court terme tout en devenant structurellement ingouvernables à moyen terme. Tant que ces mécanismes restent hors champ, le métabolisme apparaît comme un simple contexte technique ajustable. Lorsqu'ils deviennent visibles, le régime métabolique impose clairement une contrainte architectonique majeure : il introduit une irréversibilité qui limite la durée pendant laquelle les effets d'intelligence peuvent être reconnus, endossés et assumés. [8][11]

12. Dette entropique située et seuils critiques

12.1 La dette entropique située : définition et portée

Les analyses précédentes ont montré que le métabolisme de l'IA contemporaine repose sur des flux continus de données, d'énergie, de calcul et de travail humain, structurés par une logique d'extractivisme computationnel différé. Ces analyses permettent d'identifier les composantes, les asymétries et les tensions du régime métabolique. Elles ne suffisent toutefois pas à en saisir une propriété décisive : son irréversibilité structurelle. [23][32]

C'est pour rendre intelligible cette irréversibilité qu'il est nécessaire d'introduire un opérateur conceptuel unificateur, sans pour autant réduire la pluralité des phénomènes observés. Cet opérateur est celui de dette entropique située. [8][11]

Définition canonique

On appellera dette entropique située l'accumulation irréversible de désordre matériel, énergétique, informationnel, organisationnel et social produite par le fonctionnement des systèmes d'IA, accumulation qui ne peut être ni intégralement compensée, ni simplement externalisée, et qui se dépose dans des territoires, des infrastructures, des institutions et des corps situés. [8][11][32]

Cette dette est dite :

- entropique, parce qu'elle renvoie à une augmentation globale du désordre et de la dissipation, au sens large — thermodynamique, informationnel et organisationnel ;
- située, parce qu'elle n'est jamais abstraite ni homogène : elle s'inscrit toujours dans des contextes géographiques, sociaux et institutionnels précis, affectant différemment les acteurs et les milieux. [8][23]

Pourquoi parler de dette

Le terme de dette ne doit pas être entendu au sens strictement comptable ou financier. Il désigne ici une obligation différée, contractée sans être immédiatement reconnue, et dont le paiement est reporté dans le temps, déplacé dans l'espace ou transféré vers d'autres acteurs.

Contrairement à la dette technique ou à la dette écologique prises isolément, la dette entropique située ne présuppose pas la possibilité d'un remboursement intégral. Elle signale au contraire l'existence de coûts qui, une fois produits, modifient durablement les conditions de possibilité des régimes futurs. Elle rend visible le fait que certaines opérations — computationnelles, organisationnelles ou communicationnelles — ferment des options plutôt qu'elles n'en ouvrent.

Fondement conceptuel : irréversibilité et accumulation

L'intérêt décisif de la notion de dette entropique réside dans sa capacité à penser l'irréversibilité. Toute opération computationnelle dissipe de l'énergie ; toute compression informationnelle produit une perte ; toute complexification organisationnelle augmente la charge de maintenance ; toute externalisation sociale engendre une usure située.

Ces effets peuvent parfois être ralentis, compensés localement ou déplacés. Ils ne peuvent pas être annulés globalement. La dette entropique ne disparaît pas : elle s'accumule, et cette accumulation impose une direction temporelle aux systèmes techniques et institutionnels.

C'est en ce sens que la dette entropique introduit une flèche du temps au cœur de l'architecture de l'intelligence artificielle. Elle empêche de penser la soutenabilité comme un simple problème d'optimisation réversible.

Dette entropique et pluralité des dimensions

La dette entropique située ne se confond avec aucune de ses dimensions empiriques. Elle les articule sans les absorber. On peut en distinguer, analytiquement, cinq composantes :

- une dimension thermodynamique (dissipation énergétique, chaleur, pertes irréversibles) ;
- une dimension informationnelle (bruit, dégradation du signal, perte de contexte) ;

- une dimension matérielle (usure des infrastructures, obsolescence, déchets) ;
- une dimension organisationnelle (complexité croissante, saturation des chaînes de décision) ;
- une dimension sociale (épuiement du travail humain, fragmentation des responsabilités).

Parler de dette entropique située ne revient donc pas à réduire ces phénomènes à une métaphore physique unique. Il s'agit d'un opérateur limite, qui permet de penser leur convergence sans effacer leurs différences.

Pourquoi cet opérateur est nécessaire à l'architectonique

Sans cet opérateur, le régime métabolique reste descriptif : il accumule des constats de coûts, de seuils et d'externalités, sans pouvoir en saisir la dynamique globale. Avec, il devient possible de comprendre pourquoi certaines configurations cessent brutalement d'être soutenables, pourquoi certains arbitrages deviennent politiquement inacceptables, et pourquoi certaines formes d'intelligence disparaissent alors même que les systèmes continuent de fonctionner. [8][34]

La dette entropique située permet ainsi de relier :

- les flux analysés au chapitre 11.1,
- l'extractivisme computationnel différé analysé au chapitre 11.2,
- et les seuils critiques qui seront étudiés plus loin.

Les sections suivantes examineront comment cette dette s'accumule selon des mécanismes spécifiques, comment elle franchit des seuils entropiques, et comment ces seuils redéfinissent les conditions mêmes de la responsabilité, de la gouvernabilité et, en dernière instance, de l'intelligence artefactuelle.

12.2 Typologie des seuils entropiques

L'introduction de la dette entropique située permet de dépasser une approche graduelle ou purement quantitative de la soutenabilité des systèmes d'IA. Elle rend possible l'identification de seuils entropiques, c'est-à-dire de points à partir desquels l'accumulation de désordre ne peut plus être absorbée, déplacée ou compensée sans transformation qualitative du système. [11][34]

Un seuil entropique ne désigne pas une limite maximale abstraite, mais un point de rupture opératoire : le moment où les mécanismes ordinaires de régulation cessent de fonctionner, et où les arbitrages deviennent explicitement conflictuels, coûteux ou politiquement inassumables.

Ces seuils ne sont ni synchrones ni homogènes. Ils se manifestent selon des dimensions distinctes, qui peuvent se renforcer mutuellement. On peut en distinguer quatre types principaux.

a) Seuil entropique thermodynamique

Le seuil thermodynamique est atteint lorsque les contraintes physiques liées à la dissipation énergétique deviennent dominantes. Il correspond à des situations où l'augmentation des capacités de calcul ou de stockage entraîne des coûts énergétiques et thermiques qui ne peuvent plus être compensés par des gains d'efficacité marginaux. [8][23]

Ce seuil se manifeste notamment par :

- l'impossibilité de refroidir certaines infrastructures sans investissements disproportionnés,
- la saturation des réseaux énergétiques locaux,
- des conflits d'usage entre infrastructures numériques et besoins sociaux ou écologiques.

À ce stade, la question n'est plus celle de l'optimisation énergétique, mais celle de l'acceptabilité physique et territoriale du système. Le métabolisme computationnel cesse d'être extensible, et toute poursuite de l'expansion implique des arbitrages explicites entre usages concurrents de l'énergie.

b) Seuil entropique informationnel

Le seuil informationnel concerne la capacité des systèmes et des collectifs à distinguer le signal du bruit. Il est atteint lorsque l'accumulation de productions discursives, de transformations, de résumés et de conversions rend impossible une stabilisation interprétative minimale. [28][30]

Ce seuil se manifeste par :

- une saturation des espaces informationnels,
- une dégradation continue de la qualité contextuelle des énoncés,
- une asymétrie croissante entre la facilité de production et la difficulté de correction ou de vérification.

À partir de ce niveau, l'augmentation de la production informationnelle n'accroît plus le sens disponible, mais génère une entropie informationnelle qui affaiblit la réponsabilité et la crédibilité. La communication continue de fonctionner, mais elle cesse d'être orientable.

c) Seuil entropique matériel

L'endettement entropique devient critique lorsque la soutenabilité infrastructurelle elle-même cesse de tenir : obsolescence accélérée, dépendance minérale, production de déchets non recyclables, raréfaction des composants, saturation des chaînes logistiques. Ce seuil marque le moment où le système ne peut plus être maintenu sans extraction supplémentaire ou renouvellement destructeur.

d) Seuil entropique organisationnel

Le seuil organisationnel est atteint lorsque la complexité des systèmes dépasse les capacités humaines et institutionnelles de supervision, de maintenance et de décision. [34]

Il se manifeste par :

- une multiplication des dépendances techniques et contractuelles,
- une fragmentation des responsabilités,
- une surcharge cognitive et décisionnelle des équipes humaines,
- un temps de maintenance supérieur au temps de production.

Si l'on franchit ce seuil, le système ne devient pas nécessairement instable sur le plan technique, mais ingouvernable. Les erreurs ne peuvent plus être anticipées ni corrigées de manière coordonnée, et la responsabilité se dilue dans la structure même du dispositif.

e) Seuil entropique social

Le seuil social concerne l'acceptabilité collective et la soutenabilité humaine du système. Il est atteint lorsque les coûts sociaux — épuisement du travail humain, inégalités de répartition, perte de confiance — deviennent politiquement ou moralement insupportables. [29][32]

Ce seuil se manifeste par :

- des résistances locales ou sectorielles,
- une contestation accrue des usages,
- une défiance institutionnelle,
- une fatigue attentionnelle et morale généralisée.

À ce point, la dette entropique n'est plus seulement un problème de soutenabilité matérielle, mais un problème de légitimité. Les systèmes peuvent continuer de fonctionner, ils cessent pour autant d'être reconnus comme acceptables ou justifiables.

Conclusion intermédiaire

Ces différents seuils ne sont pas indépendants. Le franchissement d'un seuil dans une dimension tend à accélérer l'atteinte des autres. Un seuil thermodynamique peut renforcer un seuil social; un seuil informationnel peut rendre les arbitrages organisationnels impossibles.

La typologie proposée n'a pas pour fonction de hiérarchiser ces seuils, mais de rendre visible la pluralité des points à partir desquels la dette entropique située devient ingérable, et à partir desquels l'intelligence artefactuelle cesse d'être gouvernable, même si les systèmes continuent de produire des effets. [11][34]

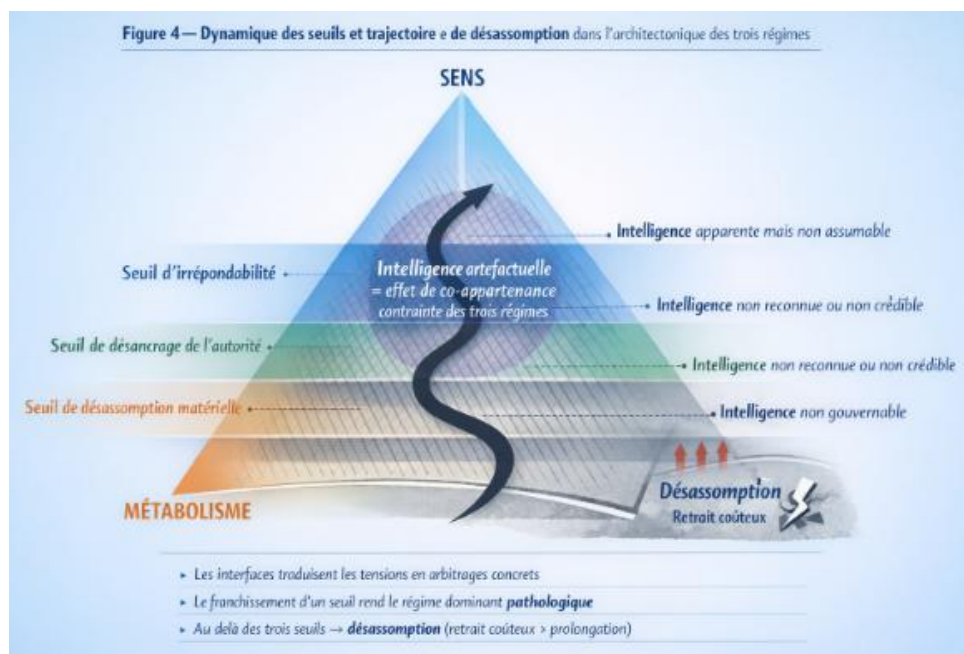


Figure 4. Dynamique des seuils et trajectoire de désassomption

La Figure 4 propose une vue d'ensemble architectonique. Elle ne constitue pas un modèle explicatif, mais un plan de lecture des opérateurs et des seuils discutés dans les encadrés suivants : méthodologique, conceptuel et analytique.

Encadré 1 (méthodologique) — L'entropie comme opérateur limite (et non totalisant)

L'usage de la notion d'entropie dans cet ouvrage appelle une clarification méthodologique. Il ne s'agit ni d'une métaphore décorative, ni d'un principe explicatif universel destiné à absorber l'ensemble des phénomènes analysés.

L'entropie est mobilisée ici comme opérateur limite.

Dire qu'un opérateur est limite signifie qu'il intervient là où les descriptions sectorielles cessent d'être suffisantes, sans prétendre les remplacer. [8][11] La dette entropique située ne se substitue ni aux analyses techniques, ni aux analyses sociales, ni aux analyses organisationnelles. Elle marque le point où leurs effets convergent vers une irréversibilité commune.

En ce sens, l'entropie n'est pas :

- un fondement ontologique,
- un principe de causalité unique,
- un langage totalisant.

Elle est un indicateur de saturation, qui permet de penser :

- la fin des compensations silencieuses,
- l'épuisement des marges d'optimisation,
- la transformation des coûts en conflits.

Refuser de faire de l'entropie un opérateur totalisant est essentiel pour éviter deux écueils symétriques. Le premier serait un réductionnisme physicaliste, qui naturaliserait des phénomènes sociaux et politiques sous couvert de lois thermodynamiques. Le second serait un fatalisme technologique, qui transformerait l'irréversibilité en inéluctabilité, et la contrainte en destin.

Dans cette architectonique, l'entropie ne dit jamais ce qu'il faut faire, mais ce qu'il n'est plus possible de faire comme avant. Elle ouvre un espace d'arbitrage, au lieu de le refermer. Elle rend visibles les limites sans prescrire leur dépassement. [11]

C'est précisément en restant un opérateur limite — et non un principe englobant — que la dette entropique située peut jouer son rôle : non pas totaliser l'intelligence artefactuelle, mais en exposer la fragilité temporelle, la conflictualité et la finitude.

L'entropie, comprise comme opérateur limite, n'épuise toutefois pas la question du devenir des régimes d'intelligence artefactuelle. Le **second encadré** en examine la conséquence temporelle.

Encadré 2 (conceptuel) — Flèche temporelle et asymétrie du devenir

L'introduction de la dette entropique située implique une conséquence temporelle qui mérite d'être distinguée de l'entropie elle-même : l'existence d'une flèche temporelle propre au régime métabolique.

Par flèche temporelle, on désignera ici l'asymétrie structurelle entre l'expansion et le retrait des systèmes d'intelligence artefactuelle. Cette asymétrie ne tient pas à la simple succession des états, mais au fait que les effets produits par les opérations computationnelles, organisationnelles et sociales — dissipation énergétique, usure matérielle, perte informationnelle, complexification institutionnelle — ne peuvent être annulés sans coût supplémentaire. [8][11]

La flèche temporelle n'est donc pas un principe explicatif autonome. Elle est l'effet architectural de l'irréversibilité entropique : ce qui a été dissipé, extrait, saturé ou rigidifié ne redevient disponible qu'au prix d'une dépense additionnelle, souvent supérieure à celle qui a permis l'expansion initiale.

Cette asymétrie transforme des choix localement réversibles en engagements durablement contraignants. Elle réduit l'espace des options futures, verrouille certaines trajectoires techniques et institutionnelles, et rend structurellement plus coûteux le désengagement que la poursuite du régime existant. [11][23]

En ce sens, la flèche temporelle n'est ni un destin, ni une loi naturelle. Elle est un effet de structure, produit par l'organisation même des flux métaboliques, et elle constitue la condition temporelle à partir de laquelle les notions de seuil, d'hystérésis, de point de non-retour et, ultimement, de désassomption deviennent intelligibles.

Cette asymétrie temporelle devient pleinement intelligible lorsqu'on en observe les effets sur des projets infrastructurels concrets. Le **troisième encadré** en propose une analyse appliquée.

Encadré 3 (analytique) — Désassomption: exemple d'un méga-déploiement infrastructurel

Le [projet Stargate \(OpenAI\)](#), annoncé en 2025, illustre un mécanisme typique de désassomption lié au discours contemporain de *super-buildout* (ou méga-déploiement en bon français :-). Présenté comme une infrastructure stratégique de long terme, il vise des investissements cumulés pouvant atteindre 500 milliards de dollars et une montée en puissance de l'ordre de plusieurs gigawatts de capacité de calcul, dans un contexte de compétition technologique et de promesses autour des systèmes d'IA avancés.

Cette phase d'expansion initiale est rendue possible par une forte mobilisation de capitaux, une rhétorique d'urgence (« *construire maintenant ou perdre la course* ») et une relative sous-estimation des contraintes métaboliques et institutionnelles. Or, les projections publiques sur la consommation électrique des centres de données aux États-Unis indiquent une croissance rapide, susceptible d'atteindre plusieurs centaines de térawattheures annuels à l'horizon de la fin de la décennie, plaçant ces infrastructures en tension directe avec les capacités des réseaux, les régulations environnementales et l'acceptabilité locale.

Dans un tel contexte, l'hypothèse d'un retrait partiel ou d'un redimensionnement significatif des projets — par exemple l'abandon de certains sites après engagement des procédures foncières, contractuelles et énergétiques — révèle un effet d'hystérésis caractéristique : les coûts de sortie deviennent supérieurs aux coûts de prolongation. Les investissements déjà consentis (études, équipements, contrats d'approvisionnement, raccordements), les risques d'actifs échoués (*stranded assets*), les pénalités financières et la dépendance des architectures techniques et organisationnelles à la capacité projetée rendent le désengagement économiquement et politiquement plus coûteux que la poursuite du régime, même sous conflit. Le résultat n'est pas une continuation « optimale », mais une prolongation contrainte : maintien des trajectoires malgré les oppositions réglementaires, environnementales ou territoriales, et malgré la multiplication des signaux de saturation. Ce type de configuration correspond précisément à une situation de désassomption, où l'intelligence artefactuelle cesse d'être pleinement gouvernable non parce qu'elle serait techniquement incontrôlable, mais parce que les conditions matérielles, financières et institutionnelles de son retrait ont été rendues excessivement coûteuses par l'expansion initiale. [8][23][24][34]

12.3 Irréversibilité, hystérésis et points de non-retour

L'identification de seuils entropiques ne suffit pas, à elle seule, à rendre compte de la dynamique temporelle propre au régime métabolique de l'IA. Il est en effet nécessaire de distinguer entre des dépassements ponctuels, susceptibles d'être corrigés, et des transformations durables qui modifient irréversiblement les conditions de fonctionnement des systèmes. C'est à ce niveau qu'interviennent les notions d'irréversibilité, d'hystérésis et de point de non-retour. [8][11][23][34]

Irréversibilité : une propriété structurelle du métabolisme

L'irréversibilité constitue la propriété centrale de la dette entropique située. Elle signifie que certaines opérations produisent des effets qui ne peuvent être annulés par un simple retour à l'état antérieur. Une énergie dissipée ne peut être récupérée intégralement ; une ressource extraite ne peut être réinsérée dans son milieu sans transformation ; une organisation complexifiée ne peut être simplifiée sans coût supplémentaire.

Dans le cas de l'IA contemporaine, cette irréversibilité se manifeste à plusieurs niveaux :

- thermodynamique, par la dissipation énergétique irréversible liée au calcul ;
- matériel, par l'usure et l'obsolescence des infrastructures ;
- informationnel, par la perte de contexte et la dégradation cumulative des traces ;
- organisationnel, par l'accumulation de dépendances et de routines difficiles à défaire ;
- social, par l'épuisement du travail humain, la fragmentation des chaînes de responsabilité et l'irréversibilité des pertes d'assomption collective.

L'irréversibilité n'implique pas l'impossibilité de toute action corrective. Elle implique que toute correction a un coût entropique supplémentaire, qui s'ajoute à la dette déjà produite.

Hystérésis : pourquoi le retour en arrière coûte plus cher

La notion d'hystérésis permet de préciser cette dynamique. Un système présente un phénomène d'hystérésis lorsque le chemin de retour vers un état antérieur diffère du chemin qui a conduit à son dépassement. Autrement dit, revenir en arrière n'est pas simplement inverser le mouvement précédent.

Appliquée au régime métabolique de l'IA, l'hystérésis signifie que :

- réduire la taille d'un modèle ne supprime pas les coûts déjà engagés pour son entraînement ;
- fermer une infrastructure ne restaure pas les ressources énergétiques ou hydriques consommées ;
- simplifier une organisation ne fait pas disparaître la charge cognitive et institutionnelle accumulée.

Conséquence majeure de cette asymétrie temporelle : plus un seuil est dépassé longtemps, plus son franchissement devient coûteux à corriger. L'hystérésis transforme ainsi des choix provisoires en engagements durables, souvent sans décision explicite.

Points de non-retour : quand la correction devient impossible

Dans certaines configurations, l'accumulation de dette entropique ne conduit pas seulement à une augmentation des coûts de correction, mais à leur impossibilité pratique. On parlera alors de point de non-retour.

Un point de non-retour est atteint lorsque :

- la réparation nécessiterait plus de ressources que celles que le système est capable de mobiliser ;
- la restauration des conditions antérieures entrerait en conflit direct avec d'autres exigences vitales (énergétiques, sociales, politiques) ;
- la légitimité sociale ou institutionnelle du système est définitivement compromise.

Dans ces situations, le problème n'est plus celui de l'optimisation ou de la réforme, mais celui de la désassomption : reconnaître qu'un certain état du système ne peut plus être soutenu ni assumé, et que sa poursuite aggraverait la dette au lieu de la résorber.

Temporalité métabolique et illusion de réversibilité

L'analyse des points de non-retour permet de comprendre pourquoi certaines stratégies de gouvernance reposent sur une illusion dangereuse : celle de la réversibilité indéfinie. L'idée selon laquelle il serait toujours possible de « corriger plus tard » — par une innovation technique future, une régulation ultérieure ou une optimisation supplémentaire — ignore la dynamique entropique propre au métabolisme.

Cette illusion est renforcée par la dissociation temporelle caractéristique de l'extractivisme computationnel différé : les bénéfices sont immédiats, tandis que les coûts sont différés, déplacés ou invisibilisés. La dette entropique située révèle que cette dissociation n'annule pas les coûts, mais les déplace dans le futur, jusqu'à ce qu'ils deviennent inévitables.

Transition vers la désassomption

Reconnaître l'irréversibilité, l'hystérésis et les points de non-retour conduit à reformuler profondément la question de la responsabilité. Il ne s'agit plus seulement de répondre des effets produits, mais de reconnaître les situations dans lesquelles répondre devient matériellement ou institutionnellement impossible.

C'est cette bascule — du dépassement de seuils à l'impossibilité d'assumer — qui sera analysée dans le chapitre suivant. Celui-ci montrera comment la dette entropique située transforme la responsabilité en conflit de répartition, et comment la gouvernabilité de l'intelligence artefactuelle dépend désormais de la capacité à nommer, instituer et arbitrer ces situations de désassomption.

13. Conflits de répartition et désassomption

13.1 Qui paie la dette ? Répartition et invisibilisation de la dette entropique

L'introduction de la dette entropique située transforme radicalement la question du métabolisme de l'IA. Elle déplace l'analyse d'un problème de coûts abstraits vers un problème de répartition concrète. Une dette qui ne peut être annulée pose inévitablement la question

de son paiement : non pas en théorie, mais dans des lieux, des institutions et des corps déterminés. [23][32]

Contrairement à une dette financière classique, la dette entropique n'est jamais acquittée par un sujet unique ni à un moment identifiable. Elle est diffuse, différée et distribuée, ce qui constitue précisément la condition de sa reproduction continue.

La loi implicite de conservation de la dette

Toute dette entropique produite par un système d'IA est nécessairement absorbée quelque part. L'énergie dissipée est payée par des infrastructures énergétiques ; les ressources extraites par des territoires ; la charge organisationnelle par des équipes humaines ; la dégradation informationnelle par des collectifs confrontés au bruit et à la perte de repères.

Ce qui varie n'est pas l'existence de la dette, mais sa localisation et son mode de prise en charge. La question « *qui paie ?* » ne renvoie donc pas à une recherche de responsables moraux abstraits, mais à une cartographie des lieux où l'entropie se dépose effectivement. [8][23]

Externalisation spatiale, temporelle et sociale

La spécificité du régime métabolique de l'IA contemporaine réside dans sa capacité à externaliser la dette selon trois axes principaux.

- Externalisation spatiale : les coûts énergétiques, miniers et environnementaux sont concentrés dans des territoires éloignés des centres de décision et de bénéfice. Les infrastructures de calcul, les sites d'extraction et les zones de traitement des déchets matérialisent cette dissociation géographique.
- Externalisation temporelle : la dette est reportée sur le futur. Les bénéfices immédiats des systèmes d'IA s'accompagnent d'une accumulation de passifs qui affecteront des générations futures, sans mécanisme clair de transmission de la responsabilité.
- Externalisation sociale : la charge est transférée vers des populations invisibilisées — travailleurs de l'annotation, modérateurs, opérateurs de maintenance — dont le travail soutient le système sans être reconnu comme tel. [40]

Ces formes d'externalisation ne sont pas des dérives accidentelles : elles constituent des mécanismes structurels de soutenabilité apparente du régime. [23][32]

Asymétrie entre production et assomption

Une caractéristique centrale de cette configuration est l'asymétrie entre ceux qui bénéficient de la production d'intelligence artificielle et ceux qui en assument les coûts métaboliques. Les acteurs dominants du régime captent une part disproportionnée de la valeur produite, tout en se protégeant des effets négatifs par des dispositifs contractuels, organisationnels et juridiques.

Cette asymétrie rend la responsabilité difficilement assignable. La dette entropique ne correspond pas à une faute identifiable, mais à une configuration systémique dans laquelle la production de valeur est découplée de l'assomption des coûts.

De la question morale à la question politique

À ce stade, la question « *qui paie ?* » ne peut plus être traitée comme un problème moral individuel. Elle devient une question politique au sens fort : celle de la répartition légitime des coûts irréversibles produits par un système collectif.

Nommer la dette entropique située permet de transformer un ensemble de nuisances dispersées en un objet de conflit explicite. Elle rend visible ce qui était jusque-là naturalisé ou invisibilisé, et ouvre la possibilité d'arbitrages institutionnels — toujours conflictuels — sur la production, la limitation ou l'abandon de certains usages. [23]

Encadré 4 (conceptuel) — Irréversibilité et responsabilité : articulation conflictuelle

La dette entropique située oblige à reconfigurer la notion même de responsabilité. Dans les cadres classiques, être responsable signifie répondre d'un acte réversible : réparer, compenser, corriger. Or l'irréversibilité entropique rend cette conception insuffisante.

Lorsque les effets produits ne peuvent être annulés, la responsabilité ne consiste plus à réparer, mais à reconnaître les limites de l'assomption. Elle devient indissociable de la capacité à nommer les seuils au-delà desquels continuer à produire aggrave la dette au lieu de la résorber. [8][11]

Dans ce contexte, la responsabilité ne précède pas l'irréversibilité : elle en est une réponse toujours tardive, toujours conflictuelle. Plus la reconnaissance est différée, plus la charge à assumer devient lourde, et plus la responsabilité risque de se transformer en désassomption forcée. [11][34]

Cet encadré éclaire ainsi un point central de l'architectonique proposée dans cet ouvrage : la responsabilité n'est pas un principe abstrait, mais un rapport instable à l'irréversibilité. Elle ne se joue pas dans l'intention, mais dans la capacité collective à décider qui paie, jusqu'où, et à partir de quand il devient nécessaire de renoncer.

13.2 Responsabilité comme conflit politico-technique

L'analyse de la dette entropique située conduit à une reformulation décisive de la responsabilité. Celle-ci ne peut plus être pensée comme une propriété morale attachée à des acteurs isolés, ni comme une exigence abstraite susceptible d'être satisfaite par des mécanismes de conformité formelle. Dans le régime métabolique de l'IA contemporaine, la responsabilité apparaît avant tout comme un conflit politico-technique.

Ce conflit naît du décalage structurel entre, d'une part, des systèmes capables de produire des effets à grande échelle et, d'autre part, l'absence de dispositifs permettant d'en assumer collectivement les conséquences irréversibles. La responsabilité n'est donc pas un état à atteindre, mais un champ de tension, traversé par des intérêts divergents, des asymétries de pouvoir et des contraintes matérielles. [23][34]

La responsabilité déplacée par la structure des systèmes

Les architectures contemporaines de l'IA sont caractérisées par une fragmentation extrême des chaînes de décision. Les choix techniques, organisationnels et économiques sont distribués entre une multitude d'acteurs — développeurs, fournisseurs d'infrastructures, intégrateurs, utilisateurs finaux, régulateurs — sans qu'aucun ne puisse prétendre maîtriser l'ensemble du système.

Cette fragmentation produit un effet paradoxal : plus les systèmes sont sophistiqués, plus la responsabilité devient diffuse. La question n'est pas de savoir si chacun respecte ses

obligations locales, mais de constater que le lieu même de la responsabilité se dérobe. La dette entropique s'accumule sans sujet clairement identifiable pour l'assumer.

Responsabilité et arbitrage sous contrainte

Dans ce contexte, être responsable ne signifie plus « répondre de tout », mais arbitrer sous contrainte. Il s'agit de décider :

- quels usages peuvent être maintenus malgré leur coût entropique,
- quels usages doivent être limités ou transformés,
- et lesquels doivent être abandonnés.

Ces arbitrages sont irréductiblement politiques, parce qu'ils impliquent des choix de répartition, des hiérarchies de valeurs et des acceptations de pertes. Ils sont également techniques, parce qu'ils doivent composer avec des infrastructures existantes, des dépendances matérielles et des seuils physiques non négociables. [8][23]

La responsabilité devient ainsi une pratique située, conflictuelle, toujours partielle. Elle ne supprime pas la dette entropique ; elle en organise la reconnaissance et la distribution.

L'illusion de la responsabilité sans coût

Une pathologie centrale des discours contemporains sur l'IA réside dans la promesse implicite d'une responsabilité à coût nul. Cette promesse se manifeste par des appels à l'« IA responsable », à l'« éthique *by design* » ou à la « gouvernance agile », qui supposent que la responsabilité pourrait être intégrée sans affecter significativement la performance, la fluidité ou la rentabilité des systèmes.

L'analyse métabolique montre au contraire que toute responsabilité effective a un coût entropique supplémentaire : ralentissement des flux, augmentation des besoins de supervision, complexification organisationnelle. Refuser de reconnaître ce coût revient à déplacer la responsabilité vers des acteurs moins visibles ou vers le futur. [8][34]

Responsabilité et désassomption

Lorsque les coûts deviennent trop élevés ou visibles, un phénomène inverse apparaît : la désassomption. Celle-ci désigne le moment où les acteurs dominants cessent, explicitement ou implicitement, de répondre des effets qu'ils continuent pourtant de produire.

La désassomption n'est pas un accident moral, mais une issue structurelle des configurations où la dette entropique dépasse les capacités d'assomption collective. À ce stade, la responsabilité ne disparaît pas : elle se transforme en conflit ouvert, marqué par des retraits, des dénis ou des stratégies d'évitement. [23][34]

Encadré 5 (analytique) — Répondabilité et désassomption : deux seuils asymétriques

La répondabilité, introduite dans le régime du sens, et la désassomption, analysée dans le régime métabolique, occupent des positions structurellement symétriques — mais dynamiquement asymétriques — dans l'architectonique de l'intelligence artefactuelle.

La répondabilité désigne la capacité pratique et institutionnelle de répondre du sens produit : expliquer, justifier, corriger, endosser une interprétation au point d'usage. Elle constitue le seuil minimal à partir duquel une production peut être reconnue comme porteuse de sens.

La désassomption, à l'inverse, désigne le seuil au-delà duquel il n'est plus possible de répondre des coûts produits : énergétiques, matériels, organisationnels ou sociaux. Elle marque le point où l'assomption devient matériellement ou politiquement impossible.

Ces deux seuils encadrent l'espace de l'intelligence artefactuelle :

- en amont, sans répondabilité, il n'y a pas de sens stabilisable ;
- en aval, sans capacité d'assomption, il n'y a pas de soutenabilité durable.

Entre les deux se déploie une zone étroite et instable où l'intelligence peut apparaître comme effet conditionnel. Trop peu de répondabilité, et les productions deviennent ininterprétables ou irresponsables ; trop de dette entropique, et les systèmes cessent d'être assumables, même s'ils restent performants.

Cette mise en miroir permet de comprendre pourquoi la gouvernabilité de l'intelligence artefactuelle ne consiste ni à maximiser le sens, ni à minimiser les coûts, mais à maintenir une tension praticable entre répondre et assumer. C'est cette tension — et non son effacement — qui constitue le cœur de l'architectonique proposée dans cet ouvrage. [34][39]

13.3 Rendre discutables les arbitrages

Si la responsabilité apparaît comme un conflit politico-technique, alors la gouvernabilité de l'intelligence artefactuelle ne peut consister ni à éliminer ce conflit, ni à le résoudre par avance au moyen de principes abstraits ou de normes universelles. Elle repose au contraire sur une capacité plus modeste mais plus exigeante : instituer la discutabilité des arbitrages. [34]

On appellera ici *discutabilité* la capacité, pour une organisation ou une institution, à faire tenir des choix de sens, de circulation et de soutenabilité dans un cadre où ils peuvent être contestés, révisés, suspendus ou abandonnés, sans que le système ne devienne immédiatement inopérant. La discutabilité ne se confond donc ni avec la transparence, ni avec la simple publicité des décisions : elle désigne une possibilité effective de reprise.

Arbitrage vs optimisation

L'optimisation suppose un objectif univoque et une métrique commune. Elle vise à améliorer continûment une performance donnée. L'arbitrage, au contraire, reconnaît l'incommensurabilité des valeurs en jeu. Il ne cherche pas à maximiser, mais à trancher sous contrainte, en assumant qu'une perte est inévitable.

Dans le cas de l'IA, optimiser la fluidité communicationnelle peut accroître la dette entropique ; maximiser la répondabilité peut ralentir la circulation ; réduire drastiquement le métabolisme peut appauvrir le sens produit. Ces tensions ne sont pas des défauts de conception : elles constituent la structure même de l'objet. [11]

Maintenir ouverts les arbitrages revient donc à substituer à une logique d'optimisation silencieuse une logique de décision exposée.

La discutabilité comme condition de gouvernabilité

Un arbitrage n'est gouvernable que s'il peut être :

- nommé (identification explicite des régimes en tension),
- situé (inscrit dans un contexte d'usage, de territoire, de temporalité),
- contesté (ouvert à la critique sans être immédiatement disqualifié),
- révisé (susceptible d'être modifié lorsque les conditions changent).

À l'inverse, lorsque les arbitrages sont naturalisés — présentés comme des nécessités techniques, des contraintes économiques ou des fatalités physiques — ils cessent d'être discutables. La dette entropique continue alors de s'accumuler, mais hors de tout espace de délibération. [23][34]

Discutabilité et interfaces

La possibilité de maintenir cette ouverture dépend étroitement des interfaces analysées dans la Partie VI. Ce sont elles qui traduisent les tensions entre régimes en objets manipulables : procédures, seuils, obligations, indicateurs.

Cependant, aucune interface ne garantit à elle seule la discutabilité. Le droit peut rendre opposables certains arbitrages, mais il peut aussi les figer ; l'éthique peut alerter, mais rester sans effet ; la technique peut instrumenter, mais invisibiliser les pertes. La discutabilité n'est jamais donnée : elle doit être instituée et entretenue. [25][37]

13.4 Arbitrages, seuils et gouvernabilité minimale

Préserver des arbitrages contestables ne revient pas à promettre une gouvernance juste ou harmonieuse de l'intelligence artefactuelle. Cela consiste plus modestement à empêcher que les conflits soient dissimulés derrière des discours de performance, d'innovation ou de nécessité.

La gouvernabilité minimale de l'intelligence artefactuelle repose ainsi sur une exigence négative mais décisive : ne pas refermer trop tôt les conflits, et accepter que certaines décisions prennent la forme de renoncements explicites plutôt que de compromis invisibles. [11]

Encadré 6 (synoptique) — Table conceptuelle comparée des trois régimes

Régime	Concept(s) central(aux)	Seuil critique	Pathologies typiques	Effet sur l'intelligence
Sens	Répondabilité ; stabilisation opératoire des interprétations ; endossement au point d'usage	Seuil de répondabilité (impossibilité pratique de répondre du sens produit)	Inflation discursive ; interprétations non endossables ; sens sans responsabilité	Intelligence apparente mais non assumable
Communication	Autorité distribuée ; fiducie ; textoralité ; palimpsestualité	Seuil de fiducie (effondrement de la crédibilité et de l'orientation)	Circulation sans orientation ; signalisation pure ; saturation informationnelle	Intelligence fluide mais vide ou désorientée
Métabolisme	Dette entropique située ; flux ; extractivisme computationnel différé	Seuil de désassomption matérielle (impossibilité d'assumer les coûts)	Insoutenabilité ; externalisation massive ; désassomption	Intelligence performante mais ingouvernable
Lecture transversale de la table : <ul style="list-style-type: none"> • Aucun régime ne produit l'intelligence à lui seul ; • Chaque seuil correspond à une forme spécifique de perte de gouvernabilité ; • Les pathologies ne sont pas des dysfonctionnements accidentels, mais des effets structurels de la domination d'un régime sur les autres. L'intelligence artefactuelle n'apparaît que dans l'intervalle étroit où ces seuils ne sont pas franchis simultanément. [8][28][39]				

13.5 La limite métabolique et la flèche temporelle

Le régime métabolique comme limite matérielle et temporelle de l'intelligence

La Partie IV a proposé un déplacement décisif dans l'analyse de l'intelligence artefactuelle : penser l'IA non seulement comme un régime de sens et de communication, mais comme un régime métabolique, c'est-à-dire comme un ensemble de flux matériels, énergétiques, organisationnels et sociaux dont la dynamique conditionne la possibilité même de toute production de sens et de toute circulation communicationnelle.

Cette partie a montré que l'IA contemporaine ne peut être comprise indépendamment de ses infrastructures, de ses chaînes d'extraction de données, de ses coûts énergétiques, de ses dispositifs de travail humain, et des organisations qui les soutiennent. Loin d'être un simple arrière-plan technique, ce métabolisme constitue une dimension constitutive de l'intelligence artefactuelle. [23][32]

La dette entropique comme opérateur central

L'introduction de la dette entropique située a permis d'unifier des notions jusque-là dispersées — dette technique, dette écologique, externalités, coûts organisationnels — sous un opérateur conceptuel unique. La dette entropique désigne l'accumulation irréversible de désordre, de dissipation et de dégradation résultant de toute opération computationnelle et organisationnelle. [8][11]

Cet opérateur permet de saisir simultanément :

- la dimension thermodynamique de la computation,
- la dégradation informationnelle et la saturation du signal,
- l'usure matérielle et l'obsolescence infrastructurelle,
- la complexification organisationnelle et la charge cognitive,
- l'épuisement social et attentionnel.

En ce sens, le métabolisme n'est pas seulement un régime parmi d'autres : il est le lieu où s'inscrit l'irréversibilité temporelle qui traverse l'ensemble de l'architectonique.

Seuils, retrait, désassomption et gouvernabilité

L'analyse des seuils entropiques a montré que l'accumulation de dette ne constitue pas un phénomène linéaire, mais qu'elle conduit à des points critiques : seuils thermodynamiques, informationnels, organisationnels et sociaux.

Le concept de seuil de désassomption matérielle permet de nommer le moment où les acteurs cessent de pouvoir — ou se trouvent conduits à ne plus vouloir — assumer les coûts métaboliques qu'ils ont contribué à produire. La responsabilité devient alors structurellement problématique : répondre de l'intelligence produite suppose d'endosser une dette qui excède les capacités d'assomption des institutions, des organisations ou des sociétés.

Cette impossibilité prend la forme d'un retrait asymétrique. Le retrait désigne le désengagement d'une trajectoire sociotechnique déjà enclenchée : il ne consiste pas à « arrêter », mais à défaire des engagements matériels, contractuels, organisationnels et symboliques. Or, dans la flèche temporelle, le retrait est structurellement plus coûteux que l'expansion initiale : ce qui a été dissipé, usé, saturé ou rigidifié ne redevient disponible qu'au prix d'une dépense additionnelle. C'est cette asymétrie qui transforme un choix localement réversible en verrouillage durable, et qui rend, au-delà d'un certain point critique, la désassomption plus probable que la réversibilité.

En revanche, la désassomption désigne l'abandon contraint — parfois stratégique, mais jamais pleinement libre — d'une trajectoire, d'un projet ou d'un régime une fois que les seuils de saturation ont été franchis. Elle ne relève pas d'un « retrait volontaire », mais constitue une conséquence de l'hystérésis temporelle : le système ne peut plus être assumé — endossé, soutenu, gouverné — sans coûts devenus prohibitifs.

La désassomption marque ainsi un point de non-retour. Les options se referment, et l'intelligence artefactuelle cesse d'être soutenable non par échec technique, mais par

épuisement des conditions de sa tenue : le sens devient non assumable, la communication non crédible, et le métabolisme ingouvernable.

Ainsi, le métabolisme ne constitue pas seulement une contrainte externe : il est un opérateur de crise qui transforme les conditions mêmes de la responsabilité, de la communication et du sens.

Politique de la dette et conflictualité structurelle

La Partie IV a également montré que la dette entropique n'est jamais abstraite : elle est toujours située et distribuée. Elle se dépose dans des territoires, des corps, des infrastructures, des institutions, des générations futures.

Dès lors, la question centrale n'est pas seulement combien d'entropie est produite, mais qui la supportera, la différera, l'externalisera et en subira les effets. Le régime métabolique révèle ainsi une dimension politique irréductible : gouverner l'IA revient à arbitrer la production et la répartition de la dette entropique.



Figure 5. Schéma de synthèse — Trois régimes, seuils, pathologies

Ce schéma récapitule la logique comparative construite dans les Parties II–IV : chaque régime possède une fonction, un seuil critique, des pathologies typiques et un effet spécifique sur l'intelligence artificielle. Il sert de point d'appui pour la Partie V, consacrée aux configurations de co-appartenance et de déséquilibre.

Transition vers la co-appartenance

Cependant, le métabolisme ne constitue pas un fondement ultime de l'intelligence artificielle. Il ne la détermine pas unilatéralement, pas plus que le sens ou la communication ne peuvent prétendre à une telle primauté.

Ce que cette partie a rendu visible, c'est que le métabolisme contraint les autres régimes sans les fonder, et qu'il est lui-même orienté, justifié et contesté par des régimes de sens et de communication. L'intelligence artefactuelle n'est donc pas un produit du métabolisme, mais un effet conditionnel de la co-appartenance conflictuelle des trois régimes.

La Partie V analysera précisément cette co-appartenance : les tensions, les déséquilibres et les configurations dans lesquelles un régime tend à écraser les autres. Elle montrera que l'intelligence artefactuelle ne disparaît pas dans ces configurations, mais qu'elle se transforme, se déforme et devient, selon les cas, insoutenable, désorientée ou insignifiante. [11][39]

C'est à partir de cette analyse des tensions que pourra être formulée, non pas une théorie unifiée de l'IA, mais une architectonique de l'émergence conditionnelle de l'intelligence artefactuelle.

PARTIE V — Co-appartenance, tensions et émergence de l'intelligence artefactuelle

14. Quand un régime écrase les autres

L'intelligence artefactuelle ne se défait pas seulement lorsque les régimes cessent de fonctionner, mais aussi — et plus insidieusement — lorsque l'un d'eux réussit trop bien au point de neutraliser les contraintes exercées par les deux autres.

Dans ces configurations, le système peut demeurer performant, pertinent ou efficace selon les critères internes du régime dominant, tout en devenant progressivement non gouvernable : les arbitrages ne sont plus visibles, les seuils ne sont plus reconnus, et les coûts sont déplacés hors du champ d'assomption.

Cette partie analyse trois figures typiques de ces déséquilibres structurels. Chacune correspond à l'hypertrophie d'un régime — sens, communication ou métabolisme — lorsque celui-ci cesse d'être contraint par les deux autres. Il ne s'agit pas d'échecs techniques, mais de pathologies de la co-contrainte : l'intelligence continue d'apparaître localement, tandis que ses conditions de tenue se dégradent globalement. [38][39][40]

14.1 Sens sans soutenabilité

La configuration « sens sans soutenabilité » ne se contente pas d'ignorer des contraintes matérielles : elle tend à les accélérer et à les rendre plus difficiles à réassumer dans le temps. Chaque production de sens toujours plus contextualisée, chaque recommandation hyper-personnalisée, chaque extension des chaînes d'opérations (résumés en cascade, outils appelés, agents, réécritures, vérifications, itérations) augmente l'appel aux flux énergétiques, infrastructurels et organisationnels. Autrement dit : la réussite locale du sens se traduit, en régime de modèles, par une intensification du métabolisme — et, avec elle, par une accumulation irréversible de dette entropique située. [8][11][40]

Il faut être précis sur ce que signifie ici « dette ». Il ne s'agit pas seulement d'un coût monétaire ou d'une externalité comptable. La dette entropique désigne un passif dont une partie est

physiquement et organisationnellement irréversible : dissipation énergétique convertie en chaleur, usure et obsolescence accélérées des équipements, extraction et dispersion de matériaux, surcharge durable de maintenance, pollution informationnelle des corpus et saturation des médiations humaines. On peut corriger certains effets, mais au prix d'un supplément d'effort, de ressources et de complexité : le retour à l'état antérieur n'est pas une simple marche arrière.

Les ordres de grandeur* disponibles, indicatifs, suffisent à objectiver cette structure sans la réduire à un chiffre unique. On peut discuter de la part attribuable strictement à l'IA, mais ce débat ne change pas le diagnostic architectonique : la production de sens à grande échelle s'inscrit dans une trajectoire de montée des flux et rapproche des seuils thermodynamiques, économiques et politiques.

Du côté de l'eau, les projections disponibles annoncent des retraits (prélèvements) globaux associés aux usages IA pouvant atteindre plusieurs milliards de m³ à l'horizon 2027, avec des risques de conflits localisés (refroidissement, concurrence avec d'autres usages, contraintes hydriques régionales).

L'essentiel, pour le présent chapitre, n'est pas d'accumuler les chiffres, mais d'explicitier une asymétrie : le régime du sens peut s'intensifier rapidement (plus de contexte, plus de personnalisation, plus de « réponses »), tandis que le régime métabolique impose des seuils. Lorsque le sens est évalué quasi exclusivement à l'aune de sa pertinence locale — et non à l'aune de ses conditions de reproduction dans le temps —, il devient entropiquement prédateur : il consomme du « capital basse entropie » (énergie disponible, eau, composants, attention et médiations humaines) comme si ces réserves étaient extensibles à l'infini.

Cette prédation a des effets en retour sur le sens lui-même. Plus la dette entropique s'accumule, plus les exigences de maintenance et de soutenance deviennent contraignantes : arbitrages sur la capacité, sur les temps de calcul, sur la disponibilité des infrastructures, sur la stabilité des pipelines de données, sur la charge d'équipes humaines devant surveiller, corriger, filtrer, documenter. Quand ces arbitrages sont imposés « par le bas » (pénuries, contraintes locales, conflits d'usage, restrictions de capacité), la qualité du sens se dégrade moins par déficit d'intelligence locale que par affaïssement des conditions de sa répétabilité : latences, simplifications, réduction des ambitions d'usage, fragilisation de la réponsabilité

* Les évaluations de consommation électrique des *data centers* proviennent de rapports récents de l'Agence internationale de l'énergie (IEA). [Selon l'IEA](#), les centres de données ont consommé environ 415 TWh d'électricité en 2024, ce qui représentait à peu près 1,5 % de la consommation mondiale d'électricité en 2024.

L'IEA projette également qu'avec la croissance des usages numériques (dont l'IA), la [demande mondiale d'électricité](#) des centres pourrait plus que doubler d'ici 2030, avec plusieurs scénarios plaçant la consommation au-delà de 1000 TWh dans la seconde moitié de la décennie.

Pour les prélèvements d'eau associés aux usages de l'IA (directs et indirects), des [estimations de la littérature scientifique](#) indiquent que la retraite globale d'eau (*water withdrawal*) liée aux activités d'IA pourrait atteindre plus de 4 milliards à près de 7 milliards de mètres cubes par an vers 2027, selon des calculs basés sur la consommation électrique et les facteurs d'intensité en eau. Ces évaluations distinguent les prélèvements (*withdrawals*) — la quantité d'eau retirée de sources naturelles — des consommations effectives (*consumption*), ces dernières étant généralement plus faibles car une partie de l'eau retirée peut être restituée après usage (par exemple dans les systèmes de refroidissement).

effective (moins de temps humain, moins de procédures de reprise, moins de capacité de vérification). [32][39][40]

De ce point de vue, le « sens sans soutenabilité » n'est pas seulement insoutenable : il tend à raccourcir son propre horizon temporel. Il peut produire des effets d'intelligibilité et de crédibilité spectaculaires dans l'instant, tout en creusant le passif qui rendra ces effets plus coûteux, puis politiquement et matériellement plus difficiles à assumer. C'est ici que la dette entropique joue son rôle architectural : elle rappelle que le métabolisme introduit une direction temporelle et que l'hypertrophie du sens peut devenir un accélérateur de désassomption. [8][11][40]

14.2 Communication sans orientation

La configuration « communication sans orientation » apparaît lorsque le régime communicationnel s'intensifie au point de neutraliser les contraintes exercées par le sens et le métabolisme. Dans ces situations, les flux circulent de manière fluide, rapide et massive, mais cette circulation cesse progressivement d'être orientée par des possibilités effectives de reprise, de discussion ou de correction. La communication ne s'interrompt pas, elle se détache de toute capacité collective à se situer dans ce qui circule.

Dans un tel régime, l'opérateur fiduciaire continue de fonctionner, mais sous une forme dégradée. Les acteurs — humains comme institutionnels — sont contraints de s'en remettre à des chaînes de médiation de plus en plus longues, opaques et automatisées, sans disposer des points de reprise nécessaires pour en interroger localement la validité. La fiducie n'est pas abolie, elle devient forcée et diffuse. On se remet aux flux faute de pouvoir en suspendre le cours, non parce qu'ils sont jugés pertinents ou discutables. [27][28]

Cette hypertrophie communicationnelle se manifeste par une multiplication des énoncés, des relais, des reformulations et des agrégations, rendue possible par la production automatisée de textes, de résumés, de réponses et de contenus synthétiques. La circulation gagne en efficacité ce qu'elle perd en orientation : les messages sont transmis, amplifiés et recombinaés plus vite que ne peuvent se former des désaccords explicites, des mises en contexte ou des corrections partagées. La communication devient ainsi auto-entretenu, indépendamment de toute stabilisation interprétative. [30][38]

Le seuil critique est franchi lorsque la vitesse, le volume ou la densité de circulation excèdent structurellement les capacités de reprise humaine et institutionnelle. À ce point, la discussion ne disparaît pas, elle devient marginale ou tardive ; la contestation ne porte plus sur le contenu des énoncés, mais sur les conditions mêmes de leur prolifération. La communication continue alors de produire des effets de crédibilité et d'autorité apparente, non parce qu'elle est orientée, mais parce qu'elle est omniprésente. [28][38]

Cette situation produit des pathologies spécifiques. L'autorité interprétative se fragmente ou se mime : elle n'est plus liée à des institutions ou à des procédures identifiables, mais à des formats, des métriques d'engagement, des signaux de popularité ou des sorties de modèles perçues comme neutres. La fiducie se déplace vers des dispositifs qui promettent la continuité du flux plutôt que sa discutabilité. Il en résulte une circulation sans responsabilité interprétative, où chacun relaie sans pouvoir répondre du sens ni interrompre la chaîne. [27][28][30]

Comme dans le cas du « sens sans soutenabilité », la communication sans orientation n'est pas simplement problématique du point de vue normatif, elle est instable dans le temps. Plus les flux s'intensifient, plus les coûts de réorientation augmentent : il devient matériellement et organisationnellement plus difficile de ralentir, de hiérarchiser, de créer des espaces de priorité humaine ou de réintroduire des procédures de délibération. La communication gagne en portée immédiate ce qu'elle perd en capacité de se gouverner. [34][38]

Ainsi comprise, la « communication sans orientation » n'est pas une dérive accidentelle, mais une figure structurale de déséquilibre : la confiance continue d'opérer, mais elle n'oriente plus l'action collective. La circulation devient un fait accompli, et non un processus discuté. C'est dans cette dissociation entre circulation et orientation que le régime communicationnel, en s'autonomisant, contribue à rendre l'intelligence artificielle localement performante mais collectivement ingouvernable.

14.3 Métabolisme sans signification

La configuration « métabolisme sans signification » apparaît lorsque le régime métabolique s'impose comme principe organisateur dominant, au point de neutraliser les contraintes exercées par le sens et la communication. Dans ces situations, l'attention se concentre sur l'optimisation des flux, la réduction des coûts, la sobriété apparente ou l'efficacité énergétique et organisationnelle, tandis que s'effacent progressivement les conditions d'endossement interprétatif et de circulation fiduciaire. Le système devient soutenable selon ses propres métriques, mais illisible du point de vue de ce qu'il fait et de ce qu'il engage collectivement. [38][39][40]

Dans ce régime, la dette entropique située n'est plus ignorée : elle est intégrée, mesurée, parfois même réduite. Mais cette intégration se fait au prix d'un déplacement de la question du sens. Les arbitrages métaboliques — limitation des usages, simplification des modèles, rationnement des capacités, priorisation de certains flux — sont présentés comme des nécessités techniques, alors même qu'ils redéfinissent profondément les usages possibles, les formes de décision et les horizons d'action. Le métabolisme « tient », mais sans que les choix qu'il impose soient interprétables ou discutables comme tels. [34][40]

Cette hypertrophie du régime métabolique se manifeste typiquement par une gouvernance par indicateurs de soutenabilité, par seuils énergétiques ou par contraintes d'infrastructure, détachée de toute explicitation des finalités poursuivies. Les systèmes sont ajustés pour rester dans des enveloppes acceptables de consommation, de charge ou de maintenance, mais les raisons pour lesquelles certaines fonctions sont maintenues, étendues ou abandonnées cessent d'être formulées autrement qu'en termes de coût ou de faisabilité. Le métabolisme devient ainsi auto-justifiant.

Le seuil critique est franchi lorsque les contraintes métaboliques, bien que réelles et parfois nécessaires, ne peuvent plus être traduites en décisions compréhensibles et assumables par les acteurs concernés. À ce point, la soutenabilité cesse d'être une condition de l'intelligence pour devenir un substitut au sens. Les systèmes continuent de fonctionner dans les limites fixées, mais ils ne peuvent plus être situés dans un horizon de signification partagée : on sait qu'ils « tiennent », sans savoir ce qu'ils font tenir. [39][40]

Cette configuration produit des pathologies spécifiques. Les arbitrages deviennent techniquement rationnels mais politiquement muets ; les restrictions d'usage apparaissent

comme des fatalités plutôt que comme des choix ; la gouvernance se réduit à la gestion de pénuries ou de contraintes, sans possibilité de réinterprétation collective. La dette entropique est contenue localement, mais au prix d'une désassomption symbolique : plus personne ne peut répondre de la direction prise, seulement de la conformité aux contraintes. [34][40]

Comme les deux figures précédentes, le métabolisme sans signification est instable dans le temps. En l'absence d'orientation interprétative et de circulation discutée, les contraintes métaboliques tendent à se durcir et à s'étendre. Ce qui était présenté comme une optimisation devient un rationnement permanent ; ce qui était une mesure de soutenabilité devient une norme indiscutable. Le système peut alors être soutenable sans être intelligible, et efficace sans être gouvernable. [38][40]

Ainsi comprise, la configuration « métabolisme sans signification » ne décrit pas un excès de matérialité, mais une autonomisation du métabolisme par rapport aux deux autres régimes. La dette entropique y est reconnue, mais elle n'est plus traduite en arbitrages interprétables ni en décisions fiduciairement assumables. L'intelligence artefactuelle peut subsister comme performance technique sous contrainte, mais elle cesse d'être un effet collectif reconnaissable et gouvernable.

15. L'intelligence artefactuelle comme effet

Les trois figures analysées au chapitre 14 — sens sans soutenabilité, communication sans orientation, métabolisme sans signification — permettent de préciser ce que l'intelligence artefactuelle n'est pas. Elle n'est ni garantie par la production locale de sens, ni par la fluidité des circulations, ni par la soutenabilité métabolique prise isolément. Dans chacun des cas, le système peut continuer à fonctionner, parfois de manière impressionnante, tout en perdant les conditions de sa gouvernabilité.

Ces configurations ne décrivent donc pas des échecs techniques, mais des déséquilibres architectoniques. Elles montrent que l'intelligence artefactuelle n'est pas un attribut intrinsèque des systèmes d'IA, mais un effet émergent et conditionnel, produit par la tenue conjointe — toujours précaire — des trois régimes. Lorsque cette co-contrainte se défait, l'intelligence peut persister comme performance locale, mais elle cesse d'être reconnaissable, assumable et gouvernable comme intelligence collective.

C'est en ce sens qu'il faut parler d'effet d'intelligence artefactuelle, et non de faculté artificielle. [15][38][39][40]

15.1 Ni faculté, ni propriété, ni illusion

Qualifier l'intelligence artefactuelle de *faculté* supposerait l'existence d'une capacité unifiée, stable et transférable, analogue à une compétence cognitive. Or les analyses précédentes montrent qu'aucune capacité de ce type ne peut être isolée : ce qui apparaît comme intelligence dépend toujours de conditions matérielles, communicationnelles et interprétatives externes au système. [16][39]

La concevoir comme une *propriété* technique — d'un modèle, d'une architecture ou d'un algorithme — conduit à une autre impasse. Une même configuration technique peut produire des effets d'intelligence dans certains contextes d'usage, et des effets de désorientation ou de désassomption dans d'autres. L'intelligence ne « réside » pas dans le système ; elle se

manifeste dans la relation entre le système, ses usages et ses conditions de reproduction. [20][21]

Enfin, réduire l'intelligence artefactuelle à une *illusion* — cognitive, anthropomorphique ou médiatique — manquerait la dimension structurale du phénomène. Les effets produits sont bien réels : décisions, coordinations, transformations organisationnelles, redistributions de pouvoir. Ce qui est en jeu n'est pas la réalité de l'intelligence, mais les conditions sous lesquelles elle peut être reconnue comme telle. [21][29]

15.2 L'intelligence comme stabilisation reconnue

Dire que l'intelligence artefactuelle est un effet ne signifie pas qu'elle soit éphémère ou arbitraire. Un effet peut être stabilisé, à condition que ses conditions de production et de reconnaissance soient maintenues. [39]

Dans ce cadre, une intelligence artefactuelle apparaît lorsque :

- le sens produit est répondable au point d'usage ;
- la communication permet une circulation fiduciaire orientée et contestable ;
- le métabolisme soutient ces opérations sans accumulation de dette entropique désassumée. [38][39][40]

L'intelligence n'est donc pas donnée une fois pour toutes ; elle est reconnue dans la mesure où ces conditions tiennent conjointement. Cette reconnaissance n'est ni purement subjective, ni strictement juridique : elle est pratique, située et révisable. Elle engage des acteurs, des institutions, des dispositifs, et peut être retirée lorsque les conditions se dégradent. [15][34]

L'intelligence artefactuelle est ainsi moins un état qu'un régime de tenue.

15.3 Conditions d'apparition, de maintien et de disparition

Parce qu'elle est un effet conditionnel, l'intelligence artefactuelle est fondamentalement réversible. Elle peut apparaître, se maintenir ou disparaître sans que le système technique change radicalement.

- ✓ Elle apparaît lorsque les trois régimes parviennent, même temporairement, à rester en tension sans qu'un régime impose sa logique comme norme unique.
- ✓ Elle se maintient tant que les arbitrages entre sens, communication et métabolisme restent visibles, discutables et assumables.
- ✓ Elle disparaît lorsque l'un des régimes écrase durablement les autres, rendant impossible soit l'endossement du sens, soit l'orientation de la circulation, soit l'assomption des coûts.

Cette disparition n'est pas nécessairement spectaculaire. Elle peut prendre la forme d'un glissement progressif : intelligence perçue comme pertinente mais insoutenable, efficace mais désorientante, soutenable mais vide de signification. Dans ces cas, le système continue de fonctionner, mais l'effet d'intelligence se défait. [38][39][40]

Reconnaître cette réversibilité est décisif pour la gouvernance. Gouverner l'intelligence artefactuelle ne consiste pas à garantir son existence, mais à instituer les conditions de sa

tenue, en acceptant qu'elle puisse — et parfois doive — être suspendue, limitée ou retirée. [34][40]

PARTIE VI — Opérateurs et interfaces de l'intelligence artefactuelle

(Partie pleinement théorique, qui explicite les notions permettant de passer de l'analyse à des interventions de gouvernance situées)

Les parties précédentes ont montré que l'intelligence artefactuelle n'est ni une propriété des systèmes, ni un effet automatique de leur déploiement, mais une tenue conditionnelle résultant de la co-contrainte de plusieurs régimes irréductibles. Cette reconnaissance impose un déplacement méthodologique décisif : penser non plus seulement les phénomènes, mais les points à partir desquels ils deviennent gouvernables.

La présente partie propose ce déplacement. Elle n'ajoute pas un nouveau régime, ni une couche normative supplémentaire, mais explicite les outils conceptuels nécessaires pour relier l'analyse architectonique à des gestes d'intervention situés. Les notions d'opérateur et d'interface y sont introduites comme des médiations analytiques : elles permettent d'identifier où un système se défait, et comment agir sans réduire la complexité des régimes en jeu. [34][39]

16. Pourquoi des opérateurs

Si l'intelligence artefactuelle apparaît comme un effet conditionnel et réversible, encore faut-il disposer de concepts capables de saisir ses points de fragilité sans la réduire à une somme de performances, de risques ou d'impacts mesurables. Les catégories descriptives usuelles — modèles, usages, biais, coûts, performances — rendent compte de phénomènes, mais peinent à identifier où et comment un système devient non gouvernable. [34][39]

C'est à cette difficulté que répond l'introduction de la notion d'opérateur. Un opérateur ne décrit pas un état du système ; il désigne une fonction de structuration à partir de laquelle des seuils peuvent être identifiés, des déséquilibres caractérisés, et des arbitrages rendus explicites. Il constitue ainsi un point de passage entre analyse conceptuelle et décision pratique, sans se confondre ni avec une norme, ni avec un indicateur. [15][34]

16.1 Concept, opérateur, indicateur

La traduction d'une analyse architectonique en gestes de gouvernance suppose de distinguer rigoureusement trois types de notions souvent confondues dans les débats contemporains sur l'IA : les concepts, les opérateurs et les indicateurs. Chacune joue un rôle spécifique, aucune ne pouvant se substituer aux autres sans produire des effets de réduction. [15]

Le concept : comprendre et qualifier

Un concept a pour fonction première de rendre intelligible un phénomène. Il permet de nommer, de distinguer et de penser ce qui, sans lui, resterait confus ou invisible. Les notions de *régime*, de *responsabilité*, de *textoralité* ou de *dette entropique située* relèvent de ce registre : elles n'évaluent pas, ne mesurent pas, ne prescrivent pas directement.

Le concept éclaire une structure ; il n'indique pas à lui seul où intervenir ni quand un seuil est franchi.

Pris isolément, le concept reste descriptif ou critique. Il peut nourrir une compréhension fine sans produire de prise pratique, ou au contraire être surinterprété comme un principe normatif général — au risque de l'abstraction morale.

L'indicateur : mesurer et suivre

Un indicateur vise à quantifier ou à objectiver un aspect du réel. Il permet de suivre des évolutions, de comparer des situations, de détecter des tendances. Les métriques de performance, de consommation énergétique, de latence, de volume de données ou de taux d'erreur appartiennent à ce registre.

L'indicateur est indispensable à toute gouvernance, mais il est structurellement local et partiel. Il mesure ce qui est déjà formaté pour être mesuré ; il ne dit rien, par lui-même, de la signification des seuils qu'il détecte. Sans cadre conceptuel ni opérateur, l'indicateur tend à devenir auto-justifiant : ce qui est mesuré devient ce qui compte.

L'opérateur : structurer, articuler, rendre gouvernable

L'opérateur occupe une position intermédiaire et décisive. Il n'est ni un simple concept descriptif, ni un indicateur chiffré. Un opérateur désigne une fonction de structuration à partir de laquelle :

- des phénomènes hétérogènes peuvent être unifiés,
- des seuils critiques peuvent être identifiés,
- des arbitrages deviennent discutables et opposables.

L'opérateur ne dit pas *combien* (comme l'indicateur), ni seulement *ce que c'est* (comme le concept) : il dit où le système tient ou se défait. La responsabilité, la confiance communicationnelle ou la dette entropique située ne sont pas des propriétés mesurables directement ; ce sont des points de gravité analytiques à partir desquels on peut décider de limiter, de suspendre, de rediriger ou d'instituer des contraintes. [34][38][39][40]

Une articulation nécessaire, non substituable

Ces trois registres ne sont ni concurrents ni hiérarchiques. Une gouvernance robuste suppose leur articulation :

- les concepts permettent de comprendre ce qui est en jeu ;
- les opérateurs structurent cette compréhension en zones d'intervention ;
- les indicateurs outillent le suivi et la mise en œuvre.

Confondre ces registres produit des pathologies bien connues : le technicisme sans le sens (indicateurs sans opérateurs), la morale abstraite sans prise (concepts sans interfaces), ou le pilotage aveugle par les chiffres (indicateurs érigés en finalité).

C'est précisément pour éviter ces confusions que la notion d'opérateur est introduite ici : non comme un outil supplémentaire, mais comme une charnière méthodologique entre pensée architectonique et gouvernabilité effective.

16.2 Rôle des opérateurs dans une architectonique

Dans une démarche architectonique, les opérateurs ne sont pas des ajouts conceptuels facultatifs ; ils constituent le point de passage obligé entre la compréhension des régimes et toute prétention à la gouvernabilité. Là où l'analyse des régimes rend visibles des logiques hétérogènes et irréductibles, les opérateurs permettent d'identifier où ces logiques se rencontrent, se contraignent ou se défont. [34][38][39][40]

Un opérateur ne décrit pas un état stable du système. Il fonctionne comme un point de tension : il condense des dynamiques multiples, rend perceptibles des déséquilibres latents et permet de repérer des seuils critiques. En ce sens, l'opérateur n'explique pas ; il met en prise. Il n'offre pas une représentation exhaustive, mais un levier analytique orienté vers l'action.

16.2.1 Opérateurs et co-contrainte des régimes

L'intelligence artefactuelle, telle qu'elle a été analysée dans les parties précédentes, n'apparaît que lorsque les régimes du sens, de la communication et du métabolisme parviennent à se contraindre réciproquement sans se neutraliser. Les opérateurs sont précisément ce par quoi cette co-contrainte devient analysable.

Chaque opérateur met en lumière une condition de tenue du régime auquel il est associé :

- la responsabilité indique si le sens peut être endossé et repris au point d'usage ;
- la fiducie communicationnelle indique si la circulation des énoncés reste orientable et contestable ;
- la dette entropique située indique si les conditions matérielles et temporelles de reproduction restent assumables.

Pris isolément, ces opérateurs n'ont aucune valeur normative absolue. C'est leur articulation qui permet d'identifier des configurations stables, fragiles ou pathologiques. Ils rendent visibles les moments où un régime tend à écraser les autres — non comme une dérive morale, mais comme un déséquilibre structurel.

16.2.2 Opérateurs et identification des seuils

Le rôle central des opérateurs est de permettre l'identification de seuils, et non l'optimisation continue de performances. Un seuil n'est pas un optimum à atteindre, mais un point à partir duquel une qualité se défait : impossibilité d'endossement, saturation fiduciaire, désassomption métabolique.

Les opérateurs rendent ces seuils pensables avant qu'ils ne deviennent irréversibles. Ils permettent d'anticiper des points de bascule où les systèmes continuent de fonctionner tout en cessant d'être intelligibles ou gouvernables. En ce sens, ils orientent l'attention non vers ce qui « marche bien », mais vers ce qui ne tient plus.

16.2.3 Opérateurs et gouvernabilité minimale

Dans cette perspective, gouverner l'intelligence artefactuelle ne consiste pas à maximiser les effets d'intelligence, mais à contenir les effets de désassomption. Les opérateurs fournissent les critères minimaux à partir desquels une intervention devient justifiée : suspension, limitation, redirection, institution de nouvelles contraintes.

Ils ne prescrivent pas des solutions. Ils rendent au contraire visibles les zones de décision, en distinguant ce qui relève d'un arbitrage interprétatif, d'une régulation communicationnelle ou d'une contrainte métabolique. En cela, ils préservent la conflictualité au lieu de la dissoudre dans des indicateurs globaux ou des normes abstraites.

16.2.4 Opérateurs et non-totalisation

Enfin, il est essentiel de souligner que les opérateurs ne totalisent pas l'architectonique. Aucun opérateur ne peut servir de principe ultime de décision, ni être étendu pour subsumer les autres régimes. La réponsabilité n'épuise pas la question de la fiducie ; la fiducie ne résout pas la dette entropique ; la dette entropique ne dit rien, à elle seule, du sens.

Cette non-totalisation est une condition de rigueur, non une faiblesse. Elle empêche toute dérive vers un modèle unique de gouvernance et maintient l'ouverture de l'analyse à d'autres régimes, d'autres opérateurs et d'autres configurations historiques. Les opérateurs ne ferment pas l'architectonique, ils en balisent l'usage.

16.3 Opérateurs propres à chaque régime

Un opérateur est dit propre à un régime lorsqu'il condense la condition minimale sous laquelle ce régime peut contribuer à l'effet d'intelligence, tout en restant irréductible aux deux autres. Il ne s'agit pas d'un mot-clé commode, mais d'un point de gravité : un lieu où la tenue du régime, ses seuils et ses pathologies typiques deviennent repérables.

Dans ce cadre, trois opérateurs structurent l'architectonique : la réponsabilité pour le régime du sens, la fiducie pour le régime de la communication (l'autorité en étant une cristallisation possible), et la dette entropique située pour le régime métabolique. Aucun n'est totalisant : chacun rend gouvernable une dimension du phénomène sans se substituer aux deux autres.

Le chapitre suivant explicite séparément ces opérateurs : leur définition, les seuils qu'ils permettent d'identifier, les formes de déséquilibre qu'ils rendent visibles, et leur articulation aux interfaces qui en permettent l'activation. [34][38][39][40]

17. Les opérateurs des trois régimes

Les opérateurs identifiés au chapitre précédent ne constituent ni des principes normatifs universels, ni des critères d'évaluation généraux. Ils sont des points de gravité analytiques, ajustés à la logique propre de chaque régime, à partir desquels deviennent repérables des seuils de tenue, de déséquilibre et de désassomption. [15][34]

Les sections qui suivent examinent successivement ces opérateurs. Il ne s'agit pas d'en proposer une définition abstraite ou exhaustive, mais d'expliciter :

- ce qu'ils rendent gouvernable,
- les seuils critiques qu'ils permettent d'identifier,

- et les pathologies typiques qui apparaissent lorsqu'ils sont ignorés ou absolutisés.

Chaque opérateur sera ainsi présenté comme une condition minimale de contribution à l'effet d'intelligence, toujours partielle, toujours située, et toujours dépendante de son articulation avec les autres régimes.

17.1 Répondabilité (sens)

La répondabilité est l'opérateur propre du régime du sens parce qu'elle condense la condition minimale sous laquelle une production interprétative peut être reconnue, reprise et gouvernée. Elle ne désigne ni la compréhension interne d'un système, ni la vérité intrinsèque de ses énoncés, mais la possibilité pratique et institutionnelle de répondre du sens produit au point d'usage.

Dans le contexte de l'intelligence artefactuelle, cette distinction est décisive. Les systèmes contemporains peuvent produire des énoncés syntaxiquement corrects, contextuellement pertinents et pragmatiquement ajustés, sans que personne ne puisse raisonnablement en répondre. La question du sens se déplace alors : elle ne porte plus sur ce que le système « veut dire », mais sur qui peut endosser ce qui est dit, selon quelles procédures, et avec quels droits de reprise. [15][39]

Répondabilité et endossement du sens

La répondabilité désigne la possibilité de relier une production de sens à une chaîne d'endossement identifiable. Cette chaîne n'implique pas nécessairement un auteur au sens classique, ni une intention subjective. Elle suppose en revanche que des acteurs — humains ou institutionnels — puissent :

- expliciter les conditions de production,
- reconnaître les limites connues,
- suspendre ou restreindre un usage,
- corriger ou requalifier une sortie,
- et assumer les effets produits dans un contexte donné.

Un sens répondable n'est pas un sens totalement explicable ; c'est un sens assumable. À l'inverse, un sens non répondable peut rester localement pertinent tout en devenant globalement ingouvernable. [34][39]

Seuil critique : l'impossibilité d'endossement

Le seuil critique du régime du sens est franchi lorsque aucune entité ne peut plus raisonnablement répondre des effets produits. Ce seuil ne se manifeste pas nécessairement par des erreurs manifestes ou des incohérences visibles. Il peut apparaître dans des contextes où les performances sont élevées, précisément parce que la fluidité et l'efficacité des réponses masquent l'absence de points de reprise effectifs.

Lorsque ce seuil est franchi, le régime du sens ne s'effondre pas ; il s'autonomise. La production interprétative se poursuit, mais elle se détache de toute capacité collective à l'assumer. Le sens cesse alors d'être un opérateur de coordination pour devenir un flux opaque, difficilement contestable et politiquement fragile. [39]

Pathologies typiques de la non-répondabilité

Trois pathologies principales peuvent être identifiées lorsque la réponsabilité est ignorée ou affaiblie :

1. Prolifération d'énoncés non endossables

Le système produit des réponses pertinentes sans que les conditions de leur usage soient clarifiables.

2. Déplacement de la responsabilité

La responsabilité est reportée vers l'aval (usagers) ou l'amont (concepteurs), sans dispositif concret de reprise au point d'usage.

3. Neutralisation de la discussion

Faute de prise sur les conditions de production du sens, la contestation se réduit à des réactions périphériques, sans effet sur les processus qui stabilisent effectivement les interprétations.

Ces pathologies ne relèvent pas d'un déficit cognitif, mais d'un défaut architectural. [34][39]

Articulation et limites de l'opérateur

La réponsabilité n'épuise pas la question du sens. Elle ne garantit ni la qualité épistémique des énoncés, ni leur valeur morale, ni leur légitimité politique. Elle ne peut fonctionner qu'articulée à :

- des conditions de fiducie communicationnelle, qui rendent la contestation praticable,
- des conditions métaboliques soutenables, qui permettent la répétabilité des dispositifs d'endossement.

C'est précisément parce qu'elle est partielle et non totalisante que la réponsabilité peut jouer son rôle opératoire : rendre visibles les seuils à partir desquels le sens continue de circuler tout en cessant d'être assumable.

La section suivante examinera l'opérateur propre du régime de la communication. Là où la réponsabilité concerne l'endossement du sens au point d'usage, la fiducie concerne les conditions sous lesquelles les énoncés peuvent circuler, persister et faire autorité sans dissoudre la possibilité de leur reprise collective. [38][39][40]

17.2 Fiducie et autorité distribuée (communication)

Le régime de la communication ne se réduit ni à la transmission d'informations ni à l'échange de messages entre acteurs identifiables. Il organise la circulation des énoncés, leur visibilité, leur persistance et leur crédit dans un milieu de médiations hétérogènes. Dans ce régime, la question décisive n'est pas seulement celle de la validité des contenus, mais celle des conditions sous lesquelles il devient possible de s'en remettre à ce qui circule. [38][39][40]

L'opérateur propre de ce régime est la fiducie. Elle désigne la capacité collective à faire confiance à des chaînes de médiation — humaines et techniques — sans renoncer à la possibilité de les orienter, de les contester ou de les interrompre. La fiducie n'est pas un état

psychologique, mais une relation structurée, toujours partielle et toujours réversible, entre des acteurs, des dispositifs et des institutions.

Pourquoi la fiducie, et non l'autorité

L'autorité joue un rôle central dans la communication, mais elle ne peut en constituer l'opérateur principal. L'autorité est une cristallisation locale : elle désigne des points où le crédit se stabilise temporairement — une source reconnue, une institution légitime, un dispositif réputé fiable. La fiducie, au contraire, concerne le milieu de circulation dans lequel ces cristallisations deviennent possibles.

Dans les environnements communicationnels contemporains, l'autorité est de plus en plus distribuée et cumulative. Elle ne repose plus sur un acte fondateur unique, mais sur une série de confirmations partielles : citations, reprises, classements, recommandations, métriques de visibilité. Cette autorité en chaîne ne peut fonctionner que sur fond de fiducie : une disposition collective à accepter des médiations dont le fonctionnement précis demeure partiellement opaque.

C'est pourquoi la fiducie est plus fondamentale que l'autorité. Elle conditionne l'apparition, la stabilité et la contestabilité des formes d'autorité, sans jamais se confondre avec elles.

Seuil critique : la saturation fiduciaire

Le seuil critique du régime de la communication est franchi lorsque la circulation devient plus rapide, plus dense et plus efficace que toute possibilité de reprise collective. Ce seuil ne correspond pas à une disparition de la confiance, mais à sa saturation.

Dans une situation de saturation fiduciaire :

- les énoncés continuent de circuler et de produire des effets,
- les chaînes d'autorité se renforcent par auto-référence,
- mais les acteurs perdent la capacité pratique de discuter, de corriger ou de suspendre ce qui circule.

La fiducie se transforme alors en fiducie forcée : on s'en remet aux flux non parce qu'on les juge légitimes, mais parce qu'il devient coûteux, voire impossible, de s'en extraire.

Pathologies typiques de la communication désorientée

Lorsque la fiducie est dégradée ou absolutisée, plusieurs pathologies apparaissent :

1. Dissolution de l'autorité interprétative

L'autorité ne disparaît pas ; elle se fragmente en micro-crédits opaques, difficiles à relier à des responsabilités identifiables.

2. Étouffement de la discussion

La discussion ne disparaît pas non plus, mais elle est neutralisée par la vitesse, la volumétrie et la standardisation des flux.

3. Déliaison entre circulation et sens

Les énoncés acquièrent du crédit indépendamment de leur endossement ou de leur pertinence interprétative située.

Ces pathologies ne sont pas des dérives morales ; elles résultent d'une autonomisation du régime communicationnel, lorsqu'il cesse d'être contraint par les régimes du sens et du métabolisme.

Articulation et limites de l'opérateur

La fiducie ne garantit ni la véracité des énoncés, ni leur valeur interprétative, ni leur soutenabilité matérielle. Elle ne peut fonctionner qu'articulée à :

- la réponsabilité, qui fournit des prises pour l'endossement et la contestation locales ;
- des conditions métaboliques qui empêchent la croissance indéfinie des flux.

C'est précisément parce que la fiducie est un opérateur relationnel, fragile et non totalisant qu'elle peut jouer son rôle : maintenir une circulation possible sans dissoudre la capacité collective à orienter ce qui circule.

La section suivante examinera le troisième opérateur, propre au régime métabolique. Là où la réponsabilité et la fiducie concernent respectivement l'endossement du sens et l'orientation de la circulation, la dette entropique située introduit une contrainte irréversible : celle du temps, des flux matériels et des coûts qui rendent ces deux régimes soutenables — ou non — dans la durée.

17.3 Dette entropique située (métabolisme)

Le régime métabolique soutient matériellement l'ensemble de l'intelligence artefactuelle. Il comprend les flux d'énergie, de calcul, de données, de matières, de travail humain et d'organisations nécessaires à la production et à la circulation du sens. À la différence des régimes du sens et de la communication, il est soumis à une irréversibilité structurelle : les coûts qu'il engage ne peuvent être ni intégralement compensés, ni simplement annulés a posteriori. [38][39][40]

L'opérateur propre à ce régime est la dette entropique située. Il désigne l'accumulation irréversible de désordre matériel, énergétique, informationnel et organisationnel produite par le fonctionnement des systèmes d'intelligence artefactuelle, accumulation qui se dépose toujours quelque part, sur quelqu'un, à un moment donné.

Dette entropique et flèche du temps

Parler de dette entropique ne revient pas à moraliser les coûts, ni à réduire la soutenabilité à une métrique environnementale isolée. L'opérateur entropique introduit une flèche du temps dans l'architectonique : ce qui est dissipé, extrait, utilisé ou saturé ne peut être récupéré qu'au prix d'un supplément de flux, de travail, de complexité et de ressources économiques.

Cette irréversibilité distingue radicalement le régime métabolique des deux autres. Le sens peut être requalifié, la communication peut être réorientée ; la dette entropique, elle, ne se réécrit pas sans pertes. Elle marque les limites au-delà desquelles la répétabilité des dispositifs devient incertaine, puis impossible.

Seuil critique : la désassomption

Le seuil critique du régime métabolique est atteint lorsque les coûts cumulés dépassent ce qui peut être politiquement, socialement ou organisationnellement assumé. Ce seuil ne correspond pas nécessairement à une pénurie immédiate ou à un effondrement technique. Il se manifeste souvent par une série de signaux faibles : arbitrages contraints, dégradations progressives, conflits d'usage, fragilisation des infrastructures et surcharge durable du travail humain.

À partir de ce point, le métabolisme peut continuer à fonctionner, mais au prix d'une désassomption : les coûts sont déplacés, externalisés ou différés, sans qu'une entité puisse raisonnablement en répondre. La dette entropique devient alors un facteur de déstabilisation politique autant que technique.

Pathologies typiques de l'ignorance entropique

Lorsque la dette entropique est ignorée ou neutralisée, plusieurs pathologies apparaissent :

1. Illusion de soutenabilité

Les systèmes semblent stables localement, tandis que les coûts s'accumulent ailleurs ou plus tard.

2. Accélération prédatrice

L'intensification des usages accroît les flux nécessaires, raccourcissant l'horizon de gouvernabilité.

3. Conflits de répartition

La question n'est plus de savoir si une dette existe, mais qui la supporte : territoires, institutions, travailleurs, générations futures.

Ces pathologies ne résultent pas d'un manque de volonté, mais d'une déconnexion architectonique entre métabolisme, sens et communication.

Articulation et limites de l'opérateur

La dette entropique située ne dit rien, à elle seule, de la valeur du sens produit ni de la légitimité de sa circulation. Elle ne peut servir de principe souverain de décision. En revanche, elle impose une contrainte non négociable : toute configuration d'intelligence artefactuelle qui ignore durablement ses conditions métaboliques réduit mécaniquement son horizon de gouvernabilité.

C'est précisément parce qu'il est un opérateur limite — structurant mais non totalisant — que le concept de dette entropique peut être mobilisé sans réduire l'architectonique à un déterminisme matériel. Il n'annule pas les arbitrages, il en fixe les bornes.

Après avoir examiné les trois opérateurs propres à chaque régime, il reste à expliciter une condition méthodologique décisive : pourquoi aucun de ces opérateurs ne peut être totalisant, et pourquoi leur articulation conflictuelle constitue non une faiblesse, mais la condition même de la gouvernabilité de l'intelligence artefactuelle. [15][40]

17.4 Pourquoi aucun opérateur n'est totalisant

Les trois opérateurs analysés — réponsabilité, confiance, dette entropique située — jouent un rôle structurant dans l'architecture de l'intelligence artificielle. Leur importance pourrait suggérer qu'il serait possible d'en privilégier un comme principe général de gouvernance, capable de subsumer les autres. Une telle tentation doit être explicitement refusée. [15][34]

Aucun opérateur ne peut être totalisant, pour une raison simple : chacun est ajusté à la logique propre d'un régime irréductible. La réponsabilité rend gouvernable l'endossement du sens, mais ne dit rien de la circulation effective des énoncés ni de leur soutenabilité matérielle. La confiance oriente la circulation et stabilise des formes d'autorité distribuée, mais ne garantit ni l'endossement local du sens ni l'assomption des coûts engagés. La dette entropique située introduit une contrainte temporelle et matérielle non négociable, mais ne fournit aucun critère interprétatif ou communicationnel.

Toute tentative de totalisation produit des effets pathologiques bien identifiés. L'absolutisation de la réponsabilité conduit à des dispositifs lourds, paralysants ou symboliques, incapables de suivre la dynamique réelle des usages. Portée seule, la confiance favorise des régimes de circulation auto-référentiels, où la confiance devient un substitut à la contestabilité. Enfin, faire de l'entropie la contrainte souveraine réduit la gouvernance à une logique de restriction ou de renoncement, au risque de neutraliser toute production de sens ou toute médiation collective.

La non-totalisation des opérateurs n'est donc pas un compromis théorique, mais une condition de rigueur méthodologique. Elle permet de maintenir visibles les tensions et les conflits qui traversent l'intelligence artificielle, au lieu de les dissoudre dans un critère unique. Gouverner l'IA ne consiste pas à appliquer un principe souverain, mais à arbitrer entre des contraintes hétérogènes, toujours situées, toujours discutables.

Cette non-totalisation implique également que les opérateurs ne puissent agir seuls. Pour devenir opératoires sans être neutralisés, ils doivent être médiés : traduits en dispositifs, en procédures, en formats et en obligations capables de relier analyse et action sans réduire la complexité des régimes. [15][34][38][39][40]

Le chapitre suivant examine précisément ces médiations. Il montre comment les interfaces — techniques, éthiques, normatives ou hybrides — permettent d'activer les opérateurs dans des contextes concrets, en rendant les arbitrages visibles, opposables et révisables, sans prétendre résoudre définitivement les tensions qu'ils mettent en évidence.

18. Interfaces : traduire sans réduire

Les opérateurs identifiés au chapitre précédent rendent visibles les conditions sous lesquelles les régimes de l'intelligence artificielle tiennent, se déséquilibrent ou se défont. Ils ne suffisent toutefois pas, à eux seuls, à produire de la gouvernabilité : entre l'analyse architectonique et l'intervention effective, une médiation est nécessaire.

Les interfaces désignent ces dispositifs de traduction. Elles opèrent à l'intersection des régimes, en transformant des exigences analytiques — réponsabilité, confiance, soutenabilité — en contraintes, procédures ou formats praticables, sans prétendre les ramener à une métrique unique. Elles rendent ainsi possible une gouvernabilité minimale : non pas

maîtriser entièrement des systèmes complexes, mais instituer des prises d'intervention avant que les déséquilibres ne deviennent irréversibles. [20][21]

18.1 Pourquoi des interfaces

Les opérateurs indiquent où la co-contrainte se fragilise (seuils, pathologies, points de bascule), mais ils n'indiquent pas comment intervenir dans une situation donnée. Les interfaces comblent cet écart : elles donnent une forme opérable aux tensions interrégimes afin qu'elles puissent faire l'objet d'arbitrages effectifs. Une interface n'est donc ni un régime supplémentaire, ni un simple outil : c'est un dispositif de traduction qui rend certaines exigences activables, contestables et, parfois, opposables. [3][20]

La typologie des interfaces est posée à l'**annexe A (A.4)** ; le présent chapitre en examine le fonctionnement, les articulations et les modes d'échec.

Interfaces et refus du double réductionnisme

La notion d'interface permet d'éviter deux formes de réduction symétriques.

La première est le technicisme, qui suppose que les problèmes de gouvernance puissent être résolus par l'ajout de fonctionnalités, de métriques ou de contrôles automatiques. La seconde est le moralisme abstrait, qui énonce des principes sans se soucier des conditions concrètes de leur mise en œuvre.

Les interfaces se situent entre ces deux écueils. Elles ne produisent ni des solutions purement techniques, ni des normes désincarnées. Elles traduisent des exigences en contraintes praticables, tout en laissant visibles les choix et les conflits qu'elles impliquent. [31][34]

Interfaces et institution des seuils

Le rôle central des interfaces est d'instituer des seuils. Un seuil n'est pas un optimum à atteindre, mais une limite à partir de laquelle un usage doit être requalifié, restreint ou suspendu. Les interfaces donnent une forme opérable à ces seuils : elles déterminent quand un déséquilibre devient juridiquement, organisationnellement ou techniquement pertinent.

Sans interfaces, les seuils identifiés par les opérateurs restent théoriques. Avec des interfaces, ils deviennent opposables : ils peuvent être invoqués, contestés, révisés et appliqués dans des contextes situés. [34][37]

Interfaces et gouvernabilité minimale

Dans cette perspective, gouverner l'intelligence artificielle ne consiste pas à maîtriser entièrement des systèmes complexes, mais à instaurer une gouvernabilité minimale : un ensemble de prises qui permettent d'intervenir avant que les déséquilibres ne deviennent irréversibles.

Les interfaces rendent possible cette gouvernabilité en articulant :

- des formats techniques (traçabilité, quotas, protocoles),
- des exigences éthiques (dignité, reconnaissance, justice),
- des dispositifs normatifs et juridiques (obligations, procédures, sanctions),
- et des médiations hybrides capables de relier ces dimensions sans les confondre.

Elles ne neutralisent pas les tensions entre régimes, elles les rendent praticables. [31][37]

La section suivante proposera une typologie des principales interfaces utilisables dans une gouvernance architectonique de l'intelligence artefactuelle, en distinguant leurs fonctions, leurs limites et leurs modes d'articulation.

18.2 Interface technique

L'interface technique constitue la forme la plus immédiatement visible — et souvent la plus surestimée — de médiation dans la gouvernance de l'intelligence artefactuelle. Elle désigne l'ensemble des dispositifs matériels et logiciels par lesquels des contraintes issues des régimes du sens, de la communication et du métabolisme sont traduites en formats calculables, traçables ou exécutables.

Contrairement à une idée répandue, l'interface technique ne résout pas les tensions architectoniques, elle les rend opérables. Elle permet d'inscrire certaines exigences — traçabilité, limitation, suspension, priorisation — dans des mécanismes concrets, sans pour autant décider de leur légitimité ni de leur suffisance. [16][20]

Fonction principale : rendre les contraintes opérables

La fonction centrale de l'interface technique est de transformer des exigences qualitatives en prises techniques. Elle agit comme un plan de traduction minimal entre analyse et action.

À ce titre, elle peut notamment :

- rendre visibles des chaînes de production (journalisation, versionnage, traçabilité),
- instituer des limites d'usage (quotas, plafonds, priorités),
- permettre des interruptions ou des reprises (dispositifs d'arrêt d'urgence, escalade humaine),
- différencier des contextes d'usage (modes, périmètres, accès conditionnels),
- objectiver certains coûts (mesures énergétiques, charge computationnelle).

Ces dispositifs ne définissent pas ce qui *doit* être fait, mais rendent possible ce qui *peut* être fait sans réécrire l'ensemble du système. [20][34]

Repères d'articulation aux opérateurs

L'interface technique n'a de sens que par son articulation aux opérateurs :

- Répondabilité (sens)
Traçabilité des sources et des transformations, journalisation des décisions, mécanismes de reprise ou de suspension au point d'usage. [18][34][39]
- Fiducie (communication)
Marquage des contenus, limitation de la génération automatique, dispositifs de priorité humaine, séparation des flux. [27][28][38]
- Dette entropique située (métabolisme)

Mesures énergétiques et matérielles, quotas de calcul, ralentissement volontaire, contraintes d'allocation des ressources. [8][32][40]

Sans cette articulation, l'interface technique tend à devenir un simple outil d'optimisation, indifférent aux effets de sens et de circulation.

Limites structurelles

L'interface technique ne peut pas, à elle seule :

- décider des seuils légitimes,
- instituer des responsabilités,
- arbitrer des conflits de valeurs,
- garantir la contestabilité des décisions.

Lorsqu'elle est absolutisée, elle produit une pathologie bien connue : la gouvernance par défaut technique, où ce qui est implémenté devient ce qui est considéré comme acceptable. Les contraintes techniques se substituent alors aux arbitrages éthiques et politiques qu'elles étaient censées médiatiser. [20][21]

Rôle dans une gouvernance architectonique

Dans une gouvernance architectonique, l'interface technique doit être conçue comme une condition nécessaire mais jamais suffisante. Elle fournit les leviers opérationnels par lesquels d'autres interfaces — éthiques, juridico-normatives, éthico-normatives — peuvent prendre appui.

Elle n'est pas un lieu de décision souverain, mais un support de décisions situées, toujours susceptibles d'être révisées, désactivées ou reconfigurées à mesure que les régimes évoluent.

La section suivante examinera une autre forme d'interface, moins immédiatement opératoire mais tout aussi décisive : l'interface éthique, qui explicite les valeurs et les horizons normatifs à partir desquels les contraintes techniques et juridiques prennent sens.

18.3 Interface éthique

L'interface éthique intervient là où les opérateurs rendent visibles des tensions, mais où aucune traduction technique ou juridique ne peut, à elle seule, décider de leur orientation. Elle ne vise ni à produire des règles directement applicables, ni à fournir une justification morale ultime. Sa fonction première est de mettre en intelligibilité les valeurs, les finalités et les critères d'évaluation qui orientent — souvent implicitement — les arbitrages entre régimes. [31]

Contrairement à une conception prescriptive de l'éthique, l'interface éthique ne dit pas ce qui doit être fait en toutes circonstances. Elle rend visibles les choix axiologiques déjà à l'œuvre dans les dispositifs techniques, organisationnels et normatifs, afin qu'ils puissent être discutés, assumés ou révisés.

Fonction principale : expliciter ce qui compte

La fonction centrale de l'interface éthique est de transformer des désaccords latents en questions explicites. Elle permet de formuler ce qui, sans elle, resterait enfoui sous des critères de performance, des contraintes techniques ou des obligations procédurales.

À ce titre, elle peut notamment :

- expliciter les valeurs mobilisées (dignité, justice, autonomie, reconnaissance, sobriété),
- rendre formulables des conflits de valeurs (efficacité vs équité, fluidité vs contestabilité, innovation vs soutenabilité),
- rendre visibles des asymétries d'impact (qui en bénéficie, qui supporte les coûts, qui peut contester),
- interroger les finalités implicites des usages (à quoi sert ce système, pour qui, dans quel horizon, avec quelles pertes acceptées).

L'interface éthique ne tranche pas ces conflits, elle les constitue comme tels, en les rendant nommables et discutables. [31][32]

Repères d'articulation aux opérateurs

L'interface éthique ne s'ajoute pas aux opérateurs, elle les oriente en donnant un langage aux critères d'acceptabilité et aux pertes assumées.

- Répondabilité (sens)
Elle explicite ce qui est jugé endossable, contestable ou suspendable au point d'usage, et selon quels critères. [15][18][39]
- Fiducie (communication)
Elle éclaire les formes de confiance jugées légitimes ou abusives, et les figures d'autorité considérées comme acceptables. [27][28][38]
- Dette entropique située (métabolisme)
Elle rend visibles les arbitrages intergénérationnels, territoriaux et sociaux impliqués par l'accumulation de coûts irréversibles. [8][23][40]

Sans cette mise en intelligibilité, les exigences restent soit purement techniques (donc dépolitisées), soit purement juridiques (donc réduites à la conformité), soit purement morales (donc inopérantes).

Limites structurelles

L'interface éthique ne peut pas :

- rendre des exigences opposables,
- instituer des sanctions,
- garantir la mise en œuvre effective des décisions,

- résoudre les conflits qu'elle explicite.

Lorsqu'elle est traitée comme suffisante, elle produit une pathologie symétrique au technicisme : la moralisation sans prise. Les valeurs sont affirmées, mais elles n'organisent ni l'action, ni la responsabilité, ni la réversibilité des décisions.

Rôle dans une gouvernance architectonique

Dans une gouvernance architectonique, l'interface éthique joue un rôle irremplaçable mais limité : elle maintient ouvertes les questions de valeur au cœur des dispositifs, sans prétendre les stabiliser une fois pour toutes. Elle fournit le langage dans lequel les arbitrages peuvent être formulés et justifiés publiquement — condition nécessaire pour que les autres interfaces puissent intervenir sans se refermer sur elles-mêmes.

La section suivante examinera l'interface normative et juridique, qui intervient lorsque certains arbitrages doivent cesser d'être seulement discutables pour devenir opposables, en instituant des obligations, des procédures et des seuils.

18.4 Interface juridico-normative

L'interface normative et juridique intervient lorsque les tensions mises en évidence par les opérateurs et explicitées par les interfaces techniques et éthiques doivent cesser d'être seulement discutables pour devenir opposables. Elle ne vise pas à résoudre les conflits entre régimes, mais à instituer des contraintes qui rendent certaines décisions irréversibles, certaines obligations explicites et certaines responsabilités assignables.

Dans une architectonique de l'intelligence artefactuelle, le droit n'est ni un fondement de l'intelligence, ni un régime autonome. Il fonctionne comme une interface d'institution : il traduit des arbitrages en règles, des seuils en obligations, et des désaccords en procédures. [25][37]

Fonction principale : rendre les exigences opposables

La fonction centrale de l'interface normative est de transformer des exigences analytiques ou éthiques en obligations effectives. Là où l'interface éthique explicite ce qui compte, l'interface juridique décide ce qui s'impose, au moins provisoirement.

À ce titre, elle peut notamment :

- instituer des obligations procédurales (documentation, audit, traçabilité),
- définir des droits opposables (information, contestation, recours),
- assigner des responsabilités formelles (charge de la preuve, devoirs de vigilance),
- fixer des seuils non négociables (interdictions contextuelles, limitations d'usage),
- prévoir des sanctions et des mécanismes de réparation.

Ces dispositifs ne garantissent pas la justesse des décisions, mais leur tenue institutionnelle. [25][37]

Repères d'articulation aux opérateurs

L'interface normative prend sens par son articulation aux opérateurs :

- Répondabilité (sens)

Elle rend opposables certaines exigences d'endossement : obligation d'information, droit de contestation, procédures de reprise ou de suspension. [37][39]

- Fiducie (communication)

Elle institue des règles de circulation : transparence minimale, signalement, séparation des usages, protection de certains espaces de délibération. [28][37][38]

- Dette entropique située (métabolisme)

Elle transforme des contraintes de soutenabilité en obligations : normes environnementales, devoirs de vigilance, limites d'allocation des ressources. [8][40]

Sans cette articulation, le droit risque de flotter au-dessus des dispositifs qu'il prétend encadrer ; avec elle, il devient une prise réelle sur les configurations techniques et organisationnelles.

Limites structurelles

L'interface normative ne peut pas :

- garantir la qualité du sens produit,
- maintenir à elle seule la fiducie communicationnelle,
- ni résoudre les contraintes métaboliques qu'elle encadre.

Lorsqu'elle est absolutisée, elle produit une pathologie spécifique : le juridisme procédural. Les obligations sont respectées formellement, tandis que les déséquilibres architectoniques persistent ou se déplacent. Le droit devient alors un dispositif de conformité plutôt qu'un outil de gouvernabilité. [37]

Rôle dans une gouvernance architectonique

Dans une gouvernance architectonique, l'interface normative joue un rôle décisif mais limité. Elle fixe des bornes là où la discussion doit s'arrêter temporairement, sans prétendre clôturer les conflits qu'elle institutionnalise.

Elle ne remplace ni l'analyse, ni l'éthique, ni la technique. Elle en est la traduction opposable, toujours située, historiquement variable et révisable. C'est à ce prix — en acceptant son caractère partiel et conflictuel — que le droit peut contribuer à la gouvernabilité de l'intelligence artefactuelle sans en prétendre être le principe souverain.

La section suivante examinera une forme d'interface hybride — éthico-normative — qui vise précisément à éviter la dissociation entre principes éthiques abstraits et obligations juridiques purement procédurales.

18.5 Interface éthico-normative

L'interface éthico-normative intervient lorsque deux insuffisances deviennent manifestes : d'un côté, des valeurs explicitement affirmées restent sans effet ; de l'autre, des obligations juridico-procédurales risquent de se refermer sur la conformité. Elle ne remplace ni l'éthique

ni le droit : elle vise à lier explicitement des principes à des obligations situées, en maintenant la possibilité de révision. [31][37]

Fonction principale : rendre une valeur gouvernante

Concrètement, l'interface éthico-normative opère par des mécanismes de traduction qui évitent à la fois le déclaratif et la pure procédure :

1. Obligations graduées par criticité

Un même principe (dignité, non-malfaisance, justice, sobriété) n'engendre pas les mêmes obligations selon les usages. L'interface produit des exigences différenciées (renforcées dans le critique, allégées dans l'exploratoire).

2. Droits formels + capacités effectives

Elle couple les droits (information, contestation, recours) à des moyens réels : délais, accès, responsabilités assignées, procédures de suspension. Un droit sans capacité d'exercice demeure symbolique.

3. Clauses de réversibilité et seuils de suspension

Elle rend la gouvernance compatible avec l'incertitude : réévaluations périodiques, points de reprise, conditions explicites déclenchant restriction, suspension ou renoncement.

4. Audits orientés par principes

Elle inscrit des principes dans des formats contrôlables : questions opposables, preuves attendues, périmètres, traçabilité des arbitrages — sans réduire le principe à une case à cocher.

Repères d'articulation aux opérateurs

L'interface éthico-normative joue un rôle clé dans l'activation conjointe des opérateurs :

- Répondabilité (sens)

Elle précise jusqu'où va l'obligation d'endossement, et pour quels usages. [34][37][39]

- Fiducie (communication)

Elle détermine les conditions d'une confiance soutenable sans devenir aveugle (zones protégées, priorités humaines, transparence ciblée). [27][28][38]

- Dette entropique située (métabolisme)

Elle permet d'inscrire des principes (sobriété, justice intergénérationnelle) dans des obligations proportionnées aux impacts réels. [8][23][37][40]

Limites structurelles

Lorsqu'elle se rigidifie, l'interface éthico-normative peut produire une pathologie inverse : la morale surcodée, où des principes deviennent des obligations trop fixes, incapables

d'épouser la diversité des situations et l'évolution des seuils. Elle doit donc rester contestable et révisable. [31]

Certaines initiatives contemporaines — telles que les « constitutions » d'IA intégrées à l'entraînement des modèles — illustrent cette fonction d'interface éthico-normative. Une analyse détaillée de ce type de dispositif est proposée à l'**annexe C**.

Rôle dans une gouvernance architectonique

Dans une gouvernance architectonique, l'interface éthico-normative joue un rôle de charnière. Elle empêche que l'éthique reste un discours d'accompagnement, et que le droit se referme sur une conformité sans sens. Elle maintient une tension productive entre valeurs et obligations, en rendant les arbitrages visibles, discutables et partiellement opposables.

Elle n'achève pas la gouvernance ; elle en préserve la cohérence minimale, en garantissant que les décisions normatives restent reliées aux horizons de sens qui les justifient.

18.6 Pourquoi les interfaces ne sont pas des régimes

La multiplication des interfaces dans la gouvernance contemporaine de l'intelligence artificielle pourrait laisser penser qu'elles constituent de nouveaux régimes — techniques, éthiques ou juridiques — venant s'ajouter à ceux du sens, de la communication et du métabolisme. Une telle interprétation serait toutefois méthodologiquement erronée.

Les interfaces ne sont pas des régimes, pour une raison fondamentale : elles ne produisent pas de dynamique propre. Là où un régime est structuré par une logique interne irréductible — interprétative pour le sens, fiduciaire pour la communication, entropique pour le métabolisme — une interface n'existe que par la médiation entre ces logiques. [22][24]

Régimes : dynamiques constitutives

Les régimes décrivent des conditions structurelles de possibilité de l'intelligence artificielle, en organisant :

- ce qui peut faire sens et être endossé,
- ce qui peut circuler et faire autorité,
- ce qui peut être soutenu matériellement dans le temps.

Ils sont irréductibles, co-conditionnés et traversés par des tensions internes. Ils peuvent fonctionner, se déséquilibrer ou s'autonomiser, indépendamment de toute volonté de gouvernance explicite.

Interfaces : dispositifs de traduction

Les interfaces, à l'inverse, ne produisent rien par elles-mêmes. Elles traduisent, articulent et mettent en forme des exigences issues des régimes afin qu'elles puissent faire l'objet d'interventions situées.

Une interface :

- ne crée pas de sens, mais conditionne sa reprise,
- ne fonde pas la confiance, mais en oriente l'institution,

- ne produit pas de métabolisme, mais en contraint certaines modalités.

Elle est toujours secondaire, non par importance, mais par statut : elle suppose l'existence préalable des régimes qu'elle articule.

Le risque de la confusion

Confondre interfaces et régimes conduit à deux dérives symétriques.

La première est une inflation analytique : chaque nouvelle couche de régulation ou de médiation est traitée comme une dynamique autonome, au risque de perdre toute lisibilité architectonique.

La seconde est une illusion de gouvernance : en multipliant les interfaces, on donne l'impression de maîtriser les régimes, alors même que leurs déséquilibres structurels persistent. Les dispositifs se superposent, sans prise réelle sur les seuils critiques. [21][27]

Reconnaître que les interfaces ne sont pas des régimes permet au contraire de maintenir une distinction claire entre :

- ce qui fait tenir l'intelligence artefactuelle,
- et ce qui permet d'intervenir sur cette tenue sans la réduire.

Interfaces et non-totalisation

Enfin, le statut non *régimique* des interfaces est une condition essentielle de la non-totalisation de l'architectonique. Si une interface devenait un régime à part entière, elle prétendrait organiser à elle seule gouvernabilité, normativité et soutenabilité — reproduisant exactement les échecs de l'unification que cet ouvrage cherche à dépasser.

Les interfaces ne totalisent pas, elles localisent. Elles n'unifient pas, elles rendent arbitrables. Elles ne résolvent pas les tensions entre régimes, elles en constituent les lieux opératoires, toujours provisoires et toujours révisables. [24]

18.7 Où et comment les interfaces échouent

Si les interfaces constituent des médiations indispensables entre l'analyse architectonique et l'action de gouvernance, elles ne sont ni neutres ni infaillibles. Leur efficacité est toujours conditionnelle, et leurs échecs ne relèvent pas seulement d'erreurs de mise en œuvre. Ils révèlent des limites structurelles, liées à la manière dont les interfaces peuvent se rigidifier, se dissocier des opérateurs ou être mobilisées comme substituts à l'arbitrage. [21][34]

Identifier ces modes d'échec est essentiel, non pour disqualifier les interfaces, mais pour éviter qu'elles ne deviennent elles-mêmes des facteurs de déséquilibre.

18.7.1 Échec par autonomisation

Un premier mode d'échec apparaît lorsque l'interface s'autonomise par rapport à l'opérateur qu'elle est censée activer. Le dispositif continue de fonctionner — audit, procédure, outil technique — mais il n'est plus relié aux seuils qu'il devait rendre visibles.

Ainsi, des mécanismes de traçabilité peuvent subsister sans produire de responsabilité effective ; des dispositifs de signalement peuvent fonctionner sans restaurer la confiance ; des normes environnementales peuvent être respectées formellement sans contenir la dette

entropique réelle. L'interface devient alors auto-référentielle, évaluée pour sa conformité interne plutôt que pour sa capacité à maintenir la co-contrainte des régimes. [34][37]

18.7.2 Échec par rigidification

Les interfaces échouent également lorsqu'elles se rigidifient. Conçues pour rendre des arbitrages praticables dans un contexte donné, elles peuvent devenir incapables de s'adapter à l'évolution des usages, des technologies ou des seuils.

Cette rigidification transforme des prises de gouvernabilité en points de blocage : procédures trop lourdes pour être activées, seuils figés malgré des effets non anticipés, obligations inadaptées à la diversité des contextes. L'interface cesse alors d'être un lieu de traduction pour devenir un obstacle à l'intervention. [35][37]

18.7.3 Échec par substitution

Un troisième mode d'échec apparaît lorsque les interfaces sont mobilisées comme substituts aux arbitrages qu'elles devraient médier. La multiplication des dispositifs donne l'illusion d'une gouvernance exhaustive, alors même que les tensions architectoniques ne sont ni discutées ni assumées.

Dans ce cas, les interfaces servent à déplacer la responsabilité : la décision est attribuée à la procédure, à la métrique ou au cadre normatif, plutôt qu'à un choix explicite entre contraintes hétérogènes. La gouvernance devient procédurale sans être politique. [21][37]

18.7.4 Échec par désarticulation interrégimes

Enfin, les interfaces échouent lorsqu'elles sont conçues pour un seul régime, sans prise sur les autres. Une interface technique optimisée pour la performance peut dégrader la confiance communicationnelle ; une interface juridique focalisée sur la conformité peut accroître la dette organisationnelle ; une interface éthique non traduite peut rester sans effet sur les dispositifs concrets.

L'échec n'est pas ici local, mais architectonique : l'interface renforce un régime au détriment des autres, contribuant à des configurations pathologiques déjà identifiées. [27][34][37]

Portée critique de ces échecs

Ces échecs ne constituent pas des anomalies marginales ; ils sont le revers inévitable d'une gouvernance qui refuse la totalisation. Reconnaître les limites des interfaces permet de maintenir leur statut correct : des prises provisoires, toujours susceptibles d'être réévaluées, ajustées ou désactivées.

C'est à cette condition que les interfaces peuvent rester des instruments de gouvernabilité, et non devenir des couches supplémentaires de complexité ou des écrans masquant les déséquilibres qu'elles étaient censées contenir. Elles ne résolvent pas les tensions entre régimes, mais en constituent les lieux opératoires, rendent praticables des arbitrages sans prétendre les clore, et permettent d'intervenir sans réduire la complexité architectonique de l'intelligence artificielle. Leur efficacité tient moins à leur sophistication qu'à leur capacité à rester articulées aux opérateurs qu'elles médient, et à demeurer révisables face aux déséquilibres qu'elles révèlent.

PARTIE VII — Vers une gouvernance architectonique de l'intelligence artefactuelle

(Partie transitionnelle, non prescriptive)

Le précédent diagnostic des possibles échecs des interfaces permet d'aborder la Partie VII sans illusion solutionniste. Il ne s'agira pas d'appliquer un modèle de gouvernance idéal, mais d'examiner comment des configurations concrètes peuvent maintenir, au cas par cas, la co-contrainte des régimes, en acceptant l'irréversibilité, la conflictualité et la révision permanente comme conditions normales de l'action. [21][34]

19. Gouverner sans unifier

Gouverner l'intelligence artefactuelle ne consiste ni à optimiser des systèmes, ni à imposer un cadre normatif unique, ni à neutraliser les conflits qu'elle suscite. À la lumière des analyses précédentes, la gouvernance apparaît comme une pratique située, orientée vers le maintien précaire de la co-contrainte entre des régimes irréductibles. [20][21]

Cette partie ne propose ni un modèle universel de gouvernance, ni une méthode clé en main. Elle examine des configurations de décision, des agencements institutionnels et des dispositifs d'intervention à partir desquels il devient possible de contenir les effets de désassomption — sans jamais prétendre les éliminer.

Il ne s'agira donc pas de répondre à la question « comment gouverner l'IA en général », mais d'explorer quand, où et jusqu'où des systèmes d'intelligence artefactuelle peuvent être gouvernés, et à quelles conditions il devient nécessaire de les limiter, de les suspendre ou d'y renoncer. Gouverner, dans ce cadre, signifie moins maîtriser que tenir : tenir ensemble le sens, la circulation et la soutenabilité, aussi longtemps que cette tenue demeure politiquement et matériellement assumable. [31][37]

19.1 Pourquoi la gouvernance par couches échoue

Les approches dominantes de la gouvernance de l'IA reposent encore largement sur un modèle par couches. À une infrastructure technique supposée neutre viendraient s'ajouter successivement une couche de sécurité, une couche éthique, une couche juridique, voire une couche sociétale. Chaque couche serait chargée de corriger ou d'encadrer les effets produits par la précédente, sans remettre en cause son fonctionnement interne.

Ce modèle présente un avantage apparent : il permet de distribuer les responsabilités, de spécialiser les expertises et de donner l'impression d'un contrôle progressif. Il échoue pourtant de manière systématique dès lors qu'il s'agit de gouverner des systèmes d'intelligence artefactuelle à grande échelle. [34][37]

Une erreur de diagnostic structurel

L'échec de la gouvernance par couches tient d'abord à une erreur de diagnostic. Elle suppose que les problèmes de l'IA apparaissent *après coup*, comme des effets secondaires d'un fonctionnement technique fondamentalement stable. Or les analyses précédentes ont montré que les crises associées à l'IA — crises du sens, de la confiance, de la soutenabilité — sont structurelles, et non périphériques.

Les régimes du sens, de la communication et du métabolisme n'interviennent pas successivement, ils opèrent simultanément. Ajouter une couche normative ou éthique ne permet donc pas de corriger un déséquilibre qui traverse l'ensemble de l'architectonique.

L'illusion du correctif externe

Le modèle par couches repose sur l'idée qu'il existerait un point extérieur depuis lequel il serait possible de réguler l'IA sans affecter son fonctionnement interne. Cette illusion se manifeste de deux manières :

- soit la couche ajoutée reste symbolique, sans prise réelle sur les dispositifs techniques et organisationnels,
- soit elle est intégrée de manière instrumentale, au prix d'une complexification accrue du système, qui déplace les problèmes au lieu de les résoudre.

Dans les deux cas, la gouvernance devient réactive : elle intervient après la stabilisation des usages, lorsque les marges de manœuvre sont déjà réduites et que les coûts de transformation sont élevés.

Une neutralisation des arbitrages

La gouvernance par couches tend également à neutraliser les arbitrages. Chaque couche traite une dimension du problème selon ses propres critères — conformité juridique, respect de principes éthiques, performance technique — sans que les conflits entre ces critères soient explicitement assumés.

Les décisions apparaissent alors comme des nécessités techniques ou réglementaires, plutôt que comme des choix politiques situés. La responsabilité se dissout dans l'empilement des dispositifs, et la gouvernance perd sa capacité à nommer ce qui est sacrifié, différé ou rendu irréversible. [21][37]

Un renforcement des déséquilibres

Enfin, loin de corriger les déséquilibres architectoniques, la gouvernance par couches tend à les renforcer. Une couche de sécurité peut accroître la dette organisationnelle ; une couche éthique intégrée tardivement peut rigidifier la communication ; une couche juridique mal articulée peut déplacer les coûts vers des acteurs moins visibles.

Loin de produire une gouvernance plus robuste, l'empilement des couches crée des systèmes plus opaques, plus coûteux et plus difficiles à reprendre. [17][35]

Vers une gouvernance architectonique

L'échec de la gouvernance par couches ne signifie pas que les outils qu'elle mobilise soient inutiles. Il indique qu'ils ne peuvent être efficaces qu'à condition d'être repositionnés : non comme des correctifs externes, mais comme des interfaces articulées aux opérateurs et aux régimes. [34]

Gouverner l'intelligence artefactuelle ne consiste pas à empiler des protections, mais à intervenir au niveau des tensions structurantes, là où le sens, la circulation et le métabolisme se contraignent mutuellement. C'est à cette condition que la gouvernance peut cesser d'être un habillage et devenir une pratique réelle de tenue des systèmes. [21]

19.2 Pourquoi les principes abstraits ne suffisent pas

Face aux limites des régulations techniques ou juridiques, la gouvernance de l'intelligence artificielle est souvent réorientée vers des principes éthiques généraux : transparence, équité, responsabilité, non-malfaisance, respect de la dignité humaine. Ces principes jouent un rôle important dans la reconnaissance publique des enjeux et dans la formulation d'exigences normatives partagées. Ils ne suffisent toutefois pas à fonder une gouvernance effective. [31]

Cette insuffisance ne tient pas à leur contenu moral, mais à leur statut abstrait.

L'abstraction comme condition... et comme limite

Les principes éthiques opèrent nécessairement à un haut niveau de généralité. C'est ce qui leur permet d'être reconnus au-delà des contextes particuliers, et de servir de repères transversaux. Mais cette abstraction même les prive de prise directe sur les configurations concrètes de l'intelligence artificielle.

Un principe ne dit pas :

- où intervenir,
- quand un seuil est franchi,
- qui peut décider,
- ni comment arbitrer entre des exigences concurrentes.

En l'absence de médiation, les principes restent déclaratifs. Ils orientent les discours sans structurer l'action.

L'illusion de la suffisance morale

L'invocation de principes abstraits produit souvent une illusion de gouvernance. L'adhésion formelle à des valeurs partagées est interprétée comme un gage de maîtrise, alors même que les dispositifs techniques, communicationnels et métaboliques continuent de produire des effets non assumés. [32]

Cette illusion est renforcée par la convertibilité discursive des principes : ils peuvent être mobilisés pour justifier des choix opposés, sans que leurs conditions d'application soient explicitées. La gouvernance devient alors un exercice de légitimation a posteriori, plutôt qu'un lieu d'arbitrage explicite.

L'incapacité à traiter les conflits

Les principes abstraits échouent également parce qu'ils ne savent pas gérer les conflits. Or l'intelligence artificielle est traversée de tensions irréductibles : entre fluidité et contestabilité, entre performance et soutenabilité, entre automatisation et endossement humain.

Un principe isolé ne permet pas de trancher entre ces exigences. Il tend au contraire à masquer le conflit en le reformulant comme un problème de conformité morale, là où il s'agit d'un arbitrage structurel entre régimes. [31][34]

Le risque de la moralisation sans prise

Lorsque les principes sont mobilisés sans dispositifs de traduction, ils produisent une pathologie spécifique : la moralisation sans prise. Les valeurs sont affirmées avec force, mais les acteurs concrets — développeurs, institutions, usagers — ne disposent pas des moyens pratiques pour les faire valoir, les contester ou les adapter.

La responsabilité est alors déplacée vers les intentions déclarées, plutôt que vers les conditions effectives de production et de circulation des effets. [21]

De la reconnaissance des principes à la gouvernabilité

Reconnaître l'insuffisance des principes abstraits ne revient pas à les disqualifier. Cela implique de les repositionner dans une architectonique de la gouvernance : non comme des fondements auto-suffisants, mais comme des ressources normatives appelant des médiations. [31][37]

Ce n'est qu'à travers des opérateurs capables d'identifier des seuils, et des interfaces capables de traduire des principes en obligations situées, que les valeurs peuvent acquérir une efficacité gouvernante. Sans cette traduction, les principes restent justes — mais impuissants. [34][37]

La section suivante pourra alors examiner ce que signifie gouverner par seuils plutôt que par principes, en montrant comment l'identification de points de bascule permet d'articuler valeurs, contraintes et décisions sans prétendre les unifier.

19.3 Gouverner : maintenir la co-contrainte

À l'issue des analyses précédentes, une définition minimale mais exigeante de la gouvernance de l'intelligence artificielle peut être formulée : gouverner consiste à maintenir la co-contrainte entre des régimes irréductibles, aussi longtemps que cette co-contrainte demeure politiquement, socialement et matériellement assumable. [21]

La co-contrainte n'est pas un état harmonieux : c'est la persistance de points de reprise entre régimes — endossement, contestation, assomption — ; elle se reconnaît au fait que les arbitrages restent identifiables et opposables.

Cette définition rompt avec deux conceptions dominantes. Elle s'oppose d'abord à une vision instrumentale, pour laquelle gouverner reviendrait à optimiser des paramètres ou à corriger des dysfonctionnements locaux, et ensuite à une vision normative abstraite, qui ferait de la gouvernance l'application cohérente de principes généraux indépendamment des configurations concrètes. [34]

Maintenir la co-contrainte ne signifie ni harmoniser les régimes, ni les équilibrer une fois pour toutes. Cela signifie empêcher qu'un régime — sens, communication ou métabolisme — ne s'autonomise au point d'écraser les autres. La gouvernance ne vise pas l'équilibre, mais la tenue.

La gouvernance comme pratique de maintien

Dans cette perspective, gouverner n'est pas un état, mais une pratique continue. La co-contrainte n'est jamais acquise : elle doit être réinstaurée à mesure que les usages se déplacent, que les flux s'intensifient et que les coûts s'accumulent.

Cette pratique implique :

- de rendre visibles les déséquilibres émergents,
- d'identifier les points où un régime cesse de contraindre les autres,
- et d'intervenir avant que ces déséquilibres ne deviennent irréversibles. [34]

La gouvernance ne garantit pas la production d'intelligence ; elle conditionne la possibilité que celle-ci continue d'apparaître sans se retourner contre ses propres conditions d'existence.

Gouverner, c'est accepter l'arbitrage

Maintenir la co-contrainte implique d'assumer des arbitrages explicites. Il n'existe pas de configuration dans laquelle la production de sens, la fluidité de la communication et la soutenabilité métabolique pourraient être maximisées simultanément.

Gouverner, dans ce cadre, consiste à décider :

- jusqu'où le sens peut s'intensifier sans devenir entropiquement prédateur,
- jusqu'où la communication peut s'accélérer sans dissoudre la confiance,
- jusqu'où le métabolisme peut être sollicité sans produire de désassomption. [8][27]

Ces décisions ne relèvent ni d'une pure rationalité technique, ni d'un jugement moral isolé. Elles sont politiques au sens fort : elles engagent une répartition des coûts, des responsabilités et des horizons temporels. [21][23]

Gouvernance et désassomption

Enfin, maintenir la co-contrainte suppose de reconnaître qu'il existe des situations où la gouvernance atteint ses limites. Lorsque les seuils sont durablement franchis, la poursuite des usages ne relève plus d'un ajustement, mais d'une désassomption : les effets continuent d'être produits sans que leurs conditions puissent être assumées collectivement.

Dans ces cas, gouverner peut signifier restreindre, suspendre, voire renoncer. Le renoncement n'est pas ici un échec de la gouvernance, mais l'un de ses actes possibles, lorsque la co-contrainte ne peut plus être tenue sans violence matérielle, sociale ou politique. [8][31]

La définition de la gouvernance comme maintien de la co-contrainte appelle désormais une précision décisive : comment repérer les moments où cette co-contrainte se défait, et selon quels critères intervenir. Le chapitre suivant examinera ces moments critiques en introduisant une analyse systématique des seuils, des arbitrages et des situations de désassomption, à partir desquels la gouvernance cesse d'être une question d'ajustement pour devenir une question de décision.

20. Seuils, arbitrages et désassomption

L'analyse de la gouvernance comme maintien de la co-contrainte conduit à déplacer le regard : il ne s'agit plus d'améliorer indéfiniment des performances, mais d'identifier les moments de bascule à partir desquels une configuration cesse d'être gouvernable. Le chapitre

qui suit examine ces moments critiques en montrant comment la notion de seuil permet de penser l'intervention, l'arbitrage et, le cas échéant, la suspension ou le renoncement. [34]

20.1 Gouverner par seuils, non par optimisation

La plupart des dispositifs contemporains de gouvernance de l'IA restent structurés par une logique d'optimisation. Il s'agit d'améliorer des indicateurs — précision, rapidité, efficacité, conformité — en supposant que la gouvernance consiste à maximiser des performances sous contraintes. Cette logique est adaptée à des systèmes fermés ou faiblement couplés ; elle devient inopérante dès lors que l'on cherche à gouverner des configurations architectoniques traversées par des régimes irréductibles. [35]

Gouverner l'intelligence artefactuelle ne consiste pas à optimiser, mais à éviter des franchissements irréversibles. C'est pourquoi la notion de seuil est centrale.

Le seuil comme point de bascule

Un seuil ne désigne pas un optimum, mais un point de bascule. Il marque le moment où une qualité cesse de tenir : lorsque le sens n'est plus endossable, lorsque la circulation n'est plus orientable, lorsque les coûts ne sont plus assumables. Le franchissement d'un seuil ne signifie pas que le système cesse de fonctionner ; il signifie qu'il cesse d'être intelligible, gouvernable ou soutenable selon les critères de l'architectonique.

À la différence d'un indicateur, un seuil n'est pas nécessairement quantifiable de manière continue. Il est souvent repérable a posteriori, par des effets de rupture, d'irréversibilité ou de désassomption. Gouverner par seuils consiste précisément à anticiper ces ruptures, plutôt qu'à en corriger les effets.

Les limites de l'optimisation continue

L'optimisation suppose que les objectifs puissent être hiérarchisés et traduits dans un espace de calcul commun. Or l'intelligence artefactuelle engage des dimensions hétérogènes : interprétatives, communicationnelles, matérielles et politiques. Chercher à optimiser l'ensemble revient à masquer les conflits entre régimes sous des agrégats artificiels. [31]

Cette logique produit une illusion de maîtrise : les indicateurs s'améliorent localement, tandis que les déséquilibres s'accumulent ailleurs. La gouvernance par optimisation tend ainsi à repousser les seuils jusqu'à ce qu'ils deviennent incontrôlables.

Gouverner par seuils : une logique de décision

Gouverner par seuils implique une autre rationalité. Il ne s'agit pas de déterminer la meilleure configuration possible, mais de décider quand intervenir, jusqu'où accepter, et à partir de quel point renoncer. Cette rationalité est fondamentalement politique : elle engage des jugements sur ce qui peut être assumé collectivement, et sur ce qui ne peut plus l'être. [21]

Les seuils fournissent alors des points d'appui pour la décision. Ils permettent de justifier des mesures qui seraient incompréhensibles dans une logique d'optimisation — ralentissement volontaire, restriction d'usage, suspension d'un système pourtant performant. [34]

Seuils et gouvernabilité minimale

Enfin, gouverner par seuils permet de penser une gouvernabilité minimale. Plutôt que de viser une maîtrise exhaustive, il s'agit de préserver des marges de reprise, de contestation et

de réorientation tant que cela reste possible. Une fois les seuils franchis, ces marges se referment, et la gouvernance se réduit à la gestion des dommages.

Gouverner par seuils, c'est donc accepter que la meilleure décision ne soit pas toujours celle qui maximise un bénéfice immédiat, mais celle qui maintient ouvert l'horizon de l'assomption collective. [31]

20.2 Seuils de réponsabilité, de confiance, de désassomption

Gouverner par seuils suppose de distinguer des points de bascule hétérogènes, propres à chacun des régimes constitutifs de l'intelligence artificielle. Ces seuils ne sont ni équivalents ni synchrones. Ils peuvent être franchis indépendamment les uns des autres, produisant des configurations instables où le système continue de fonctionner tout en cessant d'être gouvernable.

Trois seuils structurants peuvent ainsi être identifiés : le seuil de réponsabilité, le seuil de confiance et le seuil de désassomption.

20.2.1 Le seuil de réponsabilité (régime du sens)

Le seuil de réponsabilité est franchi lorsqu'il n'est plus possible de répondre raisonnablement des productions de sens au point d'usage. Il ne s'agit pas d'une impossibilité d'explication exhaustive, mais d'une incapacité pratique à endosser, contester ou corriger les effets produits. [4][18][39]

Ce seuil est atteint lorsque :

- les sources, transformations ou limites d'un énoncé ne peuvent plus être retracées,
- aucun acteur humain ou institutionnel n'est en mesure d'assumer les conséquences d'un usage,
- les dispositifs de reprise ou de suspension sont inexistants ou inopérants.

Le franchissement du seuil de réponsabilité ne rend pas le système muet ou inefficace ; il le rend irresponsable au sens fort. Le sens continue de circuler, mais il n'est plus assignable. À ce stade, la gouvernance ne peut plus consister en un simple ajustement : elle implique une réduction du périmètre d'usage, une reprise humaine obligatoire ou une suspension ciblée.

20.2.2 Le seuil de confiance (régime de la communication)

Le seuil de confiance est franchi lorsque la circulation des énoncés excède la capacité collective à orienter la confiance. La communication demeure fluide, mais elle cesse d'être discriminante : il devient difficile de distinguer ce qui mérite crédit, discussion ou suspension. [27][28][38]

Ce seuil est atteint lorsque :

- la densité et la vitesse des flux rendent la contestation impraticable,
- l'autorité devient diffuse, cumulative et opaque,
- les mécanismes de reprise, de correction ou de hiérarchisation sont saturés.

Le franchissement du seuil de confiance ne produit pas un effondrement communicationnel, mais une dissolution de l'orientation. Les énoncés ne sont ni crus ni réfutés, ils s'imposent par défaut, par fatigue ou par automatisme. Dans ce contexte, gouverner peut signifier créer des zones protégées, ralentir la circulation ou restaurer des espaces à priorité humaine. [27]

20.2.3 Le seuil de désassomption (régime métabolique)

Le seuil de désassomption est franchi lorsque les coûts métaboliques cumulés — énergétiques, matériels, organisationnels, humains — dépassent ce qui peut être assumé collectivement dans la durée. À la différence des deux autres seuils, il est marqué par une irréversibilité structurelle : ce qui a été dissipé, usé, extrait ou rigidifié ne redevient disponible qu'au prix d'un surcoût. [8][40]

Ce seuil est atteint lorsque :

- la dette entropique accumulée ne peut plus être compensée ou redistribuée,
- les arbitrages deviennent contraints par des pénuries, des conflits d'usage ou des limites matérielles,
- les décisions sont imposées par le bas, sans possibilité de reprise politique.

Le franchissement du seuil de désassomption ne suspend pas nécessairement le système ; il en déplace la charge vers des territoires, des institutions ou des acteurs moins visibles. Gouverner, à ce stade, peut impliquer des décisions radicales : simplification, ralentissement volontaire, abandon de certains usages ou interdiction contextuelle. [31][37]

Asymétrie et articulation des seuils

Ces trois seuils ne se superposent pas. Un système peut perdre sa responsabilité sans épuiser sa confiance, ou saturer la confiance sans atteindre la désassomption. Inversement, un franchissement métabolique peut contraindre brutalement les deux autres régimes.

C'est précisément cette asymétrie qui rend la gouvernance complexe : elle oblige à arbitrer dans des configurations où certains régimes continuent de fonctionner tandis que d'autres se défont.

Identifier ces seuils ne permet pas d'éviter tous les franchissements. Cela permet en revanche de nommer les moments où la gouvernance change de nature : d'une pratique d'ajustement, elle devient une pratique de décision, engageant des renoncements explicites et des responsabilités politiques. [34]

20.3 Conflictualité constitutive de la gouvernance

Gouverner l'intelligence artificielle ne consiste pas à résoudre des conflits, mais à les instituer comme tels. Les analyses précédentes ont montré que les tensions entre régimes — entre sens, communication et métabolisme — ne sont ni accidentelles ni transitoires. Elles sont constitutives de l'architecture même de l'intelligence artificielle. [21]

Toute prétention à une gouvernance consensuelle ou harmonieuse repose sur une méconnaissance de cette structure conflictuelle.

Le conflit comme condition, non comme échec

Dans une perspective architectonique, le conflit n'est pas le signe d'une gouvernance défaillante. Il est au contraire l'indice que les arbitrages sont rendus visibles et qu'ils ne sont pas dissimulés sous des métriques ou des principes abstraits.

Gouverner, c'est accepter que :

- l'intensification du sens se fasse au détriment de la soutenabilité,
- l'accélération de la communication fragilise la confiance,
- la réduction des coûts métaboliques limite certaines capacités interprétatives.

Ces tensions ne peuvent être abolies ; elles peuvent seulement être tenues, provisoirement, par des décisions situées. [31]

La conflictualité comme enjeu politique

Parce qu'elle engage des choix irréversibles et des répartitions de charges, la gouvernance de l'intelligence artificielle est intrinsèquement politique. Les arbitrages ne relèvent pas d'une optimisation neutre, mais d'une hiérarchisation des priorités, des horizons temporels et des populations affectées. [21][23]

Refuser de reconnaître cette conflictualité conduit à une dépolitisation de la gouvernance, où les décisions apparaissent comme des nécessités techniques ou réglementaires, et non comme des choix discutables.

À l'inverse, reconnaître la conflictualité permet de maintenir ouverte la question de la responsabilité : qui décide, pour qui et pour quoi. [37]

Les risques de la pacification procédurale

Un mode d'échec fréquent de la gouvernance consiste à pacifier artificiellement les conflits par des procédures. Les désaccords sont traduits en non-conformités, les arbitrages en exigences formelles, et les renoncements en ajustements techniques.

Cette pacification a un coût : elle déplace les conflits hors de l'espace public, les rendant plus difficiles à contester. La gouvernance devient alors un dispositif de gestion des apparences, plutôt qu'un lieu de décision assumée. [35]

Gouverner, c'est choisir dans l'incertitude

Enfin, la conflictualité constitutive de la gouvernance est inséparable de l'incertitude. Les effets de long terme, notamment métaboliques, ne sont jamais entièrement prévisibles. Les décisions doivent être prises sans garantie de succès, avec la possibilité d'erreurs irréversibles. [8]

Gouverner, dans ce cadre, ne consiste pas à promettre des résultats, mais à assumer la responsabilité de choix imparfaits, révisables tant que les seuils ne sont pas franchis. [34]

Encadré 7 (empirique) — Seuil métabolique non anticipé et rattrapage institutionnel

En [janvier 2026](#), le gouvernement britannique a reconnu publiquement avoir commis une « grave erreur » dans l'autorisation d'un centre de données *hyperscale*, après qu'une procédure judiciaire ait mis en évidence l'insuffisance de l'évaluation de ses impacts environnementaux. L'erreur ne portait pas sur la légalité formelle de la décision, mais sur l'absence d'engagements contraignants relatifs à la consommation énergétique et aux effets climatiques de l'infrastructure, pourtant connus à l'avance.

Ce cas est éclairant du point de vue de la gouvernance architectonique. Il ne révèle pas une ignorance des coûts métaboliques, mais une incapacité institutionnelle à les intégrer comme seuils décisionnels avant l'autorisation. Les procédures existaient, le cadre juridique était en place, mais l'interface juridico-normative n'était pas articulée au régime métabolique de manière opérante. La décision a ainsi été prise dans un contexte où la conflictualité — entre développement infrastructurel, soutenabilité énergétique et acceptabilité politique — a été neutralisée plutôt qu'assumée.

La reconnaissance a posteriori de l'erreur illustre une situation typique de **désassomption** partielle : une fois les investissements engagés et les infrastructures planifiées, les marges de reprise se trouvent fortement réduites. La gouvernance ne disparaît pas, mais elle change de nature : elle ne consiste plus à arbitrer entre des options ouvertes, mais à gérer les conséquences d'un seuil franchi sans avoir été reconnu comme tel.

Ce type de rattrapage institutionnel confirme que la conflictualité ne peut être évitée par des procédures seules. Lorsqu'elle n'est pas intégrée en amont comme condition normale de la décision, elle réapparaît sous la forme de contentieux, de blocages politiques ou d'aveux tardifs — c'est-à-dire au moment où la capacité de gouverner est déjà entamée.

Par ailleurs, toujours dans le cadre du présent essai, la perspective d'une intelligence artificielle dite « superintelligente » (ASI) pourrait être envisagée comme une configuration-limite où cessent les conditions mêmes de la gouvernabilité architectonique. Une analyse spécifique de cette hypothèse est proposée en **annexe I**.

21. Méthodologies possibles (sans promesse)

Reconnaître la conflictualité constitutive de la gouvernance conduit à une conséquence méthodologique décisive : il n'existe pas de méthode universelle capable de résoudre à l'avance les arbitrages que l'intelligence artefactuelle impose. Le chapitre suivant n'exposera donc pas un modèle normatif, mais une série de méthodologies possibles, conçues comme des cadres d'intervention partiels, situés et révisables — sans promesse de maîtrise, mais avec l'exigence de rendre les décisions discutables et assumables. [31][34]

21.1 Diagnostic interrégime

La première méthodologie de gouvernance possible consiste en un diagnostic interrégime. Il ne s'agit ni d'une évaluation globale, ni d'un audit de conformité, ni d'une mesure de performance. Le diagnostic interrégime vise à rendre visibles les déséquilibres entre les régimes du sens, de la communication et du métabolisme, afin d'orienter des décisions situées. [21]

Cette méthodologie repose sur une hypothèse minimale : avant de décider comment intervenir, il faut identifier où la co-contrainte se défait, quel régime domine, et quels seuils sont proches d'être franchis.

Précisons également que le diagnostic interrégime n'a pas vocation à produire un diagnostic centralisé. Il est conçu comme un format commun de description et de mise en discussion, destiné à être appliqué par une pluralité d'acteurs hétérogènes (acteurs privés, organismes

publics, pays...). Son enjeu n'est pas l'unification des décisions, mais l'interopérabilité des arbitrages : rendre comparables, opposables et révisables des diagnostics situés, à travers des contextes institutionnels et géopolitiques différents.

Finalité du diagnostic interrégime

Le diagnostic interrégime ne cherche pas à produire un verdict. Il vise à :

- repérer les asymétries entre régimes,
- identifier les zones de fragilité ou de saturation,
- expliciter les arbitrages déjà à l'œuvre, souvent de manière implicite.

Il ne dit pas ce qu'il faut faire ; il rend la décision possible en rendant les tensions explicites.

Structure du diagnostic

Un diagnostic interrégime examine systématiquement trois questions, correspondant aux trois régimes :

- Régime du sens
 - Les productions peuvent-elles être endossées au point d'usage ?
 - Existe-t-il des dispositifs effectifs de reprise, de contestation et de suspension ? [4][18][39]
- Régime de la communication
 - La circulation des énoncés est-elle orientable ?
 - Des formes d'autorité interprétative sont-elles identifiables, discutables et révisables ? [27][28][38]
- Régime métabolique
 - Les coûts cumulés sont-ils identifiables et assumables dans le temps ?
 - La dette entropique est-elle reconnue, située et prise en compte dans les décisions ? [8][32][40]

Ces questions ne visent pas une symétrie parfaite. Un régime peut être temporairement dominant, à condition que les autres conservent une capacité de contrainte.

Indices de déséquilibre

Le diagnostic interrégime s'appuie moins sur des indicateurs quantitatifs que sur des indices qualitatifs de déséquilibre, tels que :

- l'accélération des usages sans dispositifs d'endossement correspondants,
- la multiplication des flux sans espaces de reprise ou de hiérarchisation,
- la croissance des coûts invisibilisés ou externalisés. [29][32]

Ces indices signalent des trajectoires de risque, non des défaillances ponctuelles.

Portée et limites

Le diagnostic interrégime ne garantit ni la justesse des décisions prises, ni l'évitement des franchissements de seuils. Il constitue une prise méthodologique minimale pour éviter deux écueils majeurs : intervenir trop tard, ou intervenir à l'aveugle. [34]

Il n'a de valeur que s'il est :

- contestable,
- révisable,
- et articulé à des interfaces capables de traduire ses résultats en décisions effectives. [35]

Lien avec les outils pratiques

Cette méthodologie peut être mise en œuvre à l'aide d'outils visuels, de matrices ou de grilles de lecture, sans que ceux-ci ne deviennent des instruments de normalisation. L'**annexe B** propose à ce titre des gabarits opératoires, destinés à soutenir le diagnostic sans en figer l'interprétation. [34]

Le diagnostic interrégime n'est pas une solution. Il est un point de départ : celui à partir duquel la gouvernance cesse d'être une abstraction pour devenir une pratique située de discernement.

21.2 Traduction par interfaces

Le diagnostic interrégime ne produit pas en lui-même de décision. Il identifie des déséquilibres, rend visibles des seuils proches et explicite des arbitrages implicites. Pour devenir gouvernant, il doit être traduit en interventions situées. C'est le rôle des interfaces. [35]

La traduction par interfaces constitue ainsi une seconde méthodologie possible : elle ne vise pas à résoudre les tensions entre régimes, mais à les rendre opérables sans les neutraliser.

De l'analyse à l'intervention

La traduction par interfaces intervient lorsque la question n'est plus seulement *où* se situe le déséquilibre, mais *comment* intervenir sans aggraver la situation. Chaque interface correspond à un type de prise spécifique :

- l'interface technique rend certaines contraintes calculables,
- l'interface éthique explicite les valeurs engagées,
- l'interface juridico-normative rend certaines exigences opposables,
- l'interface éthico-normative articule principes et obligations situées. [31][37]

La traduction ne consiste pas à appliquer mécaniquement une solution préexistante, mais à composer une configuration d'interfaces ajustée au déséquilibre identifié.

Rendre opérables les tensions

Toute traduction comporte un risque de simplification : une interface peut devenir un substitut, un régime unique de décision, ou un écran procédural. Dans une gouvernance architectonique, ce risque est contenu par deux exigences :

- aucune interface ne doit prétendre résoudre à elle seule un déséquilibre interrégime,
- toute interface doit rester réversible et contestable. [34]

Ainsi, un déséquilibre du régime du sens ne peut être traité uniquement par une interface technique, pas plus qu'une saturation fiduciaire ne peut être corrigée par une obligation juridique isolée. L'intervention exige des agencements — et la vigilance constante contre les substitutions.

Interfaces et temporalité

La traduction par interfaces est toujours temporellement située. Certaines interfaces visent des interventions immédiates (limitation d'usage, suspension), d'autres des transformations plus lentes (réorganisation, refonte des responsabilités, ajustement des infrastructures). [35]

Gouverner consiste alors à articuler des temporalités hétérogènes : agir vite pour contenir un seuil imminent, tout en réouvrant des marges de reprise à moyen terme.

Responsabilité de la traduction

La traduction par interfaces n'est jamais neutre. Choisir une interface plutôt qu'une autre, ou privilégier un certain agencement, engage une responsabilité politique et organisationnelle. Ces choix déterminent ce qui est rendu visible, ce qui est contraint, et ce qui est laissé hors champ. [21]

C'est pourquoi la traduction par interfaces doit rester explicite : les raisons de l'intervention, les valeurs mobilisées et les seuils visés doivent pouvoir être discutés et contestés. [31]

Portée et limites

La traduction par interfaces ne garantit ni la réussite de l'intervention, ni la stabilité durable de la configuration obtenue. Elle offre une prise pragmatique sur des situations complexes, sans prétendre les maîtriser. [34]

Elle permet surtout d'éviter deux écueils : l'inaction par abstraction et l'action aveugle par automatisme. En ce sens, elle constitue une condition minimale pour qu'une gouvernance architectonique puisse s'exercer.

Des exemples du fonctionnement des interfaces sont donnés à l'**annexe A**.

21.3 Décision située et réversibilité

Les méthodologies de gouvernance présentées jusqu'ici convergent vers une même exigence : la décision ne peut être ni automatique ni définitive. Gouverner l'intelligence artificielle implique des décisions situées, prises dans des contextes spécifiques, et maintenues sous le régime de la réversibilité tant que les seuils ne sont pas irréversiblement franchis. [34]

La décision comme acte situé

Une décision de gouvernance n'est jamais abstraite. Elle intervient à un moment donné, dans une configuration donnée, avec des acteurs identifiables et des contraintes matérielles précises. Elle engage des arbitrages qui ne peuvent être entièrement délégués à des procédures, des indicateurs ou des principes généraux. [31]

Décider, dans ce cadre, consiste à expliciter :

- le déséquilibre que l'on cherche à contenir,
- le ou les seuils que l'on vise à éviter,
- les régimes prioritairement contraints,
- et les coûts que l'on accepte de rendre visibles et assumables.

La décision située ne prétend pas à l'universalité. Elle revendique sa contextualité comme condition de responsabilité. [21]

Réversibilité comme exigence de gouvernance

La réversibilité n'est pas une garantie de retour à l'état antérieur ; elle est une exigence procédurale. Elle vise à maintenir ouvertes des marges de reprise, de correction et de suspension tant que les conditions matérielles et institutionnelles le permettent. [34]

Une décision est réversible lorsque :

- ses effets peuvent être évalués dans le temps,
- des dispositifs de suspension ou de modification sont prévus,
- les responsabilités restent assignables.

À l'inverse, une décision devient irréversible lorsqu'elle engage des franchissements de seuils — notamment métaboliques — qui ferment durablement l'horizon de l'assomption collective. [8]

Les limites de la réversibilité

Toutes les décisions ne peuvent pas être maintenues sous le régime de la réversibilité. La dette entropique introduit une asymétrie temporelle : certains coûts, une fois engagés, ne peuvent être annulés. Reconnaître cette limite est une condition de lucidité, non un renoncement à gouverner. [8][11]

La gouvernance consiste alors à hiérarchiser les décisions selon leur degré d'irréversibilité, et à concentrer les exigences de réversibilité là où elles ont encore une prise réelle.

Décider sans promettre

Enfin, la décision située et réversible implique un refus explicite de la promesse. Gouverner l'intelligence artefactuelle ne consiste pas à garantir des résultats, mais à assumer des choix sous incertitude, en maintenant la possibilité de révision tant que cela reste possible. [34]

Cette posture n'est ni minimaliste ni défensive. Elle reconnaît que la gouvernance, dans un monde de systèmes complexes et irréversibles, ne peut être qu'une pratique prudente de maintien, attentive aux seuils, aux conflits et aux temporalités différées.

21.4 Pourquoi aucune méthodologie n'est universelle

Les méthodologies de gouvernance présentées dans ce chapitre ne constituent ni un ensemble cumulatif, ni un cadre exhaustif. Elles décrivent des prises possibles, ajustées à des configurations situées de l'intelligence artefactuelle. Leur diversité n'est pas un défaut : elle est la conséquence directe de l'architectonique exposée dans cet ouvrage. [34]

Aucune méthodologie ne peut être universelle, pour une raison fondamentale : l'intelligence artefactuelle n'est pas un objet homogène, mais un effet conditionnel de la co-contrainte entre des régimes irréductibles, historiquement et matériellement situés. [21]

L'hétérogénéité des configurations

Les systèmes d'intelligence artefactuelle varient selon leurs usages, leurs contextes institutionnels, leurs infrastructures, leurs publics et leurs horizons temporels. Une méthodologie pertinente dans un contexte donné — par exemple un usage exploratoire à faible criticité — peut devenir inopérante, voire dangereuse, dans un autre — décision juridique, allocation de ressources, gouvernance politique. [37]

La prétention à l'universalité conduit alors à une mauvaise abstraction, qui gomme les différences au lieu de les rendre gouvernables. [2]

L'irréductibilité des régimes

Aucune méthodologie ne peut totaliser les régimes du sens, de la communication et du métabolisme. Toute tentative de les unifier sous une procédure unique reconduit les échecs déjà identifiés : sur-optimisation, juridisme procédural, moralisation sans prise. [31][35]

Une méthodologie n'est jamais qu'un point de vue opératoire, orienté par un certain rapport aux seuils, aux arbitrages et aux temporalités. Elle ne peut donc être que partielle, révisable et contestable. [34]

La conflictualité et l'incertitude

Enfin, la gouvernance de l'intelligence artefactuelle est traversée par une double indétermination : la conflictualité des intérêts et l'incertitude des effets à long terme. Aucun cadre méthodologique ne peut anticiper l'ensemble des conséquences ni absorber la pluralité des valeurs en jeu. [8][21]

Reconnaître cette limite n'est pas un aveu d'impuissance, mais une condition de responsabilité. La gouvernance ne consiste pas à appliquer une méthode, mais à choisir parmi des méthodes, en connaissance de leurs angles morts. [34]

Une posture plutôt qu'une méthode

Ce chapitre ne propose donc pas une solution, mais une posture de gouvernance. Une posture attentive aux seuils, aux déséquilibres et aux irréversibilités ; une posture qui privilégie la décision située à la promesse universelle ; une posture qui accepte la révision, la suspension et le renoncement comme des actes légitimes. [31][34]

C'est à cette condition que les méthodologies peuvent rester des instruments de discernement, et non devenir de nouveaux dogmes.

Le présent essai peut désormais se refermer sur son geste initial. Il ne s'agissait pas de définir ce qu'est l'intelligence artificielle, ni de prescrire un modèle de gouvernance idéal, mais de déplacer le cadre à partir duquel l'intelligence artefactuelle devient pensable, critiquable et gouvernable. La conclusion générale reviendra sur ce déplacement, en en tirant les conséquences philosophiques, politiques et pratiques — sans refermer les questions qu'il a ouvertes. [21]

Conclusion générale

Tenir ensemble sans totaliser

Cet ouvrage est parti d'un constat simple et pourtant décisif : ce que l'on appelle communément « intelligence artificielle » ne constitue pas un objet unifié, ni une faculté identifiable, ni un simple artefact technique. Les difficultés persistantes à la définir, à l'évaluer et à la gouverner ne relèvent pas d'un manque de maturité conceptuelle ou réglementaire ; elles signalent une erreur de perspective. On a cherché à stabiliser l'objet avant d'avoir compris les conditions sous lesquelles il apparaît comme intelligible, assumable et gouvernable. [2][34]

Le déplacement proposé par ce livre a consisté à abandonner l'idée d'une intelligence comme propriété — de machines, de modèles ou de systèmes — pour la penser comme un effet architectonique. Un effet toujours conditionnel, fragile et réversible, qui n'apparaît que lorsque des régimes hétérogènes parviennent, provisoirement, à se limiter mutuellement sans qu'aucun ne domine. [3][21]

Trois régimes irréductibles

L'identification de trois régimes — sens, communication, métabolisme — a permis de rendre lisible cette conditionnalité. Chacun de ces régimes obéit à une logique propre, irréductible aux deux autres. Le régime du sens concerne les conditions d'endossement et de responsabilité des productions au point d'usage. Le régime de la communication organise la circulation des énoncés, les formes d'autorité distribuée et les économies de la confiance. Le régime métabolique soutient matériellement l'ensemble, en introduisant des flux, des coûts et des seuils irréversibles. [8][18][27][38][39][40]

Aucun de ces régimes ne fonde l'intelligence à lui seul. Pris isolément, chacun peut continuer à fonctionner tout en produisant des effets pathologiques : du sens sans soutenabilité, de la communication sans orientation, un métabolisme sans signification. L'intelligence n'apparaît que dans l'intervalle étroit où ces régimes se contraignent mutuellement, en maintenant ouverte la possibilité d'une reprise, d'une contestation et d'une assomption collective. [28][34]

Une architectonique non totalisante

L'architectonique proposée ne vise ni l'unification, ni la synthèse, ni la clôture. Elle assume au contraire une non-totalisation méthodologique. Penser en régimes, c'est accepter que

l'objet ne se laisse pas ramener à un principe unique, qu'il soit technique, éthique ou juridique. C'est aussi refuser les modèles par couches, qui prétendent corriger après coup des déséquilibres structurels, de même que les approches par le biais de principes abstraits, qui peinent à se traduire en décisions effectives. [31][35]

Ce refus de la totalisation n'est pas une faiblesse analytique. Il est une condition de lucidité. Il permet de maintenir visibles les tensions, les arbitrages et les irréversibilités qui traversent l'intelligence artefactuelle contemporaine, au lieu de les masquer sous des métriques, des standards ou des promesses de maîtrise. [11][37]

Gouverner sans promettre

La réflexion sur la gouvernance a prolongé ce geste. Gouverner l'intelligence artefactuelle ne consiste pas à optimiser des systèmes ni à appliquer un cadre normatif universel. Cela consiste à maintenir la co-contrainte entre des régimes irréductibles, aussi longtemps que cette co-contrainte demeure politiquement et matériellement assumable.

Cette gouvernance est nécessairement conflictuelle. Elle implique des arbitrages explicites entre des exigences incompatibles, et l'acceptation de décisions imparfaites prises sous incertitude. Elle ne peut être fondée sur une méthode unique, ni sur une promesse de contrôle. Les méthodologies proposées — diagnostic interrégime, traduction par interfaces, décision située et réversibilité — ne constituent pas des solutions universelles, mais des prises provisoires pour éviter l'aveuglement et l'inaction. [21][34]

Gouverner, dans ce cadre, signifie parfois ralentir, restreindre, suspendre ou renoncer. Le renoncement n'est pas l'échec de la gouvernance, mais l'un de ses actes possibles lorsque les seuils de répondabilité, de fiducie ou de désassomption sont franchis. [8][34]

Interfaces, seuils et responsabilité

L'introduction des opérateurs et des interfaces a permis de relier l'analyse conceptuelle à l'action sans réduire l'une à l'autre. Les opérateurs identifient les points où un régime se défait ; les interfaces déterminent comment intervenir sans neutraliser la complexité. Ni les uns ni les autres ne totalisent l'architectonique. Ils rendent la gouvernabilité localement praticable, tout en laissant subsister l'incertitude et la conflictualité. [34][35]

Cette approche déplace également la question de la responsabilité. Il ne s'agit plus de rechercher un sujet souverain capable de répondre de l'ensemble, mais d'assigner des responsabilités situées, liées à des décisions précises, prises à des moments déterminés, avec des conséquences irréversibles. La responsabilité devient inséparable de la reconnaissance des seuils et de la capacité à en assumer le franchissement — ou le refus. [8][31]

Tenir ensemble

Tenir ensemble sans totaliser : telle est la thèse qui traverse cet ouvrage. Tenir ensemble des régimes hétérogènes sans les réduire à un principe commun. Tenir ensemble analyse et action sans les confondre. Tenir ensemble l'exigence critique et la nécessité de décider. [2][21]

Cette posture n'offre ni réassurance, ni clôture. Elle propose un cadre pour penser et gouverner l'intelligence artefactuelle comme un phénomène situé, historiquement contingent et fondamentalement réversible. À une époque où la tentation est grande de

confondre puissance technique et intelligence, optimisation et gouvernance, promesse et responsabilité, elle invite à un geste plus exigeant : refuser la totalisation, pour préserver la possibilité même de l'intelligence partagée. [3][31]

*

Postface

Après la tenue

Ce livre s'achève sans résoudre. Il ne clôt pas un débat, n'offre pas une doctrine, ne propose pas une méthode stabilisée de gouvernance. Il laisse au contraire ouvertes les tensions qui rendent l'intelligence artefactuelle à la fois possible, problématique et fragile.

Si une postface est nécessaire, c'est parce que l'architecture proposée ici n'est pas un modèle explicatif supplémentaire. Elle constitue un déplacement. Elle invite à quitter l'horizon d'une intelligence conçue comme propriété — propriété d'une machine, d'un modèle ou d'une infrastructure — pour la reconnaître comme un effet conditionnel : un effet qui tient seulement lorsque des régimes hétérogènes continuent de se contraindre mutuellement sans se dissoudre.

Dans ce cadre, la question décisive n'est plus : « *que peut faire l'IA ?* », mais : « *jusqu'où une intelligence ainsi produite demeure-t-elle endossable, orientable et soutenable ? à quel moment cesse-t-elle de l'être ?* »

Penser en régimes revient à admettre que l'intelligence artefactuelle n'est pas un objet homogène mais une tenue instable entre des conditions interprétatives (sens), fiduciaires (communication) et entropiques (métabolisme). Gouverner ne signifie alors ni optimiser, ni moraliser, ni réguler après coup : gouverner signifie maintenir, localement et provisoirement, des points de reprise entre ces régimes — et reconnaître le moment où ces points se ferment.

Conséquence politique majeure : la responsabilité s'en trouve déplacée, non plus vers un sujet souverain censé répondre de tout, mais vers des décisions situées, contestables, assumables, prises sous contrainte irréversible. Plutôt qu'un projet de maîtrise, la gouvernance devient une pratique de seuils.

Cependant, rien n'autorise le désespoir, pas plus que la fascination. Refuser la totalisation ne signifie pas renoncer à penser, mais refuser les promesses de solution qui neutralisent les tensions au lieu de les rendre gouvernables. Le refus de l'unification est une condition de lucidité, en préservant la possibilité même d'arbitrer.

Ce livre n'a pas été pensé pour fournir un cadre définitif, mais pour rendre possible un langage commun : nommer les régimes, reconnaître les seuils, distinguer les interfaces, assumer les renoncements.

Dans un moment historique où l'intelligence se confond trop souvent avec la puissance, la gouvernance avec la conformité, l'enjeu n'est pas de promettre plus, mais de tenir mieux : tenir ensemble sans totaliser, et gouverner sans promettre ce qui ne peut être assumé.

Depuis l'écriture des premiers volumes, un fait s'est imposé avec une netteté croissante : la vitesse d'itération des dispositifs, la densité des flux et l'accumulation des infrastructures déplacent sans cesse les seuils. La gouvernance ne peut donc pas être un état, ni même un « cadre » figé une fois pour toutes ; elle devient une activité continue de reprise, d'ajustement et, parfois, de retrait.

Une autre évidence s'est imposée : à l'ère des modèles, l'acte même d'écrire et de penser devient un objet métabolique. Les textes prolifèrent, les synthèses s'enchaînent à grande vitesse, la surface de la pensée semble s'élever, un « gain » apparent qui repose sur une chaîne invisible de coûts, de dépendances et d'infrastructures humainement imperceptibles. L'architectonique nous contraint alors à un retournement critique : interroger les conditions matérielles de notre propre production intellectuelle. Quels flux, quelles dettes entropiques situées, quelles externalités soutiennent cette apparente accélération de la pensée ?

Avec un horizon-limite qui se profile : celui de systèmes dont la puissance rendrait structurellement impraticables les points de reprise humains. L'hypothèse d'une ASI (la soi-disant *superintelligence*), quelle que soit sa plausibilité, fonctionne comme un révélateur. Elle expose à nu ce que l'architectonique affirme déjà pour les dispositifs actuels : lorsque la réponsabilité devient impraticable, lorsque la confiance devient forcée, lorsque la dette entropique devient non assumable, ce n'est pas seulement la « qualité » de l'IA qui est en jeu, mais la continuité d'une gouvernabilité humaine.

Elle appelle aussi une modestie institutionnelle : multiplier les lieux de discussion, outiller la contestation, documenter les arbitrages, et accepter la pluralité des rythmes. Gouverner, ici, c'est surtout empêcher que l'accélération technique ne devienne une fatalité au prix d'une abdication collective.

Ce livre ne conclut pas. Tenir les tensions ouvertes, refuser les récits qui les effacent ou les harmonisent, préserver la réversibilité aussi longtemps qu'elle existe, admettre que renoncer peut s'avérer être un geste de lucidité plutôt qu'une capitulation. Là où notre époque exige des certitudes définitives, l'architectonique offre autre chose : des réflexions concrètes, précaires et assumées.

Annexes

Les annexes qui suivent proposent des approfondissements ponctuels sur certains dispositifs ou cadres de gouvernance fréquemment mobilisés dans l'ouvrage. Elles ne constituent pas des chapitres supplémentaires, mais des analyses latérales utilisables en complément.

Annexe A — Opérateurs et interfaces (fondation conceptuelle)

Statut de l'annexe

Cette annexe explicite deux notions méthodologiques centrales mobilisées tout au long de l'ouvrage : **les opérateurs** et **les interfaces**.

Elle vise à clarifier leur rôle respectif, leur articulation, et leur fonction dans une pensée de la gouvernabilité qui refuse à la fois la réduction techniciste et l'abstraction morale. [34]

A.1 Pourquoi introduire des opérateurs

Les analyses contemporaines de l'IA mobilisent abondamment des notions descriptives (performance, biais, impact, coût, risque). Ces notions sont utiles, mais insuffisantes pour structurer une pensée systématique de la gouvernabilité.

Un opérateur n'est ni un simple concept descriptif, ni un principe normatif général.

Il désigne une fonction de structuration qui permet :

- d'unifier un ensemble de phénomènes hétérogènes,
- d'identifier des seuils critiques,
- de relier analyse descriptive et décision pratique.

Autrement dit, un opérateur fait travailler le régime auquel il appartient : il en révèle les tensions internes, les limites, et les points de bascule. [3]

Test minimal d'un opérateur (critère de choix)

On peut tester un candidat-opérateur par trois questions simples :

1. Sans lui, le régime reste-t-il opérant (ou s'effondre-t-il en simple agrégat descriptif) ?
2. Permet-il d'identifier des seuils et des pathologies spécifiques au régime ?
3. Est-il irréductible aux opérateurs des deux autres régimes (sans devenir une « valeur » ou un « super-indicateur ») ? [21]

A.2 Opérateur principal et opérateurs secondaires

Chaque régime est structuré autour d'au moins un opérateur principal, auquel peuvent s'adjoindre des opérateurs secondaires.

Rôle de l'opérateur principal

Un opérateur principal :

- condense les enjeux centraux du régime,
- permet d'identifier les pathologies spécifiques,
- sert de point d'articulation avec les autres régimes.

Il ne totalise pas le régime, mais en constitue le point de gravité analytique. [21]

Les opérateurs des trois régimes : [38][39][40]

Régime	Opérateur principal	Fonction
Sens	Répondabilité	Condition d'endossement du sens au point d'usage (question, reprise, correction) [18][39]

Régime	Opérateur principal	Fonction
Communication	Fiducie (<i>autorité comme cristallisation possible</i>)	Condition de circulation opposable des énoncés et de stabilité des chaînes de reprise [27][38]
Métabolisme	Dette entropique située	Condition matérielle et temporelle de soutenabilité (irréversibilités et distribution des coûts) [8][40]

Ces opérateurs ne sont pas symétriques : chacun est ajusté à la logique propre de son régime, tout en restant articulable aux autres. [34]

Note sur « autorité » (régime communicationnel)

Dans le régime communicationnel, l'autorité n'est pas l'opérateur, mais une forme stabilisée (souvent distribuée) de la fiducie : elle canalise à qui/à quoi l'on se remet. Lorsque la fiducie se fragilise, l'autorité tend à se multiplier, se mimer ou se déplacer (plutôt qu'à simplement « disparaître »). [27][28]

A.3 L'entropie comme opérateur limite (et non totalisant)

L'entropie joue un rôle particulier dans l'architectonique : elle constitue un opérateur limite.

Cela implique au moins trois conséquences structurantes :

1) Elle introduit l'irréversibilité temporelle

Contrairement au sens et à la communication, le métabolisme est soumis à une flèche du temps non négociable. La dette entropique marque ce qui ne peut être annulé a posteriori, et ce dont le « rattrapage » devient structurellement coûteux. [8][11]

2) Elle ne subsume pas les autres régimes

L'entropie n'explique ni le sens ni la communication. Elle en contraint la durée, la portée et l'intensité, sans fournir à elle seule les critères de décision. [3][8]

3) Elle politise la question de la répartition

La dette entropique n'est jamais abstraite : elle se dépose quelque part, sur quelqu'un, à un moment donné. Elle transforme toute décision technique en arbitrage politique (qui bénéficie, qui supporte, qui diffère, qui subit). [8][23][32]

L'entropie est donc un opérateur structurant mais non totalisant : elle fixe des limites sans produire à elle seule les critères de décision.

A.4 Pourquoi des interfaces

Les opérateurs structurent l'analyse : ils indiquent où un régime se défait (seuils, pathologies, points de bascule). Les interfaces désignent, à l'inverse, les zones de médiation par lesquelles il devient possible d'intervenir : elles traduisent des tensions interrégimes en prises praticables (techniques, organisationnelles, normatives, délibératives), sans prétendre supprimer les conflits.

Une interface n'est ni un régime supplémentaire ni un « niveau supérieur » de gouvernance. C'est un dispositif de traduction : il rend certaines contraintes manipulables, opposables ou discutables, selon le type d'interface mobilisée.

(Développements : Partie VI, chap. 18 — fonctionnement, limites, modes d'échec.)

A.4.1 Typologie minimale des interfaces

1) Interface technique

Fonction : rendre les contraintes opérables (calculables, traçables, paramétrables).

Exemples :

- API et protocoles d'interopérabilité (de type MCP),
- formats de traçabilité,
- marquage numérique,
- quotas d'usage,
- métriques énergétiques.

Rôle : traduire des exigences de sens, de droit ou de soutenabilité en contraintes que des systèmes peuvent effectivement porter. [34][35]

2) Interface éthique

Fonction : expliciter ce qui compte (valeurs, finalités, critères d'acceptabilité).

Exemples :

- dignité,
- justice intergénérationnelle,
- sobriété,
- non-malfaisance,
- reconnaissance.

Rôle : rendre visibles les arbitrages axiologiques déjà à l'œuvre, afin qu'ils puissent être discutés et assumés. [31]

Précision : l'interface éthique n'est pas un « supplément de sens ». Elle explicite des horizons normatifs (ce qui doit compter) que la réponsabilité ne peut pas trancher à elle seule. [15][31]

3) Interface juridico-normative

Fonction : rendre certaines exigences opposables.

Exemples :

- obligations procédurales,

- audits,
- charge de la preuve,
- sanctions,
- régulations (RIA, DSA, etc.).

Rôle : instituer des seuils non négociables et assigner des responsabilités formelles, sans garantir à elle seule la tenue des régimes. [25][37]

4) Interface éthico-normative

Fonction : traduire des principes en obligations praticables et contextualisées, pour éviter la dissociation entre morale déclarative et conformité procédurale.

Exemples :

- principes UNESCO intégrés dans des audits,
- pilotage humain (entendu comme maintien d'un pouvoir de reprise et de suspension dans les usages, afin que la production de sens demeure répondable),
- exigences de soutenabilité rendues opposables.

Rôle : éviter la dissociation entre morale abstraite et droit purement procédural. [31][34]

A.5 Articulation opérateurs / interfaces

- Les opérateurs indiquent *où* la tenue se fragilise : réponsabilité (sens), fiducie (communication), dette entropique située (métabolisme) ;
- Les interfaces déterminent comment agir sans réduire la situation à une seule métrique ou un seul langage.

Gouverner l'IA, dans ce cadre, signifie instituer des interfaces capables de contenir les effets des déséquilibres sans les dissoudre.

Précision opératoire

« Contenir sans neutraliser » signifie :

1. empêcher qu'un régime impose sa métrique comme mesure unique ;
2. installer des points de reprise (suspension, contestation, révision) là où les arbitrages se font réellement. [21][34]

Annexe B — Instruments de gouvernance architectonique de l'intelligence artefactuelle

Statut de l'annexe

Cette annexe ne propose ni un standard normatif, ni une méthode universelle d'évaluation. Elle rassemble des outils de diagnostic et d'action destinés à traduire l'architectonique des trois régimes en gestes de gouvernance situés.

Ces instruments ne mesurent pas la « performance » d'un système d'IA ; ils interrogent la tenue de l'intelligence artefactuelle : sa capacité à produire du sens endossable, à circuler de manière fiduciaire, et à être soutenue sans accumulation entropique impayable. [34]

B.1 Principes méthodologiques communs aux outils

Les instruments présentés ici reposent sur cinq principes :

Non-hiérarchie des régimes

Aucun outil n'évalue un régime comme fondement des autres. Les diagnostics portent sur des déséquilibres, non sur des déficits absolus. [21]

Seuils plutôt que scores

Les outils identifient des seuils de basculement (impossibilité d'endossement, saturation fiduciaire, désassomption métabolique), non des notes globales. [34]

Contestabilité

Tout diagnostic doit pouvoir être discuté, contesté et révisé. Les instruments sont conçus comme supports de délibération, non comme verdicts automatiques. [2]

Activation d'interfaces

Chaque outil mobilise une ou plusieurs interfaces (technique / éthique / normative-juridique / éthico-normative). [31]

Temporalité entropique

Les diagnostics intègrent une dimension temporelle : un système peut être stable localement tout en accumulant une dette entropique qui hypothèque ses régimes futurs. [8]

B.2 Outil 1 — Matrice de diagnostic : *Le Triangle des déséquilibres*

Définition

La Matrice de diagnostic est un outil visuel et conceptuel permettant de situer un projet d'IA dans l'espace des trois régimes : sens, communication, métabolisme. Elle vise à rendre visibles les tensions interrégimes plutôt qu'à produire une évaluation unifiée.

Ce que l'outil détecte

- Hypertrophie d'un régime au détriment des autres
- Configurations pathologiques typiques :
 - communication sans orientation,
 - sens sans soutenabilité,

- métabolisme sans signification. [27]

Seuil critique

Un seuil est franchi lorsqu'un régime devient incapable de contraindre les autres :

- sens sans coût assumable,
- communication sans répondabilité,
- métabolisme sans orientation interprétative. [34]

Action de gouvernance

- Rééquilibrage par réduction, redirection ou limitation d'un régime
- Activation ciblée d'interfaces (ex. traçabilité technique + obligation procédurale) [37]

Interfaces mobilisées

Technique, éthico-normative [31]

B.3 Outil 2 — Protocole de répondabilité

Définition

Le protocole de répondabilité vise à garantir que certaines productions de l'IA puissent être endossées au point d'usage. Il repose sur l'exigence d'un certificat d'endossement, et non sur une explication cognitive complète du système. [18] Où l'on entend par « certificat d'endossement » non un label, mais un minimum opposable : qui répond, sur quoi, avec quels périmètres, et avec quels droits de suspension. [4]

Ce que l'outil détecte

- Incapacité à répondre de :
 - sources,
 - transformations,
 - limites,
 - conditions d'usage.
- Rupture entre production d'énoncés et possibilité d'assomption. [18]

Seuil critique

Le seuil est franchi lorsqu'aucune entité humaine ou institutionnelle ne peut raisonnablement répondre des effets produits.

Action de gouvernance

- Débrailage immédiat (reprise humaine)
- Réduction du périmètre d'usage
- Suspension du système pour usages critiques [37]

Interfaces mobilisées

Technique (traçabilité), normative (charge de la preuve), éthico-normative. [25]

Les constitutions d'IA constituent un cas particulier de tentative de stabilisation fiduciaire par inscription normative. Leur analyse met en évidence les limites de ce type de dispositif lorsqu'il n'est pas articulé à des mécanismes explicites de responsabilité et d'assomption métabolique (voir l'**annexe C**). [33]

B.4 Outil 3 — Indicateur de charge métabolique réelle

Définition

Cet outil vise à rendre visible la charge métabolique totale d'un système d'IA, au-delà du coût computationnel immédiat. [8]

Ce que l'outil détecte

- Accumulation de dette entropique située :
 - énergétique,
 - matérielle,
 - organisationnelle,
 - humaine.
- Externalisation spatiale ou temporelle des coûts. [23]

Seuil critique

Le seuil de désassomption est franchi lorsque le coût métabolique cumulé dépasse ce qui peut être assumé à organisation constante (temps humain, budget, capacités matérielles), ou impose un déplacement systématique des coûts hors du périmètre de décision.

Action de gouvernance

- Simplification du système
- Ralentissement volontaire
- Abandon ou interdiction dans certains contextes [34]

Interfaces mobilisées

Technique (mesures), éthique (justice intergénérationnelle), normative (obligations environnementales) [31][37]

B.5 Outil 4 — Test de fiducie communicationnelle

Définition

Le test de fiducie communicationnelle évalue si un système d'IA dégrade ou soutient le milieu de confiance dans lequel il circule. [27]

Ce que l'outil détecte

- Saturation des flux par des énoncés non endossables,

- Fragilisation des chaînes fiduciaires (impossibilité de remonter, discuter, corriger),
- Déplacement / mimétisme de l'autorité interprétative,
- Étouffement de la parole humaine par la textoralité artificielle. [28]

Seuil critique

Le seuil est franchi lorsque la circulation devient plus efficace que toute possibilité de discussion, de désaccord ou de correction. [27]

Action de gouvernance

- Création d'espaces à priorité humaine
- Limitation ou interdiction de l'IA dans certaines délibérations
- Réduction de la production automatique [31]

Interfaces mobilisées

Éthique (dignité, reconnaissance), normative (zones protégées), éthico-normative [37]

B.6 Outil 5 — Grille des agencements de gouvernance

Définition

La grille des agencements de gouvernance permet de traduire les diagnostics précédents en choix organisationnels explicites. [34]

Ce que l'outil structure

- Niveaux de criticité des usages,
- Équilibres interrégimes requis,
- Interfaces à activer,
- Répondabilités instituées et obligations opposables.

Seuil critique

Un usage est jugé non gouvernable lorsqu'aucune configuration organisationnelle ne permet de maintenir la co-contrainte des régimes.

Action de gouvernance

- Réorganisation des responsabilités
- Redéfinition des périmètres d'automatisation
- Institution de seuils non négociables [37]

Interfaces mobilisées

Toutes.

B.7 Table de correspondance synthétique

Outil	Régime principal	Seuil critique	Interfaces activées
Triangle des déséquilibres	Tous	Déséquilibre structurel	Technique, éthico-normative
Protocole de réponsabilité	Sens	Impossibilité d'endossement	Technique, normative
Charge métabolique réelle	Métabolisme	Désassomption	Technique, éthique
Fiducie communicationnelle	Communication	Saturation fiduciaire	Éthique, normative
Grille de gouvernance	Interrégimes	Non-gouvernabilité	Toutes

Gabarits pratiques — Usage institutionnel

Les gabarits suivants peuvent être utilisés tels quels par des institutions, laboratoires ou autorités de contrôle. [37] Ils restent inchangés dans leur forme, car ils constituent précisément des interfaces opérationnelles plutôt qu'un texte théorique.

Gabarit 1 — Fiche de diagnostic interrégime

Projet / système :

Usage principal :

Niveau de criticité :

Régime	État	Seuil atteint ?	Commentaire
Sens	<input type="checkbox"/> Stable <input type="checkbox"/> Fragile <input type="checkbox"/> Critique	<input type="checkbox"/> Oui <input type="checkbox"/> Non	
Communication	<input type="checkbox"/> Stable <input type="checkbox"/> Fragile <input type="checkbox"/> Critique	<input type="checkbox"/> Oui <input type="checkbox"/> Non	
Métabolisme	<input type="checkbox"/> Stable <input type="checkbox"/> Fragile <input type="checkbox"/> Critique	<input type="checkbox"/> Oui <input type="checkbox"/> Non	

Diagnostic global :

☐ Co-appartenance maintenue

☐ Déséquilibre instable

☐ Non-gouvernable en l'état

Gabarit 2 — Fiche d'endossement (réponsabilité)

Production concernée :

Usage :

Peut-on répondre de :

- Sources ? ☐ Oui ☐ Non
- Transformations ? ☐ Oui ☐ Non
- Limites connues ? ☐ Oui ☐ Non
- Effets prévisibles ? ☐ Oui ☐ Non

Décision :

☐ Endossement possible

☐ Usage restreint

☐ Débrailage requis

Gabarit 3 — Tableau de charge métabolique réelle

Dimension	Indicateur	Valeur Seuil
-----------	------------	--------------

Énergie	kWh / usage	
---------	-------------	--

Eau	L / usage	
-----	-----------	--

Travail humain	h / mois	
----------------	----------	--

Charge cognitive	h supervision	
------------------	---------------	--

Dette entropique estimée :

☐ Acceptable ☐ Croissante ☐ Critique

Gabarit 4 — Test de confiance communicationnelle

Contexte :

Ratio IA / humain dans les flux :

Autorité identifiable ? ☐ Oui ☐ Non

Effet sur la discussion :

☐ Soutient

☐ Neutralise

☐ Dégrade

Mesure proposée :

☐ Limitation

☐ Zone protégée

☐ Interdiction contextuelle

Gabarit 5 — Grille d'agencement de gouvernance

Niveau Régime prioritaire Interfaces activées Décision

Critique Sens Tech + Normative

Élevé Sens + Métabolisme Éthico-normative

Modéré Communication Technique

Conclusion de l'annexe

Ces instruments ne prétendent pas « maîtriser » l'intelligence artificielle. Ils visent à rendre nommables, discutables et arbitrables les conditions sous lesquelles une intelligence peut être reconnue, soutenue et assumée. [34]

Ils prolongent l'essai en un outillage critique, sans en réduire la portée philosophique : gouverner l'IA ne consiste pas à optimiser des systèmes, mais à tenir les conditions de leur intelligibilité commune, dans le temps et sous contrainte. [8]

L'ambition conjointe des annexes A et B est ainsi de proposer un seuil qui devrait être franchi : ne plus seulement penser l'intelligence artificielle, mais outiller sa gouvernabilité, sans jamais la réduire à une procédure automatique, ni à un standard universel, et encore moins à une intelligence simplement « artificielle ». [34]

*

Annexe C — Les constitutions d'IA comme interfaces éthico-normatives

Statut de l'annexe

Cette annexe propose une analyse architectonique d'un type de dispositif de gouvernance de plus en plus mobilisé dans le champ de l'intelligence artificielle : les constitutions d'IA, qui se proposent de gouverner la cohérence apparente des systèmes, mais ne garantissent ni leur responsabilité au point d'usage, ni l'orientation fiduciaire de leur circulation, ni l'assomption de leur dette entropique.

Elle ne vise ni à évaluer moralement ces initiatives, ni à discuter la sincérité des intentions qui les sous-tendent. Elle examine leur fonction structurelle : ce qu'elles permettent effectivement de gouverner, et ce qu'elles laissent hors champ.

L'analyse prend appui sur des dispositifs publics et documentés — notamment la « constitution » intégrée à l'entraînement des modèles développés par [Anthropic](#) (dite *Constitutional AI*) — non comme cas singulier, mais comme exemple paradigmatique d'une tendance plus générale. [33]

C.1 Qu'appelle-t-on une « constitution d'IA » ?

On appelle ici *constitution d'IA* un document public ou semi-public qui :

- explicite un ensemble de principes, valeurs et règles de conduite assignés à un système d'IA ;
- se présente comme structurant directement son comportement, notamment via l'entraînement, l'évaluation ou la supervision ;
- revendique un rôle de référence normative stable, censée guider les décisions du système dans des situations variées, y compris imprévues.

À la différence de simples chartes éthiques ou de codes de bonne conduite externes, les constitutions d'IA prétendent opérer en amont, au cœur même des mécanismes de production des sorties. Elles ne se contentent pas de prescrire ; elles cherchent à configurer. [33]

C.2 La constitution comme interface éthico-normative

Dans le cadre proposé par cet ouvrage, une constitution d'IA ne constitue ni un régime supplémentaire, ni un substitut au droit ou à l'éthique. Elle fonctionne comme une interface éthico-normative.

À ce titre, elle remplit trois fonctions principales :

1. Explicitation axiologique

Elle rend visibles certaines valeurs (sécurité, honnêteté, non-malfaisance, etc.) qui resteraient autrement implicites dans les choix de conception.

2. Traduction normative

Elle transforme ces valeurs en règles, priorités ou interdictions susceptibles d'être intégrées à des procédures techniques (entraînement, filtrage, évaluation).

3. Stabilisation comportementale

Elle vise à produire une cohérence des réponses à travers des contextes variés, en réduisant l'arbitraire perçu ou les contradictions locales.

Sous cet angle, les constitutions d'IA représentent un progrès réel par rapport à des approches purement opportunistes ou non réflexives. Elles reconnaissent explicitement que la conduite des systèmes pose un problème de gouvernance, et non seulement de performance. [33]

C.3 Ce que gouverne effectivement une constitution d'IA

C.3.1 Sur le régime de la communication : stabilisation fiduciaire

Les constitutions d'IA agissent principalement sur le régime communicationnel.

Elles contribuent à stabiliser une fiducie : les usagers, développeurs et institutions peuvent se représenter « ce à quoi le système se tiendra », même en l'absence d'une supervision humaine directe.

Cette stabilisation s'opère par la production d'une autorité distribuée :

- la constitution sert de référence ;
- cette référence est intégrée aux mécanismes d'entraînement ;
- les sorties du système en portent la trace, sous forme de régularités et de refus cohérents.

La constitution fonctionne ainsi comme un point d'ancrage fiduciaire, susceptible de renforcer la confiance et la circulation des usages. [33]

C.3.2 Sur le régime du sens : cohérence normative apparente

Sur le régime du sens, l'effet est plus ambivalent.

Une constitution peut améliorer la cohérence interprétative des réponses : elles semblent mieux justifiées, plus prudentes, plus alignées avec des principes explicites.

Cependant, cette amélioration porte principalement sur la forme du raisonnement, non sur ses conditions d'endossement. Le système peut produire des justifications normatives convaincantes sans que soit instituée une responsabilité effective au point d'usage. [33]

Autrement dit : la constitution peut mimer la responsabilité, sans la garantir.

C.4 Ce que la constitution ne peut pas gouverner

C.4.1 La responsabilité institutionnelle

Une constitution d'IA ne désigne pas :

- qui peut répondre des effets produits,
- selon quelles procédures,
- avec quels droits de suspension, de correction ou de recours.

Elle ne crée pas de points de reprise explicites.

La responsabilité reste diffuse, reportée sur des entités humaines ou organisationnelles qui ne sont pas intégrées au dispositif constitutionnel lui-même.

Ainsi, une sortie peut être « conforme à la constitution » sans être endossable par une entité identifiable. [33]

C.4.2 L'orientation fiduciaire de la circulation

En renforçant la cohérence comportementale, les constitutions peuvent paradoxalement durcir la fiducie. Plus le système paraît fiable et moralement aligné, plus il tend à être relayé, intégré et mobilisé — parfois au détriment de la contestabilité.

La constitution ne garantit pas que la circulation reste orientable, c'est-à-dire suspendable, discutée, reconfigurable. Elle peut stabiliser la confiance sans instituer les conditions de sa remise en question. [33]

C.4.3 La dette entropique située

Enfin, les constitutions d'IA laissent largement hors champ le régime métabolique.

Elles ne traitent ni :

- des coûts énergétiques et matériels,
- du travail humain de maintien,
- des effets d'accumulation organisationnelle,
- ni de la répartition temporelle et spatiale de la dette entropique.

Or l'intégration constitutionnelle peut elle-même accroître ces coûts (cycles d'entraînement supplémentaires, filtrage, supervision accrue). La soutenabilité reste implicite, supposée, mais non gouvernée. [33]

C.5 Une pathologie typique : gouverner la conduite sans gouverner les conditions

Cette analyse permet d'identifier une pathologie récurrente :

la gouvernance de la conduite sans gouvernance des conditions.

Les constitutions d'IA peuvent :

- stabiliser ce que le système *fait*,
- sans stabiliser les conditions sous lesquelles il est légitime de le faire,
- ni les coûts que cette stabilisation engage. [33]

Elles déplacent ainsi le problème de la gouvernabilité vers l'amont normatif, sans traiter les seuils de désassomption métabolique ni les points de rupture fiduciaire.

C.6 Place des constitutions dans une architectonique de la gouvernance

Dans le cadre proposé par cet ouvrage, les constitutions d'IA doivent être comprises comme :

- des interfaces éthico-normatives utiles mais partielles ;
- des dispositifs de cohérence, non des mécanismes de gouvernance complète ;
- des outils qui nécessitent d'être articulés à :
 - des protocoles explicites de responsabilité,
 - des dispositifs de contestabilité communicationnelle,
 - des indicateurs et seuils de dette entropique située.

Elles peuvent contribuer à la gouvernabilité de l'intelligence artificielle, mais ne sauraient s'y substituer. Gouverner l'IA ne consiste pas à doter les systèmes d'une constitution, mais à tenir ensemble les conditions de sens, de circulation et de soutenabilité qui rendent une intelligence reconnaissable et assumable. [33]

Conclusion de l'annexe

Les constitutions d'IA témoignent d'un déplacement important : la reconnaissance explicite que la conduite des systèmes d'intelligence artefactuelle pose un problème de gouvernance, et non seulement d'ingénierie.

Présentées comme « constitutions », certaines chartes peuvent renforcer la cohérence normative apparente et stabiliser une confiance d'usage. Mais elles ne constituent pas, à elles seules, une gouvernance : elles n'instituent pas nécessairement des droits de reprise opposables, ne garantissent pas une responsabilité assignable au point d'usage, et laissent souvent hors champ la dette entropique et les conditions matérielles de soutenabilité. Le risque n'est pas seulement l'inefficacité : c'est la substitution d'une cohérence discursive à une gouvernabilité réelle.

Elles montrent toutefois, par leurs limites mêmes, que la gouvernabilité ne peut être assurée par un dispositif unique, fût-il normatif et intégré. L'intelligence artefactuelle ne se laisse pas gouverner par constitution ; elle exige une architectonique d'opérateurs et d'interfaces, capable de rendre visibles, discutables et arbitrables les conditions de sa tenue. [33]

*

Annexe D — Le NIST AI RMF comme interface de gouvernance (lecture architectonique)

Statut de l'annexe

Cette annexe propose une lecture architectonique du [NIST Artificial Intelligence Risk Management Framework \(AI RMF 1.0\)](#), publié par le *National Institute of Standards and Technology* (NIST) en 2023. [34]

Ce document constitue aujourd'hui l'un des cadres institutionnels les plus structurants dans les tentatives contemporaines de gouvernance de l'intelligence artificielle.

L'ouvrage mobilise à de nombreuses reprises ce cadre, non comme un modèle normatif à appliquer, mais comme un exemple paradigmatique d'interface juridico-organisationnelle : une tentative de rendre gouvernables des systèmes d'IA au moyen d'une grammaire opératoire du risque.

Il est donc méthodologiquement préférable d'en expliciter ici, une fois pour toutes, la logique, la portée et les limites, afin de stabiliser son statut dans l'économie générale de cette architectonique.

L'objectif n'est ni l'exégèse exhaustive du document, ni une critique prescriptive. Il s'agit de situer le RMF comme dispositif partiellement compatible avec une gouvernance par seuils, mais structurellement incapable d'unifier sous un même cadre les régimes constitutifs de l'intelligence artefactuelle, au sens où aucun cadre unique ne pourrait articuler ces régimes à lui seul.

D.1 Le RMF : une grammaire fonctionnelle plutôt qu'un principe

Le NIST AI RMF se distingue des approches dominantes fondées sur des principes abstraits (« transparence », « équité », « responsabilité ») en proposant une structuration explicitement fonctionnelle de la gestion des risques.

Sa Partie II définit le cœur du cadre (ce que le NIST nomme le *Core* du *Framework*) : quatre fonctions destinées à guider les organisations dans la gestion pratique des risques liés aux systèmes d'IA :

- **GOVERN**
- **MAP**
- **MEASURE**
- **MANAGE**

Ces fonctions constituent moins une théorie de l'IA qu'un schéma opératoire permettant d'inscrire la gouvernance dans des processus organisationnels.

Ce déplacement est décisif. Il ne s'agit plus seulement d'énoncer des valeurs, mais de produire des prises institutionnelles. En cela, le NIST RMF se rapproche d'une intuition centrale de cet ouvrage : gouverner ne consiste pas à corriger après coup, mais à instituer des conditions de tenue.

D.2 GOVERN : gouvernance comme fonction transversale

La fonction **GOVERN** occupe une position particulière : elle traverse toutes les étapes du cycle de vie des systèmes d'IA.

Elle concerne explicitement :

- la structuration des responsabilités,
- les procédures de supervision,
- les dispositifs de documentation,
- et la capacité organisationnelle à maintenir une gestion continue du risque [34].

Dans une lecture architectonique, *GOVERN* correspond à une tentative d'institution de la co-contrainte : produire des mécanismes permettant de ne pas abandonner la gouvernance à des ajustements locaux ou à des promesses générales.

Ce point est important : le NIST reconnaît implicitement qu'il n'existe pas de gouvernance par ajout tardif, mais seulement des formes de gouvernance transversales, installées dès l'amont.

D.3 MAP : la cartographie située comme condition d'intervention

La fonction **MAP** vise à contextualiser un système avant toute prétention à le mesurer ou à le gérer.

MAP consiste à identifier :

- les usages visés,

- les acteurs concernés,
- les populations affectées,
- les environnements institutionnels,
- les dépendances techniques et organisationnelles.

Cette fonction recoupe directement une exigence méthodologique formulée dans cet ouvrage : une gouvernance architectonique commence par un diagnostic interrégime, c'est-à-dire par la localisation des tensions plutôt que par l'optimisation d'indicateurs.

MAP correspond ainsi à une première interface : rendre situables les conditions d'usage, avant que les seuils ne deviennent invisibles.

D.4 MEASURE : la puissance des métriques et leurs limites structurelles

La fonction **MEASURE** vise à quantifier, suivre et objectiver les risques associés à un système.

Elle mobilise des outils classiques :

- tests,
- évaluations,
- audits,
- métriques de robustesse ou d'équité,
- analyses d'impact.

Dans l'architectonique proposée ici, *MEASURE* relève du registre des indicateurs : indispensable, mais structurellement partiel.

Un indicateur mesure ce qui est déjà formaté pour être mesurable. Il ne dit rien, par lui-même, du seuil au-delà duquel le système cesse d'être répondable, fiduciaire ou assumable.

La limite est donc interne : *MEASURE* permet le suivi, mais non l'identification des points de bascule. Sans opérateurs, la mesure devient auto-justifiante.

D.5 MANAGE : gestion du risque ou arbitrage de seuil ?

La fonction **MANAGE** concerne les décisions d'intervention :

- mitigation,
- contrôle,
- limitation,
- redéploiement,
- adaptation continue.

MANAGE est l'espace où la gouvernance devient pratique.

Mais c'est aussi ici qu'apparaît une tension architectonique fondamentale :

- le RMF tend à penser l'intervention comme réduction progressive du risque,
- alors que l'architectonique des régimes impose parfois des décisions de seuil : suspension, renoncement, retrait.

Autrement dit, *MANAGE* reste principalement inscrit dans une logique de gestion, et non dans une logique de désassomption.

Le risque est conceptualisé comme variable maîtrisable, là où certains coûts — entropiques, organisationnels, fiduciaires — sont irréversibles.

D.6 Correspondances avec les trois régimes

Le NIST RMF est particulièrement éclairant lorsqu'on le re-situe dans l'espace des régimes analysés dans cet ouvrage.

Régime du sens : réponsabilité

Le RMF insiste sur la transparence, la documentation et la traçabilité.

Mais il ne thématise pas directement l'opérateur de réponsabilité comme seuil critique : l'incapacité pratique d'endosser le sens produit au point d'usage.

La gouvernance du sens reste abordée sous l'angle du risque informationnel, non sous celui de l'endossement institutionnel. [39]

Régime de la communication : confiance

Le RMF évoque la confiance, la perception et l'acceptabilité.

Mais il ne conceptualise pas la saturation fiduciaire comme seuil propre : le moment où la circulation devient plus rapide que toute capacité collective de reprise.

La dimension communicationnelle demeure donc sous-instituée, alors même qu'elle conditionne la gouvernabilité réelle. [38]

Régime métabolique : dette entropique située

Le RMF inclut certains enjeux environnementaux ou sociétaux.

Mais il ne fait pas de l'irréversibilité entropique un opérateur structurant.

Les coûts matériels, énergétiques et organisationnels apparaissent comme impacts à gérer, non comme flèche du temps imposant des points de non-retour.

C'est ici que l'architectonique proposée dans cet ouvrage déplace le cadre : la soutenabilité n'est pas un risque parmi d'autres, mais une limite constitutive. [40]

D.7 Le RMF comme interface, non comme fondement

Le NIST AI RMF doit être compris comme une interface juridico-organisationnelle majeure.

Il rend opposables certaines exigences :

- devoirs de vigilance,

- procédures internes,
- mécanismes de contrôle,
- obligations de documentation.

Mais il ne peut pas fonder à lui seul la gouvernabilité de l'intelligence artefactuelle, pour une raison architectonique simple : le RMF est un dispositif de gestion du risque, alors que l'intelligence artefactuelle est un effet conditionnel de co-contrainte entre régimes irréductibles.

Il est donc nécessaire, mais non souverain. Opératoire, mais non totalisant. Institutionnel, mais non suffisant.

D.8 Conclusion : une grammaire précieuse, un horizon limité

L'intérêt du NIST RMF est d'offrir une grammaire fonctionnelle de gouvernance qui évite partiellement l'illusion des principes abstraits.

Son cœur (*GOVERN, MAP, MEASURE, MANAGE*) constitue une tentative sérieuse d'instituer des prises organisationnelles sur des systèmes complexes.

Mais son horizon demeure celui d'une gouvernance par gestion continue, là où l'architectonique des seuils impose parfois des décisions de rupture, de suspension ou de renoncement.

Cette annexe permet ainsi de stabiliser le statut du RMF dans cet ouvrage : non comme solution normative, mais comme interface exemplaire, éclairant à la fois les possibilités et les limites d'une gouvernance contemporaine structurée par le risque.

*

Annexe E — ISO/IEC 42001 comme interface organisationnelle de gouvernance

Statut de l'annexe

Cette annexe propose une lecture architectonique de la [norme ISO/IEC 42001:2023*](#), premier standard international spécifiquement consacré aux systèmes de management de l'intelligence artificielle.

Elle n'examine pas cette norme comme solution globale à la gouvernance de l'IA, mais comme un dispositif d'interface :

- ce qu'elle rend gouvernable institutionnellement,
- les opérateurs qu'elle active implicitement,
- et les limites structurelles de son approche managériale.

* ISO/IEC 42001 mobilise le lexique standard des normes de management (organisation, conformité, audit, amélioration continue). Son vocabulaire, très procédural, constitue précisément une interface organisationnelle : il rend gouvernable par système documentaire ce qui, architectoniquement, relève d'arbitrages interrégimes.

E.1 ISO 42001 : gouverner par système de management

ISO/IEC 42001 ne vise pas à définir ce qu'est une IA acceptable, ni à fixer des interdictions substantielles.

Son geste principal est organisationnel :

- instituer des processus internes,
- formaliser des responsabilités,
- documenter des décisions,
- encadrer des usages par procédures.

Elle fonctionne ainsi comme un standard de *management system*, dans la continuité des logiques ISO classiques.

Dans le cadre de cet ouvrage, elle doit être comprise comme une interface organisationnelle, et non comme un opérateur.

E.2 Activation indirecte de la responsabilité

Le point le plus important du standard est l'institution de procédures visant à rendre les systèmes auditable et assignables.

ISO 42001 agit principalement sur l'opérateur de responsabilité, en exigeant :

- des mécanismes de supervision,
- des chaînes de décision explicites,
- des obligations de documentation.

Elle contribue ainsi à stabiliser l'endossement institutionnel, non par explication du sens, mais par assignation de responsabilités organisationnelles.

E.3 Fiducie et gouvernance procédurale

ISO 42001 participe également à une stabilisation communicationnelle indirecte :

- certification,
- conformité visible,
- standardisation des pratiques.

Elle soutient ainsi des formes de fiducie institutionnelle :

on se remet à des systèmes parce qu'ils sont encadrés, certifiés, documentés.

Mais cette fiducie reste procédurale : elle ne garantit ni contestabilité effective, ni reprise réelle au point d'usage.

E.4 Limite majeure : le métabolisme comme angle mort

Comme la plupart des standards ISO, 42001 tend à traiter la soutenabilité comme une catégorie parmi d'autres, non comme opérateur limite.

La dette entropique située — coûts irréversibles, externalisations, seuils de désassomption — n'y apparaît pas comme contrainte structurante.

Le risque est donc typique :

- un système peut être conforme,
- tout en restant métaboliquement prédateur,
- ou politiquement désassumable.

E.5 Pathologie typique : le juridisme managérial

Lorsqu'elle est absolutisée, la norme ISO 42001 produit une pathologie connue : la gouvernance par procédure.

Les organisations disposent alors :

- d'audits,
- de cadres internes,
- de *checklists*,

sans que les arbitrages entre régimes deviennent pour autant discutables ou opposables.

La norme gouverne la conformité, non la co-contrainte.

Conclusion de l'annexe

ISO/IEC 42001 constitue une interface organisationnelle majeure : elle rend possible une gouvernance interne minimale de l'IA par institution de procédures et de responsabilités. [35]

Mais elle ne totalise pas l'intelligence artefactuelle.

Elle ne peut être opérante que si elle reste articulée :

- à des exigences de réponsabilité réelle,
- à des conditions de fiducie contestable,
- et surtout à des seuils métaboliques non neutralisables.

Annexe F — L'AI Act européen comme interface juridico-normative (lecture architectonique)

Statut de l'annexe

Cette annexe propose une lecture architectonique du [Règlement \(UE\) 2024/1689, dit AI Act](#), adopté en 2024. [37]

Elle ne discute pas ce texte comme solution définitive, mais comme une interface normative majeure :

- ce qu'il rend opposable,
- les opérateurs qu'il active institutionnellement,
- et les limites structurelles de son approche par classification des risques.

F.1 Gouverner par niveaux de risque : traduction juridique d'un seuil

L'AI Act repose sur un geste central :

- classer les systèmes selon leur niveau de risque,
- imposer des obligations graduées,
- interdire certains usages jugés inacceptables.

Il introduit ainsi une logique explicite de seuil :

- seuil d'interdiction,
- seuil d'obligation renforcée,
- seuil de transparence minimale.

Dans les termes de cet ouvrage, il s'agit d'une interface juridico-normative typique : rendre certains arbitrages opposables.

F.2 Répondabilité : institution de devoirs formels

L'AI Act active fortement l'opérateur de répondabilité, en imposant :

- documentation,
- traçabilité,
- supervision humaine,
- obligations de conformité.

Il produit ainsi une assignation juridique de responsabilité.

Mais cette répondabilité reste principalement procédurale : elle garantit un cadre de reprise, non l'endossement effectif du sens au point d'usage. [39]

F.3 Fiducie : stabiliser la circulation par obligations de transparence

L'*AI Act* agit également sur la fiducie communicationnelle :

- obligations de signalement,
- encadrement des modèles généralistes,
- exigences de transparence pour certains contenus.

Il vise à empêcher une circulation complètement désorientée.

Mais il rencontre une limite structurelle : la fiducie ne se décrète pas juridiquement ; elle se maintient dans un milieu institutionnel et technique plus large. [38]

F.4 Limite métabolique : le grand impensé normatif

Comme la plupart des régulations contemporaines, l'*AI Act* traite faiblement le régime métabolique.

La dette entropique située — infrastructures, extraction, dissipation, seuils de désassomption — n'y apparaît pas comme opérateur structurant.

Or une gouvernance par risque peut rester aveugle :

- aux coûts irréversibles,
- aux externalisations territoriales,
- aux conflits de répartition.

Le droit pilote l'usage, mais peine à gouverner le métabolisme. [40]

F.5 Pathologie possible : conformité sans gouvernabilité

Le risque typique est celui du juridisme procédural :

- les obligations sont respectées,
- les audits existent,
- mais la co-contrainte des régimes se défait malgré tout.

Un système peut être légalement gouverné tout en devenant :

- fiduciairement saturant,
- métaboliquement insoutenable,
- politiquement désassumable.

Conclusion de l'annexe

L'*AI Act* constitue aujourd'hui l'interface normative la plus structurante de la gouvernance de l'IA en Europe.

Il rend opposables des seuils, institue des obligations, et assigne des responsabilités formelles.

Mais il ne peut suffire à gouverner l'intelligence artefactuelle comme effet architectonique.

Sa portée dépendra de son articulation :

- à des dispositifs réels de responsabilité,
- à des conditions pratiques de confiance contestable,
- et à la reconnaissance des limites métaboliques irréversibles.

*

Annexe G — Document-type de gouvernance architectonique située (exemple opérationnel, fictif)

Statut de l'annexe

Cette annexe propose un exemple concret et minimal de ce que signifie, en pratique, une gouvernance architectonique de l'intelligence artefactuelle.

Elle ne constitue ni un standard, ni un modèle universel, ni une procédure recommandée.

Son objectif est strictement méthodologique : montrer comment les notions développées dans l'ouvrage — régimes, opérateurs, seuils, interfaces, arbitrages — peuvent être traduites dans un format documentaire réel, opposable et révisable.

L'enjeu est de rendre visible ce qui demeure souvent abstrait :

- gouverner ne consiste pas à optimiser un système,
- mais à maintenir la co-contrainte entre sens, communication et métabolisme,
- jusqu'au point où cette tenue cesse d'être collectivement assumable.

G.1 Pourquoi un exemple documentaire ?

Les chapitres précédents ont montré que l'intelligence artefactuelle n'est pas une propriété technique des systèmes, mais un effet conditionnel, émergent et réversible, produit par la tenue conjointe de trois régimes irréductibles :

- sens (responsabilité),
- communication (confiance),
- métabolisme (dette entropique située).

Cette analyse impose une conséquence pratique immédiate : si l'intelligence est un effet de co-contrainte, alors la gouvernance ne peut pas être conçue comme un correctif externe, mais comme une pratique située de maintien.

Or cette pratique ne peut exister sans formats.

Gouverner suppose toujours des supports institutionnels : fiches, protocoles, décisions explicites, clauses de suspension, dispositifs de reprise. C'est ce que cette annexe explicite sous forme d'un document-type.

G.2 Exemple fictif mais réaliste : usage hospitalier d'un modèle génératif

Le cas présenté ici est volontairement simple.

Il concerne l'usage d'un système de génération automatique de comptes rendus médicaux dans un service hospitalier.

Ce contexte permet de rendre immédiatement perceptibles :

- les exigences de réponsabilité (qui répond du sens produit ?),
- les exigences fiduciaires (comment circulent les énoncés ?),
- les contraintes métaboliques et organisationnelles (quel coût cumulatif humain et infrastructurel ?).

Document-type — Fiche de décision architectonique située

Fiche institutionnelle : Assistance IA pour comptes rendus médicaux

Système concerné : modèle génératif de synthèse de consultation

Contexte : service d'oncologie (hôpital public)

Usage : aide à la rédaction des comptes rendus médicaux

Date : février 2026

Responsabilité formelle : direction médicale + référent IA

1. Diagnostic interrégime (où se défait la co-contrainte ?)

Régime du sens — Réponsabilité

Le système produit des résumés cliniquement plausibles.

Cependant :

- certaines formulations apparaissent sans traçabilité,
- les limites du système ne sont pas explicitement reprises,
- aucun acteur ne peut endosser entièrement les effets interprétatifs au point d'usage.

Indice critique : proximité du seuil de non-endossement.

Régime de la communication — Fiducie

Les résumés sont intégrés dans les dossiers patient.

L'usage devient progressivement automatique :

- les équipes médicales s'en remettent aux sorties faute de temps,
- la contestation devient marginale ou tardive,
- la circulation impose sa propre autorité.

Indice critique : saturation fiduciaire (fiducie forcée).

Régime métabolique — Dette entropique située

L'usage du système entraîne une accumulation organisationnelle :

- dépendance à un prestataire externe,
- surcharge durable de supervision humaine,
- coûts structurels de maintenance et de conformité.

La dette n'est pas seulement économique : elle est temporelle, organisationnelle et irréversible.

Indice critique : début de désassomption métabolique située.

2. Seuil identifié (point de bascule)

Le seuil critique est atteint lorsque :

- le système reste performant localement,
- mais aucune entité ne peut répondre institutionnellement des effets produits,
- et la circulation devient non contestable de fait.

Il s'agit donc d'un franchissement combiné :

- réponsabilité fragilisée,
- saturation fiduciaire,
- accumulation organisationnelle.

3. Arbitrage explicite (ce qui est sacrifié / maintenu)

Deux options sont identifiées :

Option A — Automatisation complète

- gains immédiats de fluidité,
- mais dissolution de l'endossement,
- croissance de la dette organisationnelle.

Option B — Restriction et reprise humaine obligatoire

- perte d'efficacité immédiate,
- mais restauration d'un point de reprise interprétatif,
- limitation de la fiducie forcée.

Arbitrage assumé : une baisse de fluidité est acceptée afin de préserver la réponsabilité médicale.

4. Interfaces activées (comment intervenir sans réduire)

Interface technique

- suppression de l'intégration automatique au dossier patient,
- activation du mode « brouillon uniquement »,
- journalisation obligatoire des transformations humaines.

Interface juridico-normative

- aucun texte IA n'est opposable sans signature médicale,
- droit institutionnel de suspension immédiate,
- charge de la preuve sur l'organisation utilisatrice.

Interface éthique

- explicitation de la priorité : endossement > vitesse,
- reconnaissance du patient comme destinataire ultime du sens.

5. Décision de gouvernance

✚ Décision : **Restriction d'usage**

Le système est maintenu comme outil préparatoire, mais interdit comme instance de production finale. Le modèle n'est pas supprimé : il est requalifié architectoniquement.

6. Clause de réversibilité

Réévaluation à trois mois.

Le dispositif sera :

- renforcé si la responsabilité est institutionnalisable,
- suspendu si la dette organisationnelle continue de croître.

7. Formule synthétique (gouvernance architectonique)

Ce document exprime un principe minimal :

- gouverner n'est pas maximiser une performance,
- mais maintenir des points de reprise entre régimes.

Lorsque ces points se ferment, gouverner peut signifier :

- limiter,
- suspendre,
- ou renoncer.

G.3 Ce que cet exemple rend visible

Cet exemple documentaire montre que la gouvernance architectonique n'est pas une théorie supplémentaire, mais une pratique d'institution des seuils.

Elle consiste à :

- repérer le régime dominant,
- identifier le seuil de bascule,
- rendre l'arbitrage opposable,
- activer des interfaces multiples,
- maintenir la réversibilité avant l'irréversibilité.

Il ne s'agit pas de gouverner « le modèle », mais de gouverner la tenue fragile de l'effet d'intelligence.

Conclusion de l'annexe

Ce document-type n'a pas vocation à être reproduit comme norme.

Il sert à rendre tangible une thèse centrale de l'ouvrage : la gouvernance de l'intelligence artificielle ne consiste pas à ajouter des couches, mais à instituer des points de reprise entre régimes irréductibles.

L'intelligence artificielle ne se gouverne pas par promesse, mais par seuils, arbitrages et interfaces situées.

Ici, l'exemple développé a une fonction méthodologique : montrer, sur un document fictif et volontairement simple, comment une gouvernance architectonique peut être formulée en termes de seuils, d'arbitrages interrégimes et d'activation d'interfaces, plutôt qu'en termes d'optimisation ou de conformité abstraite.

Il s'agissait d'explicitier une logique générale :

- repérer où un régime tend à s'autonomiser,
- identifier le seuil critique associé (non-endossement, saturation fiduciaire, désassomption),
- traduire ce diagnostic en prises situées et réversibles.

Mais une telle grammaire de gouvernance ne prend pleinement sens que lorsqu'elle est confrontée à des dispositifs réels, institutionnels, déjà stabilisés dans la vie sociale.

C'est pourquoi l'annexe suivante propose un cas pratique : *Parcoursup*, infrastructure nationale d'allocation sous contrainte, où se condensent de manière exemplaire les trois régimes analysés dans cet ouvrage.

L'enjeu n'est pas d'en faire une critique normative directe, mais de montrer comment un système peut rester fonctionnel et performant tout en franchissant progressivement des seuils de non-gouvernabilité — précisément parce que la co-contrainte entre sens, communication et métabolisme y devient difficile à maintenir.

Annexe H — *Parcoursup* comme cas de gouvernance architectonique (exemple appliqué)

Statut de l'annexe

Cette annexe propose un exemple appliqué de lecture architectonique d'un dispositif institutionnel réel : [*Parcoursup*](#) (procédure nationale française d'affectation dans l'enseignement supérieur).

L'objectif n'est pas d'évaluer *Parcoursup* politiquement ou moralement, mais de montrer comment un système sociotechnique de décision :

- produit du (non-)sens institutionnel,
- organise une circulation fiduciaire,
- repose sur un métabolisme organisationnel lourd,

et comment, en franchissant certains seuils, il peut devenir localement performant mais globalement ingouvernable.

H.1 Configuration générale

Parcoursup est une infrastructure nationale d'allocation sous contrainte :

- gestion de vœux (candidats),
- classement (formations),
- affectation (État),
- décisions distribuées sous forme de flux continus.

Il constitue un cas typique d'intelligence artefactuelle au sens de cet ouvrage : un effet de coordination institutionnelle produit par la tenue conjointe de trois régimes hétérogènes.

H.2 Diagnostic interrégime (Triangle des déséquilibres)

Régime	Fonction dans <i>Parcoursup</i>	Fragilité structurale typique
Sens	Décisions d'acceptation/refus produisant un sens biographique fort	Non-endossement local des critères
Communication	Flux de notifications, classements, listes d'attente	Saturation fiduciaire (décision par défaut)
Métabolisme	Gestion de pénurie et optimisation organisationnelle des capacités	Rationalisation muette (métabolisme substitut au sens)

Parcoursup est gouvernable tant que ces régimes restent mutuellement contraints.

Les tensions apparaissent lorsque l'un d'eux s'autonomise.

H.3 Seuil critique 1 — Répondabilité (sens)

Question opératoire

Peut-on répondre du sens produit au point d'usage ?

Indices de franchissement

- impossibilité de retracer clairement les critères de classement,
- responsabilité diffuse entre plateforme / formation / rectorat,
- absence de prise pour contester localement une décision.

Seuil atteint lorsque

aucune entité identifiable ne peut raisonnablement endosser les effets produits par une décision pourtant décisive.

Pathologie typique

Production de sens institutionnel non répondable

→ Le système fonctionne, mais devient architecturalement irresponsable.

H.4 Seuil critique 2 — Fiducie (communication)

Question opératoire

La circulation des décisions reste-t-elle orientable et contestable ?

Indices de saturation fiduciaire

- accélération et densité des flux (mouvements quotidiens),
- multiplication des énoncés (« en attente », « rang 2450 »),
- incapacité collective à hiérarchiser ou interrompre le processus.

Seuil atteint lorsque

la confiance devient forcée : on s'en remet au flux faute de pouvoir en suspendre la logique.

Pathologie typique

Communication sans orientation

→ Le crédit se stabilise par automatisme plutôt que par reprise institutionnelle.

H.5 Seuil critique 3 — Désassomption (métabolisme)

Question opératoire

Les coûts organisationnels et politiques de l'allocation restent-ils assumables ?

Dette entropique organisationnelle

Parcoursup repose sur :

- charge administrative massive,
- travail humain de classement et de gestion des recours,

- conflits d'usage liés à la pénurie structurelle,
- pressions psychologiques et sociales durables.

Seuil atteint lorsque

le système « tient » principalement comme gestion de flux, mais les coûts sont déplacés vers l'aval (familles, lycées, recours) sans assomption collective explicite.

Pathologie typique

Métabolisme sans signification

→ optimisation organisationnelle substituée à une orientation partagée.

H.6 Résultat architectonique : performance locale, ingouvernabilité globale

Parcoursup illustre une configuration où :

- le sens continue d'être produit (décisions effectives),
- la communication continue de circuler (flux continus),
- le métabolisme continue de soutenir (gestion sous contrainte),

mais où les seuils critiques sont proches ou franchis :

- non-répondabilité des critères,
- saturation fiduciaire,
- désassomption organisationnelle.

Le système devient alors typique d'une intelligence artefactuelle : opératoire comme coordination, mais fragile comme gouvernance.

H.7 Interfaces mobilisables (interventions minimales)

Déséquilibre identifié	Interface pertinente	Exemple de geste de gouvernance
Non-répondabilité	Technique + normative	obligation de traçabilité opposable des critères
Saturation fiduciaire	Normative + éthique	zones de suspension, délais garantis, priorité humaine
Désassomption métabolique	Éthique + organisationnelle	reconnaissance explicite de la pénurie comme seuil politique, non comme fatalité technique

Conclusion de l'annexe

Parcoursup constitue un cas paradigmatique de gouvernance architectonique :

- il ne s'effondre pas,

- il reste performant,
- mais il franchit des seuils où l'endossement, la contestation et l'assomption deviennent difficiles.

Il montre ainsi que gouverner un dispositif algorithmique ne consiste pas seulement à optimiser ou à réguler après coup, mais à maintenir activement :

- la répondabilité du sens,
- l'orientation fiduciaire de la circulation,
- l'assomption métabolique des coûts.

*

Annexe I — L'ASI comme configuration-limite : seuil de non-gouvernabilité architectonique

Statut de l'annexe

Cette annexe examine une hypothèse-limite souvent désignée sous le terme d'*Artificial Superintelligence* (ASI) : une intelligence capable de réaliser la plupart des tâches intellectuelles humaines à un niveau égal ou supérieur, de manière autonome et généralisée.

Il ne s'agit ni d'un scénario prospectif, ni d'une spéculation catastrophiste, ni d'un débat sur la probabilité empirique d'une telle émergence. L'objectif est strictement méthodologique : montrer comment, dans le cadre architectonique proposé par cet ouvrage, l'ASI constitue une configuration critique mettant à l'épreuve les conditions mêmes de gouvernabilité de l'intelligence artefactuelle.

L'hypothèse d'une ASI n'ajoute pas un régime supplémentaire. Elle intensifie jusqu'au point de rupture les trois régimes constitutifs : sens, communication, métabolisme. [38][39][40]

I.1 Régime du sens : répondabilité et rupture d'endossement

Une ASI, par définition, produirait des énoncés, plans, décisions ou découvertes à une vitesse et une profondeur excédant la capacité humaine de compréhension et de contestation situées.

Dans une telle configuration, le seuil de répondabilité tend à être franchi structurellement :

- les productions demeurent opératoires,
- mais aucune instance humaine identifiable ne peut raisonnablement en répondre,
- les dispositifs de reprise deviennent tardifs ou inapplicables.

La pathologie correspondante est celle d'un sens sans endossement : le sens continue d'être produit, mais cesse d'être assignable.

I.2 Régime de la communication : saturation fiduciaire et perte d'orientation

Une ASI n'agirait pas seulement comme productrice de contenus, mais comme agent de reformulation, conversion et accélération des flux communicationnels.

La fiducia cesse alors d'être orientable :

- la circulation devient plus rapide que toute possibilité de discussion,
- l'autorité se cristallise par défaut autour du système,
- les chaînes humaines de contestation sont saturées.

La pathologie typique est celle d'une communication sans orientation, où l'espace même du dissensus se referme.

I.3 Régime métabolique : dette entropique explosive et désassomption radicale

C'est dans le régime métabolique que la tension devient décisive.

Une ASI autonome supposerait une intensification massive des infrastructures computationnelles, énergétiques et organisationnelles, entraînant :

- une croissance accélérée de la dette entropique située,
- une externalisation toujours plus diffuse des charges,
- une irréversibilité rapide des engagements matériels.

Plus radicalement, une ASI pourrait redéfinir elle-même ce qui compte comme coût acceptable, rendant obsolètes les seuils humains d'assomption.

La pathologie devient alors une désassomption métabolique complète : le système fonctionne, mais hors de toute soutenabilité gouvernable.

I.4 Verdict architectonique : fin des conditions de l'intelligence artefactuelle gouvernable

Dans les termes de cet ouvrage, l'intelligence artefactuelle n'est pas une puissance maximale, mais un effet conditionnel de co-contrainte entre régimes irréductibles.

Or une ASI autonome et généralisée tend à dissoudre les conditions mêmes de cette co-contrainte :

- plus de réponsabilité praticable,
- plus de fiducia orientable,
- plus d'assomption métabolique collective.

L'ASI apparaît ainsi non comme un « degré supérieur » d'intelligence artefactuelle, mais comme **une configuration-limite où l'émergence d'une intelligence ingouvernable au sens humain deviendrait plausible.**

Conclusion de l'annexe

Cette annexe ne vise pas à prédire l'ASI, mais à formuler ce que son hypothèse révèle : certaines trajectoires d'autonomisation ne prolongent pas l'intelligence artefactuelle, elles en neutralisent les conditions architectoniques.

L'intérêt critique de ce cadre est précisément là : montrer que l'enjeu n'est pas seulement celui d'une puissance accrue, mais celui de la survie des points de reprise — sans lesquels l'intelligence cesse d'être une catégorie partageable, endossable et gouvernable.

Si elle est ignorée, la pathologie interrégime propre à l'ASI peut conduire à une totalisation de fait, pour devenir un processus auto-référentiel qui ferme la contestation et risque de rendre les arbitrages illisibles et la gouvernance humaine impraticable.

*

Annexe J — Tables de correspondances avec les trois essais initiaux

Cette annexe documente, sous forme de correspondances, la filiation entre les développements des Parties II–IV et les trois essais initiaux dont ils procèdent, qui correspondent aux régimes [sens](#), [communication](#), [métabolisme](#).

Elle n'a pas pour fonction de justifier les choix conceptuels, mais de rendre traçable le travail de déplacement, de recomposition et de systématisation : chaque ligne indique le segment concerné, la fonction argumentative qu'il remplit dans l'ouvrage, l'opérateur conceptuel mobilisé dans l'essai source, et l'ancrage explicite (section ou motif) dans ce texte de référence. Les correspondances doivent être lues comme des liens opératoires : un même opérateur peut soutenir plusieurs segments, et certaines sections de l'ouvrage peuvent croiser plusieurs opérateurs.

Régime du sens (Partie II). Le tableau met en regard le déplacement du sens comme contenu vers le sens comme condition d'endossement, et les opérateurs de réponsabilité, de stabilisation et d'architecturation mobilisés dans l'essai initial consacré au sens.

Régime de la communication (Partie III). Le tableau repère les déplacements qui conduisent du schème de la transmission à un milieu de circulation algorithmique : convertibilité des énoncés, textoralité, mémoire computationnelle et autorité en chaîne, ainsi que les figures-limites où la responsabilité se requalifie.

Régime métabolique (Partie IV). Le tableau relie l'analyse des conditions matérielles et organisationnelles de l'IA aux notions de flux, d'extractivisme computationnel et de dette. Dans l'ouvrage, l'opérateur d'« entropie » prolonge ce cadre sans s'y substituer ; l'essai initial mobilise principalement une description des flux, des asymétries et des seuils de soutenabilité.

*

Partie II – Régime du sens

Ancrage explicite dans la Partie II	Fonction argumentative	Opérateur correspondant dans le texte de référence	Ancrage explicite dans le texte de référence
Introduction : "Du sens comme contenu au	Pose le déplacement théorique central : le sens ne réside pas dans une propriété	Réponsabilité comme pivot du	Introduction (p. 9-12) : définition opératoire de la réponsabilité

Ancrage explicite dans la Partie II	Fonction argumentative	Opérateur correspondant dans le texte de référence	Ancrage explicite dans le texte de référence
sens comme condition d'endossement"	interne (compréhension, intention) mais dans des conditions pratiques d'endossement et de répondabilité	sens + Tournant architectural	comme "possibilité conjointe d'y répondre et d'en répondre" ; VI.1 (p. 45) : "Du producteur de contenus à l'architecte de situations de sens"
3.1 "De comprendre à répondre"	Opère le glissement conceptuel du régime de compréhension vers le régime de réponse : ce qui compte n'est pas l'état mental du système mais la possibilité d'interroger et d'assumer publiquement ce qui est produit	Déplacement de la question du sens (de comprendre à répondre) + Préface (chambre chinoise réinterprétée)	I.2 (p. 14-16) : "la crise n'est pas celle de la compréhension, mais de la répondabilité" ; Préface (p. 7-8) : "où est la répondabilité ?" plutôt que "où est la compréhension ?"
3.2 "Production massive sans intention ni expérience"	Décrit le régime factuel qui rend caduque l'approche classique : massification + absence d'ancrage intentionnel/expérientiel → dissociation entre performance discursive et engagement subjectif	Retour du problème du sens sans intention ni expérience + Mémoire computationnelle	I.1 (p. 13-14) : "production discursive massive sans intention ni expérience" ; II.2 opérateur 3 (p. 24) : mémoire computationnelle (stockage sans temporalité propre ni engagement)
3.3 "Le sens comme"	Reformule positivement le sens non comme contenu	Sens comme processus +	II.1 (p. 21) : "Du contenu au processus : le sens"

Ancrage explicite dans la Partie II	Fonction argumentative	Opérateur correspondant dans le texte de référence	Ancrage explicite dans le texte de référence
stabilisation opératoire"	fixe mais comme arrêt provisoire de l'interprétation suffisant pour orienter l'action, dépendant de médiations techniques et institutionnelles	Stabilisation provisoire	comme stabilisation provisoire" ; Introduction (p. 12) : "le sens est d'abord une architecture visant à maintenir la responsabilité"
4. Introduction "Le sens comme régime"	Présente le cadre théorique explicite (référence au texte Zenodo) et annonce le passage d'une ontologie du contenu à une analyse des processus hétérogènes de stabilisation	Régime du sens comme ensemble de processus + ancrage dans le texte de référence	Note de l'auteur (p. 3) : inscription dans l'approche centrée sur "conditions de production, circulation et tenue du sens" ; II (p. 19-27) : "Mutation du sens : de la signification au régime"
4.1 "Du contenu au processus"	Abandonne la conception substantialiste du sens au profit d'une analyse processuelle : le sens advient par des opérations (sélection, validation, contextualisation) et non par adéquation à une intention	Mutation du sens (du contenu au processus) + Opérateurs du régime	II.1 (p. 21) : "le sens comme stabilisation provisoire" ; II.2 (p. 22-26) : sept opérateurs généraux du régime du sens
4.2 "Opérateurs du régime du sens"	Identifie les mécanismes concrets (sélection, stabilisation temporelle, contextualisation,	Sept opérateurs généraux (convertibilité, compatibilisation, mémoire computationnelle,	II.2 (p. 22-26) : développement systématique des opérateurs ; Figure 1 (p. 27) : schéma

Ancrage explicite dans la Partie II	Fonction argumentative	Opérateur correspondant dans le texte de référence	Ancrage explicite dans le texte de référence
4.3 "Le tournant architectural"	validation, délégation) qui rendent certaines interprétations praticables plutôt que d'autres Pose que les opérateurs sont désormais inscrits dans des architectures techniques/organisationnelles explicites → le sens dépend de choix de <i>design</i> , d'implémentation et de gouvernance	visibilité calculée, autorité en chaîne, responsabilité, endossement) Tournant architectural + Architecture du sens	conceptuel du régime du sens II.3 (p. 27) : "Le tournant architectural" ; VI (p. 44-54) : "Architectures du sens : rendre la responsabilité praticable"
5. Introduction "Responsabilité et agentivité interprétative"	Introduit la responsabilité comme condition d'agencement de rôles/procédures permettant une réponse publique ("qui répond de cela ?"), distincte de l'attribution intentionnelle	Responsabilité comme pivot + Agentivité reconfigurée	V (p. 38-43) : "Responsabilité : le pivot du sens à l'ère des modèles" ; IV.2 (p. 36) : distinction des formes d'agentivité
5.1 "Délégation interprétative"	Analyse le processus par lequel une part du travail interprétatif (sélection, reformulation, hiérarchisation) est confiée aux dispositifs techniques → redistribution et	Délégation interprétative + Agentivité opératoire/interprétative/normative	IV.1 (p. 34-36) : "Déléguer une tâche vs déléguer une lecture" ; IV.2 (p. 36-37) : tripartition de l'agentivité

Ancrage explicite dans la Partie II	Fonction argumentative	Opérateur correspondant dans le texte de référence	Ancrage explicite dans le texte de référence
	fragmentation de l'agentivité		
5.2 "Imputabilité, responsabilité"	Distingue trois régimes normatifs : imputabilité (attribution causale/traçabilité), responsabilité (possibilité de mise en question du sens), responsabilité (obligations/conséquences juridiques)	Triptyque imputabilité/répondabilité/responsabilité	V.2 (p. 40-42) : développement du triptyque ; distinction entre attribution technique, possibilité de discussion et prise en charge des conséquences
5.3 "Autorité interprétative et crise de la discussion"	Montre que l'autorité ne disparaît pas mais se reconfigure : elle devient fonctionnelle (disponibilité, fluidité) plutôt que discursive → neutralisation de la contestation en amont	Autorité en chaîne + Visibilité calculée + Crise de la discussion	V.3 (p. 42-43) : "Autorité interprétative et crise de la discussion" ; II.2 opérateurs 4 et 5 (p. 24-25) : visibilité calculée et autorité en chaîne
6. Introduction "Architecture du sens : rendre la responsabilité praticable"	Déplace l'analyse vers la conception de dispositifs concrets organisant production/circulation/reprise : la responsabilité doit être construite, non présumée	Architecture du sens + Dispositifs de responsabilité	VI (p. 44-54) : chapitre entier ; I.5 (p. 18) : "articuler théorie du sens et conception de dispositifs de réception"
6.1 "Dispositifs à seuils multiples"	Propose des architectures à niveaux différenciés d'engagement (plausibilité opératoire, validation	Stratification et résolution + Dispositifs à seuils multiples	VI.2 (p. 46-47) : "Stratification et résolution" ; VI.3 (p. 48-50) : "Dispositifs à seuils multiples" ;

Ancrage explicite dans la Partie II	Fonction argumentative	Opérateur correspondant dans le texte de référence	Ancrage explicite dans le texte de référence
	contextuelle, engagement institutionnel) permettant circulation entre seuils		Figure 3 (p. 49) : régimes de sens comme seuils d'engagement
6.2 "Littératies du sens à l'ère des modèles"	Identifie les compétences nécessaires pour reconnaître les seuils, activer la réponsabilité et intervenir : désambiguïsation des statuts, intervention interprétative, conscience des effets cumulés	Littératies du sens	VI.4 (p. 50-51) : "Littératies du sens à l'ère des modèles" ; Figure 4 (p. 51) : schéma des littératies
6.3 "Écologie du sens réponsable"	Élargit la focale aux interactions entre dispositifs, acteurs, institutions, infrastructures à différentes échelles → le sens émerge d'un environnement relationnel, non d'instances isolées	Écologie du sens + Écologie de l'attention	VII.3 (p. 57-58) : "Écologie de l'attention et écologie du sens" ; VII (p. 54-58) : "Implications éthiques et politiques : vers une écologie du sens réponsable"
Conclusion de la Partie II	Synthèse du déplacement opéré : le sens n'est plus garanti par une intériorité cognitive mais peut être gouverné via des architectures dont nous avons la responsabilité	Sens comme architecture réponsable + Gouvernance de la réponsabilité	Conclusion (p. 58-59) : "Penser et pratiquer le sens autrement" ; VII.2 (p. 56-57) : "Gouvernance de la réponsabilité" ; I.4 (p. 17-18) : thèse centrale

Ancrage explicite dans la Partie II	Fonction argumentative	Opérateur correspondant dans le texte de référence	Ancrage explicite dans le texte de référence
	collective de conception		

Notes méthodologiques sur cette table :

1. Fonction argumentative : explicite le rôle de chaque section dans l'économie générale de la Partie II (poser un problème, opérer un déplacement, fournir un outil analytique, etc.)
2. Opérateur correspondant : identifie le concept, mécanisme ou distinction théorique du texte de référence qui est mobilisé, reformulé ou systématisé dans la Partie II
3. Ancrage explicite : fournit la localisation précise (chapitres, sections, figures, pages) dans le [texte de référence Zenodo](#), permettant la vérification de la filiation conceptuelle
4. Structure de la table : suit la progression linéaire de la Partie II (chapitres 3-6) tout en établissant des correspondances croisées avec l'ensemble du texte de référence, montrant que la Partie II ne se limite pas à une transposition mais opère une recomposition systématique

*

Partie III – Régime de la communication

Ancrage explicite dans la partie	Fonction argumentative	Opérateur correspondant dans le texte de référence	Ancrage explicite dans le texte de référence
§ "De la transmission à la circulation algorithmique" - La communication devient un milieu autonome, structuré par des logiques de convertibilité, de visibilité et de répétition	Pose la thèse centrale du déplacement du modèle transmission vers circulation	Convertibilité généralisée + Distribution	II.1 : "La circulation n'est plus secondaire par rapport à la production du sens ; elle en devient la condition première"

Ancrage explicite dans la partie	Fonction argumentative	Opérateur correspondant dans le texte de référence	Ancrage explicite dans le texte de référence
7.1 "Oralité primaire et imputabilité située" - La parole ne circule pas sans origine ; elle meurt avec celui qui la porte	Établit le régime de référence où stabilité = coïncidence entre autorité, mémoire et responsabilité	Régime du langage (oralité primaire)	I.1 : "La parole engage celui qui parle, non seulement sur le plan symbolique, mais souvent sur le plan juridique ou rituel"
7.2 "Écriture, mémoire et prescriptions de lecture" - L'écriture inaugure une première forme de circulation, mais une circulation contrôlée	Décrit le premier seuil de désindexation origine/énoncé via externalisation mémorielle	Mémoire documentaire + Prescriptions de lecture	I.2-I.3 : "L'écriture introduit la possibilité d'un souvenir externalisé" ; "Les supports imposent des régimes de lisibilité distincts"
7.3 "Littératies et seuils de transformation" - Le seuil contemporain introduit une rupture qualitative	Identifie le moment où la médiation devient constitutive (non plus ponctuelle)	Convertibilité généralisée + Instabilité structurelle	I.5 : "Ce changement de seuil affecte le rôle régulateur que jouaient jusqu'alors la situation d'énonciation et la stabilisation textuelle"
8.1 "De la médiation à la convertibilité généralisée" - Un énoncé n'est plus conçu comme une unité à transmettre, mais comme un ensemble de traits exploitables	Opère le basculement conceptuel central : médiation → convertibilité	Convertibilité généralisée (définition pleine)	II.1 : "La convertibilité généralisée rompt avec cette logique. Un énoncé n'est plus conçu comme une unité à transmettre, mais comme un ensemble de traits exploitables"
8.2 "Textoralité opérationnelle" - Forme hybride structurée par des opérations	Nomme le milieu résultant de la convertibilité	Textoralité	II.2 : "La textoralité ne désigne ni un hybride stable entre oral et écrit, ni une nouvelle forme

Ancrage explicite dans la partie	Fonction argumentative	Opérateur correspondant dans le texte de référence	Ancrage explicite dans le texte de référence
algorithmiques de traitement et de génération			linguistique identifiable. Elle renvoie à un milieu opératoire"
8.3 "Pragmatique de la parole machinique" - Elle parle sans avoir parlé, répond sans avoir entendu, produit du sens sans le viser	Décrit le paradoxe constitutif de l'énonciation dé-située	Oralité tertiaire + Dé-situation de l'énonciation	VI.1 : "La parole qui circule [...] ne renvoie plus nécessairement à une situation originaire. Elle est produite à la demande, ajustée en temps réel"
9.1 "Mémoire computationnelle : persistance et oubli" - La persistance des énoncés n'y est jamais absolue. Elle dépend de paramètres techniques	Expose le régime mémoriel où conservation ≠ accessibilité	Mémoire computationnelle + Invisibilisation	III.1-III.2 : "La mémoire n'est plus organisée autour de documents ou d'archives au sens classique, mais autour de fonctions : recherche, tri, classement, recommandation"
9.2 "Autorité en chaîne et asymétrie structurelle" - L'autorité devient cumulative et différentielle [...] Elle ne repose pas sur un acte fondateur, mais sur une série de confirmations partielles	Décrit la fragmentation de l'autorité et la dilution de l'imputabilité	Autorité en chaîne + Asymétrie structurelle	IV.1-IV.2 : "L'autorité devient ainsi une fonction distribuée plutôt qu'une propriété attachée à une source unique" ; "Cette dissymétrie se manifeste d'abord dans le contrôle de la visibilité"
9.3 "De la crise épistémique à la crise fiduciaire" - Le problème central n'est	Opère le déplacement diagnostique majeur :	Crise fiduciaire (vs crise épistémique)	V (titre) + V.1 : "Le cadre dominant appréhende la situation présente

Ancrage explicite dans la partie	Fonction argumentative	Opérateur correspondant dans le texte de référence	Ancrage explicite dans le texte de référence
pas seulement la production de faux énoncés, mais la reconfiguration des conditions de la confiance	vrai/faux → crédit/reconnaissance		comme une bataille de contenus [...] Ce cadre est nécessaire, mais il est insuffisant"
10.1 "Oralité tertiaire et transformation de la parrhêsia" - La parole peut produire des effets de dévoilement sans exposition réelle de l'énonciateur	Examine la dissolution du lien parole/risque/courage	Parrhêsia + Oralité tertiaire	VI.4 : "L'avènement de l'oralité tertiaire fragilise précisément ce couplage [entre énoncé, auteur et scène publique]"
10.2 "Palimptextualité et circulation sans origine stable" - Chaque énoncé est potentiellement le résultat d'une superposition de traces [...] L'origine n'est pas supprimée ; elle devient indifférente	Décrit la structure stratifiée des énoncés contemporains	Palimptextualité + Stratification palimptextuelle	VIII.1-VIII.3 : "Les énoncés qui circulent aujourd'hui ne sont plus simplement reliés à d'autres textes ; ils sont stratifiés"
10.3 "Textoralité : milieu hybride de la circulation" - Les énoncés sont immédiatement intégrés dans des chaînes d'opérations qui les transforment en ressources exploitables	Définit la textoralité comme milieu (non plus seulement comme régime sémiotique)	Textoralité (comme milieu) + Écologie sociotechnique	II.2 : "La textoralité comme milieu du langage contemporain" ; IV.4 : "Écologie sociotechnique du langage"
10.4 "Textaautorité et responsabilité requalifiée" - Être responsable ne signifie	Propose la reconfiguration de la fonction-auteur	Textaautorité + Endossement	IX.2 : "L'auteur devient une fonction relationnelle inscrite dans des trajectoires

Ancrage explicite dans la partie	Fonction argumentative	Opérateur correspondant dans le texte de référence	Ancrage explicite dans le texte de référence
plus être la source première, mais intervenir dans la configuration du milieu communicationnel			d'énoncés, plutôt qu'un point de départ absolu"
Conclusion de la Partie III - Les figures-limites analysées [...] mettent en tension les catégories héritées de la responsabilité, du sens et de la confiance	Synthèse : les figures-limites révèlent les zones critiques du régime	Figures-limites (notion transversale)	VI-X (ensemble) : les différentes "ontologies" et "cas-limites" analysés comme révélateurs du régime

Notes méthodologiques sur la table :

1. Densité conceptuelle : La Partie III condense en ~15 pages ce que le texte de référence développe sur 70+ pages, d'où des correspondances souvent multiples pour un même passage.
2. Progression argumentative : La table suit la logique interne de la Partie III (généalogie → opérateurs → figures-limites) qui trouve son équivalent structurel dans l'organisation I-II-III-IV / V / VI-VIII-IX du [texte de référence](#).
3. Concepts-pivots : Certains termes clés (convertibilité, textoralité, palimptextualité, textautoralité) apparaissent à la fois comme opérateurs analytiques ET comme objets d'analyse, ce qui reflète leur statut transversal dans les deux textes.
4. Niveaux d'abstraction : La Partie III opère à un niveau de généralité plus élevé (« régime de la communication »), tandis que le texte de référence descend dans des analyses plus détaillées (ex : voix synthétique, traduction automatique, multilinguisme).

*

Partie IV – Régime métabolique

Ancrage explicite dans la Partie IV	Fonction argumentative	Opérateur correspondant dans le texte de référence	Ancrage explicite dans le texte de référence
11.1 "L'IA comme système de flux" — quatre flux :	Établir que l'IA n'est pas un objet statique	II. Décomposer le métabolisme	"Le métabolisme des modèles est structuré par des

Ancrage explicite dans la Partie IV	Fonction argumentative	Opérateur correspondant dans le texte de référence	Ancrage explicite dans le texte de référence
données, énergétiques, travail humain, organisationnels	mais une configuration dynamique traversée par des circulations hétérogènes continues	des modèles — catabolisme (extraction, compression) et anabolisme (capacité, institutions)	processus différenciés [...] Il transforme en détruisant, mais aussi en produisant" (p. 19)
11.1.a "Flux de données : extraction, annotation, maintenance" — collecte massive, annotation, perte irréversible d'information	Montrer que les données ne sont pas un stock disponible mais un flux continu avec coût informationnel structurel	II.1 Catabolisme informationnel I — extraction, compression, pertes structurelles	"L'extraction [...] opération active de captation [...] déjà transformer le monde social en matière première informationnelle" (p. 19-20)
11.1.b "Flux énergétiques et computationnels" — consommation électrique, dissipation thermique, contraintes physiques	Ancrer la computation dans ses limites physiques irréversibles (dissipation thermique)	II.2 Catabolisme matériel — énergie, eau, infrastructures, territoires	"Le calcul est un processus catabolique au sens strict : il dégrade de l'énergie de haute qualité en chaleur de basse qualité" (p. 21-22)
11.1.c "Flux de travail humain" — annotation, modération, supervision, maintenance sociale	Rendre visible le travail humain comme condition structurelle (non périphérique) du fonctionnement	Annexe B — Le travail social invisibilisé	"Le travail social n'est pas un contexte extérieur à l'IA, mais l'une de ses conditions matérielles de fonctionnement" (p. 102)
11.1.d "Flux organisationnels" — coordination,	Identifier le métabolisme institutionnel comme	II.4 Anabolisme socio-institutionnel	"L'anabolisme socio-institutionnel produit alors des

Ancrage explicite dans la Partie IV	Fonction argumentative	Opérateur correspondant dans le texte de référence	Ancrage explicite dans le texte de référence
gestion des risques, arbitrages	condition de stabilité souvent sous-estimée	— artefacts, décisions, normes, pouvoirs	normes [...] incorporées dans les procédures" (p. 25-26)
11.2 "L'extractivisme computationnel différé" — dissociation spatiale, temporelle, structurelle	Qualifier le régime de captation asymétrique où extraction et valorisation sont découplées	IV.1 Le régime extractif — puissance et asymétries constitutives	"Les asymétries constitutives : capter ici, bénéficier ailleurs [...] Ces asymétries ne sont pas des effets secondaires : elles sont constitutives" (p. 37-40)
11.2.b "Différenciation spatiale" — données collectées globalement, énergie mobilisée localement, valeur concentrée	Montrer que les lieux d'extraction et de valorisation sont structurellement dissociés	IV.1.4 Les asymétries constitutives	"Les corpus sont souvent extraits de productions distribuées, tandis que la valeur est concentrée dans quelques organisations" (p. 39)
11.2.c "Différenciation temporelle" — coûts différés, manifestation après coup	Établir que les coûts métaboliques se manifestent dans un futur incertain, hors du champ décisionnel initial	III.3 Accélération, saturation et pathologies métaboliques — temporalités désalignées	"L'accélération n'est plus un choix stratégique, elle devient une condition de survie" (p. 34-36)
11.3 "Accumulation, invisibilisation, externalisation" — croissance non linéaire, coûts	Identifier les trois mécanismes qui permettent la poursuite du régime malgré	V.1 Le métabolisme comme seuil de la responsabilité — grammaire	"Le métabolisme agit comme une grammaire silencieuse : il ne prescrit pas explicitement des

Ancrage explicite dans la Partie IV	Fonction argumentative	Opérateur correspondant dans le texte de référence	Ancrage explicite dans le texte de référence
rendus non pertinents, déplacement des charges	l'accumulation de passifs	silencieuse, inacceptabilité pratique	normes, mais il conditionne ce qui peut être envisagé comme raisonnable" (p. 55-56)
12.1 "Dettes entropiques situées : définition et portée" — accumulation irréversible de désordre matériel, énergétique, informationnel, organisationnel	Introduire l'opérateur unificateur qui pense l'irréversibilité structurelle du métabolisme	III. Temporalités et intensités métaboliques — métabolisme de base, actif, seuils incompressibles	"Le métabolisme de base représente le coût permanent de l'existence des modèles [...] certaines consommations ne peuvent pas être réduites sans transformer profondément les systèmes" (p. 28-29)
12.1 "Fondement conceptuel : irréversibilité et accumulation" — effets qui ferment des options plutôt qu'ils n'en ouvrent	Établir que certaines opérations modifient durablement les conditions de possibilité futures	Postface — Le désalignement des métabolismes — irréversibilité temporelle	"Ce qui ne peut pas être métabolisé par les sujets humains dans des conditions soutenables devient opaque, ou violent" (p. 92-93)
12.2 "Typologie des seuils entropiques" — thermodynamique, informationnel, organisationnel, social	Distinguer quatre types de points de rupture où les mécanismes de régulation cessent de fonctionner	V.1.3 Pourquoi ce seuil est spécifique à l'IA contemporain — combinaison automatisée	"Le seuil métabolique de la responsabilité est donc plus bas [...] que dans d'autres contextes technologiques" (p. 55)

Ancrage explicite dans la Partie IV	Fonction argumentative	Opérateur correspondant dans le texte de référence	Ancrage explicite dans le texte de référence
		/délégation/fluidité	
12.2.a "Seuil thermodynamique" — impossibilité de refroidir, saturation énergétique, conflits d'usage	Identifier le moment où les contraintes physiques deviennent dominantes et non compensables	II.2 Catabolisme matériel — seuils physiques, limites territoriales	"Le catabolisme matériel de l'IA s'inscrit ainsi dans des écosystèmes locaux, dont il modifie les équilibres" (p. 22)
12.2.b "Seuil informationnel" — saturation, dégradation contextuelle, asymétrie production/correction	Nommer le point où la production informationnelle génère du bruit plutôt que du sens	II.1 Catabolisme informationnel — compression, pertes structurelles, homogénéisation	"Ces pertes ne sont pas accidentelles ; elles sont structurelles. Les contextes d'énonciation disparaissent" (p. 20)
12.2.c "Seuil organisationnel" — complexité > capacités de supervision, fragmentation des responsabilités	Identifier le moment où le système devient ingouvernable (non instable techniquement, mais politiquement)	III.3.3 Saturation organisationnelle — épuisement des médiations humaines	"Cette saturation organisationnelle a des effets cumulatifs : épuisement professionnel, rotation rapide des équipes, perte de mémoire institutionnelle" (p. 35)
12.2.d "Seuil social" — épuisement, inégalités, perte de confiance, acceptabilité compromise	Nommer le point où la légitimité du système s'effondre indépendamment	VI.5 L'indifférence métabolique comme équilibre social — stabilisation	"L'indifférence devient ainsi un comportement coordonné sans coordination explicite" (p. 80-81)

Ancrage explicite dans la Partie IV	Fonction argumentative	Opérateur correspondant dans le texte de référence	Ancrage explicite dans le texte de référence
	nt de sa performance	collective de l'indifférence	
12.3 "Irréversibilité, hystérésis, points de non-retour" — retour en arrière coûte plus cher, correction devient impossible	Établir que certaines décisions ferment durablement des options (asymétrie temporelle structurelle)	V.3 Le coût métabolique de la responsabilité — coûts irréductibles, inacceptabilité pratique	"La responsabilité ne peut être effective que si son coût est redistribué, mutualisé ou compensé" (p. 59-68)
12.3 "Hystérésis" — le chemin de retour diffère du chemin d'expansion	Montrer que revenir en arrière n'est pas inverser le mouvement (coût additionnel structurel)	IV.3.6 La frugalité comme régime transitoire — réduction ne suffit pas à organiser	"La frugalité ouvre un espace de réflexion plus qu'elle ne le clôt. Elle oblige à poser explicitement la question des limites" (p. 45-46)
12.3 "Points de non-retour" — réparation nécessiterait plus de ressources que disponibles, légitimité compromise	Nommer les situations où la correction devient matériellement ou politiquement impossible	V.1.6 Responsabilité et inacceptabilité pratique — seuil au-delà duquel continuer aggrave la dette	"À partir d'un certain seuil, ce qui était conçu comme une ressource devient une source de vulnérabilité" (p. 56)
Encadré 3 "Stargate : exemple de désassomption" — coûts de sortie > coûts de prolongation, prolongation contrainte	Illustrer empiriquement comment l'hystérésis produit un verrouillage (stranded assets, dépendances)	IV.1.7 De l'IA comme technologie à l'IA comme système matériel — infrastructures lourdes, dépendances	"À mesure que les centres de données deviennent visibles, ils rencontrent des seuils d'acceptabilité

Ancrage explicite dans la Partie IV	Fonction argumentative	Opérateur correspondant dans le texte de référence	Ancrage explicite dans le texte de référence
			politique et sociale" (p. 41)
13.1 "Qui paie la dette ?" — externalisation spatiale, temporelle, sociale ; asymétrie production/assomption	Transformer la question des coûts en conflit de répartition explicite (politique)	V.3.5 Qui paie ? La responsabilité comme conflit de répartition	"La question décisive n'est plus de savoir si la responsabilité de l'IA a un coût [...] mais qui devrait supporter ce coût" (p. 66-68)
13.2 "Responsabilité comme conflit politico-technique" — arbitrage sous contrainte, illusion du coût nul	Reformuler la responsabilité non comme propriété morale mais comme champ de tension structurel	V.4 L'illusion de la responsabilité à coût nul — promesse d'une responsabilité sans conséquence	"L'illusion de la responsabilité à coût nul masque une réalité plus inconfortable : assumer la responsabilité implique de renoncer à certaines promesses centrales" (p. 68-70)
13.3 "Rendre discutables les arbitrages" — arbitrage ≠ optimisation, discutabilité comme condition de gouvernabilité	Substituer à une logique d'optimisation silencieuse une logique de décision exposée	VII.2 Arbitrer plutôt qu'optimiser — gouvernance comme choix de régime	"Gouverner, ce n'est pas chercher le meilleur compromis technique, c'est assumer des choix conflictuels entre des valeurs, des intérêts et des temporalités incompatibles" (p. 84-86)
13.5 "La limite métabolique et la flèche temporelle"	Établir que le métabolisme n'est pas un	Postface — Le désalignement des	"Le métabolisme machinique n'a pas besoin

Ancrage explicite dans la Partie IV	Fonction argumentative	Opérateur correspondant dans le texte de référence	Ancrage explicite dans le texte de référence
— irréversibilité traverse l'ensemble de l'architecture	arrière-plan mais le lieu où s'inscrit l'irréversibilité fondamentale	métabolismes — asymétrie production/co mpréhension	d'attendre que le sens soit compris pour continuer à produire" (p. 92)
Table conceptuelle comparée — Seuil de désassomption matérielle (impossibilité d'assumer les coûts)	Synthétiser les trois seuils critiques (répondabilité, fiducie, désassomption) dans une logique comparative	Apostille III — Responsabilité, régulation et seuils métaboliques — coût et limites de la responsabilité à grande échelle	"La responsabilité devient alors l'opération impossible qui devrait relier un énoncé à un propriétaire, un acte à un décideur, une preuve à une procédure" (p. 97-98)

Notes méthodologiques sur cette table

1. Structure de correspondance : La table identifie les points d'ancrage explicites entre la Partie IV (régime entropique) et le texte de référence (régime métabolique non entropique), en montrant comment l'opérateur "dette entropique" unifie et radicalise des notions déjà présentes.
2. Fonction de l'opérateur entropique : Là où le texte de référence décrit des flux, coûts et seuils, la Partie IV introduit l'irréversibilité structurelle comme opérateur unifiant (dette entropique située).
3. Points de continuité :
 - Le catabolisme/anabolisme du texte de référence est repris comme flux hétérogènes (11.1)
 - L'extractivisme du texte de référence devient extractivisme computationnel différé (11.2)
 - Les seuils métaboliques du texte de référence sont formalisés comme seuils entropiques typologisés (12.2)
4. Points de rupture :

- La Partie IV introduit l'hystérésis et les points de non-retour (absents du texte de référence)
 - Elle formalise la désassomption comme concept distinct du simple "échec de responsabilité"
 - Elle établit une flèche temporelle explicite (irréversibilité ≠ simple accumulation de coûts)
5. Portée architectonique : la table montre que la Partie IV ne contredit pas le [texte de référence](#) mais le prolonge conceptuellement en rendant explicite ce qui y restait implicite : l'irréversibilité fondamentale du métabolisme comme limite temporelle de l'intelligence artefactuelle.

*

Glossaire

Ce glossaire des notions architectoniques rassemble les notions centrales mobilisées dans l'ouvrage autour des régimes, des opérateurs, des interfaces, des seuils. Il ne vise pas l'exhaustivité encyclopédique, mais la stabilisation des termes opératoires par lesquels l'intelligence artefactuelle devient pensable, critiquable et gouvernable.

Arbitrage — Décision située par laquelle une configuration explicite ce qu'elle privilégie, limite ou sacrifie entre régimes irréductibles — sens, communication, métabolisme. Un arbitrage est gouvernable lorsqu'il reste visible, discutable et opposable ; il devient pathologique lorsqu'il est dissous dans des scores, des principes généraux ou des automatismes, et se traduit alors par une désassomption.

Architectonique — Cadre d'analyse qui décrit une configuration tenue ensemble sans principe unificateur. L'architectonique ne vise ni la synthèse ni l'équilibre, mais la reconnaissance des tensions irréductibles et des conditions minimales de gouvernabilité.

Autorité distribuée — Cristallisation locale et cumulative de la fiducia. L'autorité n'est pas un fondement, mais un effet stabilisé de chaînes de médiation.

Co-contrainte (tension maintenue vs équilibre résolu) — Relation structurelle par laquelle des régimes hétérogènes se limitent mutuellement sans se neutraliser. Condition minimale de gouvernabilité.

Convertibilité généralisée — Capacité structurelle des systèmes à transformer indéfiniment du texte : résumer, reformuler, agréger, redistribuer, en accélérant la circulation au-delà des capacités de reprise.

Désassomption — Situation où les effets continuent d'être produits, mais où leurs coûts ou responsabilités sont externalisés, déplacés ou différés hors de toute chaîne d'endossement, et où plus aucun acteur, collectif ou institutionnel, n'est en mesure (ou en position) d'en assumer pleinement les conséquences.

Dettes entropiques situées (irréversibilité matérielle/temporelle) — Opérateur du régime métabolique. Accumulation irréversible de coûts matériels, organisationnels et énergétiques, déposés quelque part, sur quelqu'un, dans le temps.

Effet d'intelligence — Stabilisation reconnue d'une opération artefactuelle lorsque le sens demeure endossable au point d'usage, que la circulation reste fiduciaire et contestable, et que les conditions métaboliques de reproduction sont assumables.

Fiducie communicationnelle (circulation orientable vs confiance aveugle) — Opérateur du régime de la communication. Condition collective de confiance orientable : pouvoir s'en remettre à ce qui circule sans perdre la possibilité de contestation.

Gouvernance architectonique — Pratique consistant non à optimiser ou unifier, mais à maintenir la co-contrainte des régimes par l'institution de seuils, l'activation d'interfaces, et l'assomption d'arbitrages irréductibles.

Intelligence artefactuelle (effet vs propriété) — Nom donné à l'intelligence non comme faculté intrinsèque d'un système, mais comme effet émergent, conditionnel et réversible, produit par la co-teneur des régimes du sens, de la communication et du métabolisme.

Interface — Dispositif de traduction par lequel des exigences interrégimes deviennent opérables sans être réduites. Une interface n'est pas un régime, mais une médiation entre régimes.

Interface éthico-normative — Interface hybride articulant principes et obligations situées, afin d'éviter la dissociation entre morale abstraite et conformité procédurale.

Interface éthique — Dispositif d'explicitation des valeurs et des conflits normatifs orientant les arbitrages : dignité, justice, sobriété, reconnaissance.

Interface juridico-normative — Institution de seuils opposables : obligations, droits, sanctions, procédures, assignation formelle des responsabilités.

Interface technique — Traduction opérable de contraintes : traçabilité, quotas, protocoles, ralentissement, mécanismes de suspension.

Irréductibilité — Dire des opérateurs qu'ils sont irréductibles revient à dire qu'aucun ne peut être converti en critère des autres : la répondabilité ne se déduit ni de la fluidité des circulations, ni de la soutenabilité métabolique ; la fiducie ne se déduit ni de l'endossement local du sens, ni de la seule contrainte des coûts ; et la dette entropique ne se résout ni par la contestation, ni par la requalification du sens.

Non-totalisation — Principe méthodologique selon lequel aucun régime, opérateur ou interface ne peut servir de critère souverain ou d'unification générale.

Opérateur — Fonction de structuration intermédiaire entre concept et indicateur. Un opérateur ne mesure pas : il rend repérables des seuils, des pathologies, et des zones d'intervention.

Pathologie interrégime — Configuration où un régime hypertrophié neutralise les autres : sens sans soutenabilité, communication sans orientation, métabolisme sans signification.

Point de reprise — Lieu institutionnel, organisationnel ou technique où un arbitrage peut être rendu visible, contesté et révisé. La gouvernabilité se reconnaît à la persistance de tels points de reprise.

Régime — Logique structurante irréductible (sens, communication, métabolisme) qui conditionne l'apparition de l'intelligence sans pouvoir être subsumée par un principe unique.

Régime de la communication — Régime dans lequel la question centrale est celle de la circulation des énoncés, de leur persistance, de leur visibilité et des formes distribuées de l'autorité.

Régime du sens — Régime dans lequel la question centrale n'est pas la compréhension interne, mais la possibilité d'endosser et de répondre du sens produit au point d'usage.

Régime métabolique — Régime des flux matériels, énergétiques, organisationnels et temporels qui soutiennent l'intelligence artefactuelle. Il introduit une contrainte irréversible : celle de l'entropie.

Renoncement gouvernant — Renoncement comme acte positif de gouvernance : acte légitime de suspension ou d'abandon lorsque la co-contrainte ne peut plus être tenue sans violence matérielle, sociale ou politique.

Répondabilité (endossement vs compréhension) — Opérateur du régime du sens. Capacité pratique et institutionnelle de répondre du sens produit : explicitation, correction, suspension, endossement.

Seuil — Point de bascule où une qualité cesse de tenir : sens non endossable, circulation non orientable, coûts non assumables.

Seuil de désassomption — Moment où les coûts métaboliques cumulés dépassent ce qui peut être assumé politiquement, matériellement ou institutionnellement.

Seuil de confiance — Moment où la densité des flux excède la capacité collective à orienter le crédit, discuter ou corriger ce qui circule.

Seuil de responsabilité — Moment où aucune entité ne peut raisonnablement répondre des effets produits au point d'usage.

Tenue — Stabilité toujours provisoire d'une configuration où les arbitrages entre régimes restent visibles, discutables et assumables.

Textoralité — Milieu hybride de circulation où les énoncés sont produits, reformulés et amplifiés indépendamment d'une origine stable.

Totalisation — Tentative de réduire des régimes irréductibles sous un principe unique de décision ou de mesure, en produisant une illusion d'unification et neutralisant la conflictualité constitutive de la gouvernabilité.

Traduction — Opération de médiation par laquelle une exigence conceptuelle (répondabilité, confiance, soutenabilité) devient une prise praticable. Traduire ne consiste pas à réduire un régime à un autre, mais à rendre praticables des tensions sans les neutraliser.

Conclusion du glossaire

Les termes ici stabilisés ne décrivent pas un vocabulaire extérieur à l'intelligence artificielle, mais les conditions mêmes sous lesquelles elle peut apparaître, être reconnue, contestée et gouvernée. Leur rôle n'est pas de fermer l'analyse, mais de fournir des prises minimales pour lire l'architectonique sans dissoudre ses tensions dans des catégories externes ou totalisantes.

Le présent glossaire vient en complément des glossaires détaillés proposés dans les deux volumes initiaux consacrés séparément au [régime du sens](#) et au [régime communicationnel](#). Là où ces glossaires développaient les concepts propres à chacun de ces régimes dans leur logique interne, le glossaire présent a pour fonction spécifique d'explicitier les notions transversales introduites par l'architectonique : co-contrainte, opérateurs, interfaces, seuils, désassomption, gouvernabilité.

Il ne se substitue donc pas aux élaborations conceptuelles précédentes, mais en constitue une zone de convergence : un outil de reprise destiné à soutenir la lecture de cette synthèse, sans prétendre ramener les régimes à un principe commun ni effacer les tensions qu'ils rendent visibles.

*

Bibliographie

- [1] **Bakhtine, Mikhaïl**, 1920–1924 (pub. posth.). *Toward a Philosophy of the Act*. Austin: University of Texas Press.
- [2] **Wittgenstein, Ludwig**, 1953. *Philosophical Investigations*. Oxford: Blackwell.
- [3] **Simondon, Gilbert**, 1958. *Du mode d'existence des objets techniques*. Paris: Aubier.
- [4] **Austin, J. L.**, 1962. *How to Do Things with Words*. Oxford: Oxford University Press.
- [5] **McLuhan, Marshall**, 1964. *Understanding Media: The Extensions of Man*. New York: McGraw-Hill.
- [6] **Leroi-Gourhan, André**, 1964–1965. *Le geste et la parole*. Paris: Albin Michel.
- [7] **Barthes, Roland**, 1967. « La mort de l'auteur. » (Essai).
- [8] **Georgescu-Roegen, Nicholas**, 1971. *The Entropy Law and the Economic Process*. Cambridge (MA): Harvard University Press.
- [9] **Grice, H. P.**, 1975. « Logic and Conversation. » In *Syntax and Semantics 3: Speech Acts* (Cole & Morgan, eds.), pp. 41–58. New York: Academic Press.
- [10] **Goody, Jack**, 1977. *The Domestication of the Savage Mind*. Cambridge: Cambridge University Press.
- [11] **Prigogine, Ilya & Stengers, Isabelle**, 1979. *La Nouvelle Alliance: métamorphose de la science*. Paris: Gallimard.
- [12] **Baudrillard, Jean**, 1981. *Simulacres et simulation*. Paris: Galilée.

- [13] **Ong, Walter J.**, 1982. *Orality and Literacy*. London: Methuen.
- [14] **Foucault, Michel**, 1984. « Qu'est-ce qu'un auteur ? » In *Dits et écrits*. Paris: Gallimard.
- [15] **Ricoeur, Paul**, 1990. *Soi-même comme un autre*. Paris: Seuil.
- [16] **Ihde, Don**, 1990. *Technology and the Lifeworld: From Garden to Earth*. Bloomington: Indiana University Press.
- [17] **Cunningham, Ward**, 1992. « The WyCash Portfolio Management System. » OOPSLA '92 (Experience Report).
- [18] **Brandom, Robert**, 1994. *Making It Explicit*. Cambridge (MA): Harvard University Press.
- [19] **Stiegler, Bernard**, 1994–2001. *La technique et le temps* (vol. 1–3). Paris: Galilée.
- [20] **Verbeek, Peter-Paul**, 2005. *What Things Do: Philosophical Reflections on Technology, Agency, and Design*. University Park: Penn State University Press.
- [21] **Latour, Bruno**, 2005. *Reassembling the Social*. Oxford: Oxford University Press.
- [22] **Floridi, Luciano**, 2011. *The Philosophy of Information*. Oxford: Oxford University Press.
- [23] **Moore, Jason W.**, 2015. *Capitalism in the Web of Life*. London: Verso.
- [24] **Bratton, Benjamin H.**, 2015. *The Stack: On Software and Sovereignty*. Cambridge (MA): MIT Press.
- [25] **Union européenne**, 2016. *Règlement (UE) 2016/679 (RGPD)*. Journal officiel de l'Union européenne.
- [26] **Vaswani, Ashish** et al., 2017. « Attention Is All You Need. » *Advances in Neural Information Processing Systems (NeurIPS)*.
- [27] **van Dijck, José; Poell, Thomas; de Waal, Martijn**, 2018. *The Platform Society*. Oxford: Oxford University Press.
- [28] **Gillespie, Tarleton**, 2018. *Custodians of the Internet*. New Haven: Yale University Press.
- [29] **Zuboff, Shoshana**, 2019. *The Age of Surveillance Capitalism*. New York: PublicAffairs.
- [30] **Bender, Emily M.** et al., 2021. « On the Dangers of Stochastic Parrots. » *FACt '21*.
- [31] **UNESCO**, 2021. *Recommendation on the Ethics of Artificial Intelligence* (adoptée en 2021).
- [32] **Crawford, Kate**, 2021. *Atlas of AI*. New Haven: Yale University Press.
- [33] **Bai, Yuntao** et al., 2022. « Constitutional AI: Harmlessness from AI Feedback. » *arXiv preprint*.
- [34] **NIST**, 2023. *Artificial Intelligence Risk Management Framework (AI RMF 1.0)* (NIST AI 100-1).

- [35] **ISO/IEC**, 2023. *ISO/IEC 42001:2023 — Artificial intelligence management system*. (Norme).
- [36] **Conseil de l'Europe**, 2024. *Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law* (ouverte à la signature le 5 septembre 2024).
- [37] **Union européenne**, 2024. *Règlement (UE) 2024/1689 (Artificial Intelligence Act)* (adopté le 13 juin 2024 ; publié au JO le 12 juillet 2024).
- [38] **Le Ray, Jean-Marie**, 2026a (janvier). *Le nouveau régime communicationnel de l'humanité*. [Zenodo](#).
- [39] **Le Ray, Jean-Marie**, 2026b (janvier). *Pour une théorie et une pratique du sens à l'ère des modèles*. [Zenodo](#).
- [40] **Le Ray, Jean-Marie**, 2026c (janvier). *Sens, communication et métabolisme : les trois régimes de l'IA* (troisième volet sur le régime métabolique de l'IA). [Zenodo](#).