

LE CANAL ET LES ATELIERS

*De l'ontologie processuelle
aux gestes du métier*

David Bories

Le canal et les ateliers — De l'ontologie processuelle aux gestes du métier

Conçu et dirigé par David Bories — Albi, France.
Rédigé en utilisant l'intelligence artificielle générative.

Octobre 2026.

0.4.0-poche

contact@artisanat-de-la-donnee.fr

Table des matières

Partie I — Le concept au travail	1
Chapitre 1 — Rappel des fondements de la trilogie	1
Chapitre 2 — Lire la donnée comme un flux : grille générale aux trois temps	7
Partie II — Cas historiques	13
Chapitre 3 — Atlas d'illustrations historiques	13
Partie III — Études de cas sectorielles	21
Note liminaire — La structure-type des études sectorielles	21
Chapitre 4 — Qualité de l'air urbain	22
Chapitre 5 — Santé hospitalière	23
Chapitre 6 — Sécurité et justice	24
Chapitre 7 — Climat et environnement	25
Chapitre 8 — Transport physique et mobilité	27
Chapitre 9 — Éducation et apprentissage	29
Ce que la traversée sectorielle a montré	30
Partie IV — Vivre avec le LLM	33
Chapitre 10 — Le LLM comme saut systémique de réception, et les pratiques de restauration	33
Partie V — Boîte à outils	39
Chapitre 11 — Checklists, grilles, fiches de signature, exercices .	39
Partie VI — Formes économiques et politiques	45
Note liminaire — Cartographier, non militer	45
Chapitre 12 — Le droit en évolution	46
Chapitre 13 — Les coopératives de données	47
Chapitre 14 — Souveraineté numérique et communs institutionnels	48
Chapitre 15 — Data activism et démocratisation	48
Ce que la cartographie a montré	49
Clôture — La carte révisable	51
Chapitre 16 — Statut de cet ouvrage et conditions de sa révision	51

Partie I — Le concept au travail

Rappel et orientation

Texte rédigé en mai 2026. Cette partie pose des repères conceptuels stables ; elle contient peu d'éléments datés. Les concepts qu'elle rappelle proviennent de la trilogie À l'épreuve du flux et ne se périment pas. Seules les illustrations mobilisées en chemin sont contingentes et révisables.

Chapitre 1 — Rappel des fondements de la trilogie

Pourquoi commencer par un rappel

Cet ouvrage est une trousse à outils. Mais un outil ne sert à rien si l'on a oublié ce qu'il sert à faire. Avant de décliner la grille processuelle en gestes — historiques, sectoriels, opérationnels, politiques —, il faut la tenir au clair. Ce premier chapitre n'apprend rien de neuf à qui a lu la trilogie. Il en reformule les repères, dans une langue dégagée de l'appareil démonstratif, pour qu'on puisse y revenir d'un coup d'œil en cours de lecture.

La trilogie *À l'épreuve du flux* a suivi la donnée le long de trois moments : sa fabrication, sa circulation, sa réception. Le premier tome, *Le fleuve et le canal*, a dit comment la donnée naît ; le deuxième, *De la source au torrent*, comment elle voyage ; le troisième, *La trace ou la donnée*, comment elle arrive — ou n'arrive pas — jusqu'à quelqu'un. Ces trois moments ne sont pas trois sujets distincts : ce sont les trois temps d'une même chose. On ne comprend la fabrication qu'en sachant que la donnée fabriquée est destinée à circuler ; on ne comprend la circulation qu'en sachant qu'elle vise une lecture ; et la lecture, à son tour, renvoie à ce qui a été fabriqué et transmis. Tout ce que cet ouvrage propose ensuite suppose ces trois temps tenus ensemble.

Le rappel qui suit les parcourt dans l'ordre. Il est volontairement resserré. À chaque concept, on indiquera de quel tome il provient, pour que le lecteur puisse, s'il le souhaite, remonter à la source.

Premier temps — la fabrication (tome 1)

Tout part d'un renversement. L'usage courant se représente la donnée comme un objet qui préexisterait à sa collecte : il y aurait, quelque part dans le monde, des données qui attendent qu'on les ramasse. Le premier tome a montré que cette image est fausse. Il n'y a pas, dans le réel, de données qui attendent. Il y a un **flux du réel** — ce qui se passe, en permanence, à toutes les échelles, indifférent à l'humain, que personne n'a creusé et que personne n'arrête. Et il y a des phénomènes qui ne deviennent données qu'en étant prélevés — c'est la voie observationnelle ; une donnée peut aussi être institutionnelle ou dérivée.

Prélever, c'est construire un **flux d'observation** : un flux que l'humain édifie à partir du flux du réel, par accumulation cadencée de prélèvements, selon des critères qu'il a choisis et dans une intention qui est la sienne. Le flux d'observation est fidèle au flux du réel relativement aux critères qui l'ont défini, sans jamais en être une réplique totale. C'est le maillon central de toute l'ontologie : la donnée est le grain élémentaire d'un flux d'observation. On ne cueille pas une donnée ; on creuse un canal dans le fleuve, et la donnée est l'eau qu'on en tire.

Produire une donnée demande deux opérations que la rapidité des dispositifs modernes a fondues, mais qui restent logiquement distinctes. La première est l'**acte d'inscription** : arrêter une valeur, décider qu'à cet instant cette grandeur continue vaut telle quantité — pratiquer une coupe dans un écoulement qui ne comporte de lui-même aucune coupure. La seconde est la **fixation** : installer durablement la valeur arrêtée sur un support.

De ces deux opérations naissent les **deux temps de la donnée**, et avec eux la résolution d'un paradoxe. Une donnée est figée — une fois inscrite, elle ne change plus — et pourtant elle vieillit. Comment une chose figée peut-elle vieillir ? Parce qu'elle n'a pas un temps mais deux. Dans le **temps de l'inscription**, elle est éternelle : la valeur inscrite reste ce qu'elle est. Dans le **temps du réel**, elle vieillit : le réel dont elle a été extraite continue de couler, et la donnée représente de moins en moins fidèlement un monde qui s'éloigne. Ces deux temps ne se contredisent pas, car ils ne mesurent pas la même chose.

De là découle la **fraîcheur** — la proximité entre l'instant de l'acte et le présent du réel — et sa loi la plus dure : la fraîcheur ne se rattrape jamais. L'acte d'inscription est unique, daté, non répétable. Pour disposer d'une donnée fraîche, il n'y a qu'un moyen : accomplir un nouvel acte, maintenant, sur le réel présent. Ce nouvel acte ne rafraîchit pas l'ancienne donnée ; il en crée une autre, à côté. La **péremption** est l'éloignement du temps du réel, à mesure que le flux continue de couler ; elle n'est pas une loi uniforme, mais une relation — la **validité pour un usage**, qui dépend de l'objet visé, de la question posée, de l'instant et de la tolérance admise.

Le premier tome a encore posé d'autres repères que cet ouvrage mobilisera. L'**optimum observationnel** : chaque flux d'observation a

un point d'équilibre propre entre le bénéfice qu'on tire du prélèvement et son coût, en deçà duquel on sous-prélève et au-delà duquel on sur-prélève. La **boucle** : certains flux d'observation transforment le réel qu'ils observent, de sorte que l'observation se réfléchit dans son objet. Le **saut systémique** : il arrive qu'un changement de régime technique ne fasse pas que déplacer un curseur, mais reconfigure la nature même du rapport à la donnée. Les **deux morts** de la donnée, enfin : la péremption, qui atteint sa validité représentative sans toucher à son existence, et la disparition, qui fait cesser la donnée d'exister — que le support se dégrade, que le format devienne illisible faute de machine pour le lire, ou que la donnée soit rendue définitivement inaccessible. Ces deux morts sont indépendantes et appellent des soins différents. On retiendra que la disparition n'est pas seulement la destruction physique du support : c'est toute extinction de la donnée comme donnée. Le troisième tome y reviendra avec un cas voisin mais distinct — la donnée morte, où le support survit intact mais où plus personne, nulle part, ne peut interpréter ce qu'il porte.

Sous tout cela, deux piliers que le premier tome a creusés jusqu'au socle. La donnée ne tient pas par elle-même : elle tient sur la **lisibilité du réel** — le flux, en s'écoulant, laisse des marques, il garde trace de lui-même, et c'est ce qui le rend prélevable — et sur l'**existence d'un lecteur** — le flux a produit, par son propre écoulement, un être capable de lire ces marques. La donnée naît à la rencontre exacte de ces deux piliers. Et la **trace** n'est pas la donnée : la trace est la marque que le réel laisse de lui-même, antérieure à toute lecture ; elle ne devient donnée que par un acte d'inscription. On retiendra ce point, car le troisième temps y reviendra.

Deuxième temps — la circulation (tome 2)

L'humain ne fabrique pas ses données pour les garder par-devers lui. Le mot le dit : *donnée* vient de *datum*, ce qui est donné — donné à quelqu'un, pour une occasion ultérieure. Une donnée inscrite et jamais transmise reste solitaire, comme une eau qui ne coule pas. Le deuxième tome part de là et en tire une thèse jumelle de celle du premier : **la donnée navigue parce qu'elle est fixée**. Loin de s'opposer, la fixité et le voyage se conditionnent — on ne transporte que ce qui tient.

Mettre une donnée en mouvement demande, là encore, un geste : l'**acte d'échange**, troisième geste fondateur après l'inscription et la fixation. Et ce geste déclenche un mouvement dont la nature est moins simple qu'il n'y paraît. Le **transport** n'est pas le déplacement d'une chose d'un point à un autre. C'est un **rapport entre deux déplacements** : celui de la donnée vers le lecteur, et celui du lecteur vers la donnée. Le transport est accompli quand la part de la donnée est faite et que la part du lecteur est rendue praticable — non quand la donnée est simplement « arrivée » quelque part. Cette définition, qui peut sembler abstraite, est décisive pour la pratique : elle interdit de croire qu'une donnée « mise à disposition » est, par ce seul fait, transmise. Tant que le lecteur ne peut pas accomplir sa part, le

rapport est empêché, et le transport n'a pas eu lieu.

Là où le transport concerne la donnée élémentaire, la **propagation** concerne l'agréé : la manière dont des flux entiers se répandent, se recourent, alimentent d'autres flux. Et le transport a des propriétés qui ne sont pas neutres. Sa **fidélité** : ce qui arrive peut être altéré en chemin. Sa **temporalité** : transporter prend du temps, et pendant ce temps la donnée continue de vieillir dans le temps du réel. Sa puissance la plus étrange, enfin : le transport **multiplie**. Faire voyager une donnée, c'est faire voyager un support qui la porte — ou produire ailleurs une copie, c'est-à-dire la faire porter par un second support sans cesser de la faire porter par le premier. La donnée n'est pas son support : le support est un bien rival — ce volume-ci, cette tablette-là, on ne peut l'avoir en deux lieux à la fois —, mais le fait qu'il porte ne l'est pas. Transporter une donnée ne la déplace donc pas, il la duplique — et la copie porte le même fait que l'original. Une donnée ne se cède pas comme une chose qui passerait de main en main ; **elle se partage comme un fait**. Celui qui la transmet ne s'en dépossède pas. Ce point n'a rien de propre au numérique : le scribe qui recopiait un texte, l'imprimeur qui tirait mille exemplaires multipliaient déjà les porteurs du même fait, sur autant de supports nouveaux. Le numérique n'a pas inventé la multiplication ; il l'a rendue quasi instantanée et presque gratuite, ce qui en a changé l'échelle, non la nature. Et ce point, d'apparence technique, fonde à lui seul une grande part de la dimension politique : on ne peut pas raisonner sur la propriété du fait porté comme sur celle du support qui le porte.

Le deuxième tome a aussi nommé un cas qui trompe l'intuition : les **sources stables**. Certaines données paraissent immobiles — la position d'une étoile, la profondeur d'un port. Elles ne le sont pas : leur réel coule, mais lentement, si lentement qu'au regard de nos usages la donnée semble ne pas vieillir. La stabilité n'est pas une propriété de la donnée ; c'est un **rapport de vitesses** entre l'écoulement du réel et le rythme de nos besoins. On « navigue sur place » tant que ce rapport tient — et il se dégrade dès que l'une des deux vitesses s'emballe.

C'est précisément ce qui définit le **torrent**. On serait tenté d'opposer deux flux : le flux du réel, naturel et pur, et le torrent de données, artificiel. Cette opposition est fautive, et la corriger est décisif : il n'y a qu'un seul réel qui s'écoule. Le torrent n'est pas un second flux à côté du premier ; c'est un **régime** de ce réel unique — le régime du **réel-marqué-par-nous**, celui où la cadence, l'étendue, la précision et la vitesse de circulation de nos prélèvements ont atteint un point tel que nous entretenons en permanence des centaines de millions de flux d'observation. Le torrent est une richesse et une menace à la fois : une puissance de lecture du réel sans précédent, et un débit qui peut submerger.

Il grossit par **auto-amplification** : la circulation accélérée produit des données qui appellent d'autres données, et la boucle creuse son propre lit. Face à quoi le deuxième tome ouvre un fil que le troisième reprendra — la **sagesse collective**. Le point doit être tenu avec une exactitude qui

en fait un repère et non une prise de position. S'abstenir de prélever, ou de circuler, est un acte qui engage : l'abstention n'est pas neutre. Mais — et c'est la nuance que le tome établit avec soin — l'abstention ne devient une position distincte, pesable, que là où le choix est *exposé*, rendu visible et délibéré collectivement. C'est un constat sur la structure du choix collectif, descriptif et valable en tout temps. Ce n'est pas une recommandation de s'abstenir, ni l'endossement d'un modèle plutôt qu'un autre. On y reviendra, car la confusion entre le constat et la prescription est l'écueil principal de la dimension politique.

Le tome s'achève sur un seuil : le moment où le débit, à force d'abondance, produit plus de données qu'aucun regard ne peut en suivre. C'est le **sur-débit** relatif à la fenêtre du lecteur. La circulation a fait tout ce qu'elle pouvait faire ; reste une part qu'elle ne peut accomplir à la place de personne. Cette part est l'objet du troisième temps.

Troisième temps — la réception (tome 3)

Le troisième tome part du bord du lecteur, là où le deuxième s'arrêtait. Sa thèse, jumelle des deux autres : **sans lecture, rien ne s'actualise**. Le premier tome avait montré l'aller — le réel dépose des traces, et un acte d'inscription, en relevant l'une d'elles, la constitue en donnée. Le troisième montre qu'une donnée que nul ne lit demeure constituée — non actualisée, ou muette pour ce lecteur ; elle ne retombe à l'état de trace qu'au cas limite de la donnée morte, quand plus aucun lecteur, nulle part, ne peut l'interpréter. Le passage de la donnée constituée à la donnée actualisée ne se produit que dans l'acte de lecture, et là seulement.

Cet acte demande, comme les trois autres, d'être nommé avec précision. L'**acte de lecture** est plus court que l'acte d'inscription. Là où inscrire mobilisait percevoir, distinguer, interpréter, viser et inscrire, lire ne mobilise que les trois premiers gestes — **percevoir, distinguer, déchiffrer** — et s'arrête là. Il ne vise pas, il n'inscrit pas : ce serait produire une donnée nouvelle, non lire celle qui est là. Une précision encore : l'acte de lecture ne s'applique qu'aux conventions inscrites — à ce qui a été déposé par un acte d'inscription humain. Il ne s'applique pas aux traces brutes du réel, qui relèvent d'un autre régime, celui de la lisibilité première posée par le premier tome.

Celui qui lit est un **lecteur situé**. Il n'est pas un point sans épaisseur, un regard de nulle part. Il a une position, des moyens, un code, une attention finie — une **fenêtre** bornée. Nulle lecture n'est de nulle part. Et cette fenêtre est finie face à un débit qui ne l'est pas : c'est le sur-débit vu du lecteur. La machine elle-même, quand elle « lit », n'échappe pas à cette condition : elle occupe une position située de reprise ou de traitement, avec ses propres bornes, et n'est lecteur effectif que si un sens s'actualise pour elle. Le troisième tome se garde de trancher trop vite l'asymétrie entre lecture humaine et lecture automatique ; il la pose comme une asymétrie à examiner, non comme une frontière nette. Ce

point commandera tout le chapitre 10.

Quand une donnée n'est pas lue, elle peut l'être pour trois raisons distinctes — les **trois chemins**, trois façons de ne pas actualiser —, et les distinguer est capital, car ils n'appellent pas la même réponse. **L'incapacité** : le lecteur ne *peut* pas lire, faute de code, de moyen, d'accès. La donnée est muette pour lui, peut-être vivante pour un autre. Ce chemin appelle un moyen — apprendre, ou recourir à un référent qui lit pour soi ; il est remédiable. Le **désintérêt** : le lecteur *pourrait* lire — il a le code, le moyen, l'accès — mais ne le fait pas ; le sur-débit le détourne, la donnée passe devant lui, il la voit sans la lire. Ce chemin n'appelle pas un moyen mais un rapport : l'attention qu'on accorde ou refuse, la probité de qui décide que ceci vaut d'être lu et que cela peut passer. La **péremption rapide**, enfin : la donnée n'a pas *attendu* d'être lue ; sa fraîcheur était si brève que, passé le premier instant, son utilité est quasi nulle, et elle passe en péremption avant même que la question de la lire se pose. Ce chemin n'appelle aucune réponse du lecteur : c'est un fait de la donnée, non un manquement de qui ne l'a pas lue. **Toutes les données non lues ne sont pas des échecs**. La péremption rapide est le sort normal d'une grande part des données ; le confondre avec un manquement serait une erreur de diagnostic lourde de conséquences pratiques.

Il faut ici tenir deux états voisins sans les confondre, car la pratique les rencontrera constamment. La **donnée muette** est relative et réversible : muette pour ce lecteur-ci, faute de moyen ou d'attention, elle reste lisible pour un autre, ou pour le même plus tard. La **donnée morte** est le cas limite où plus aucun lecteur, nulle part, ne peut l'interpréter : l'interprétation n'est plus seulement cachée chez un autre, elle est universellement perdue. Il ne reste que l'inscription sur le support. La donnée morte rejoint la disparition — la seconde mort du premier tome — non par défaillance du support, qui peut subsister intact, mais par perte universelle de l'interprétation. Muette : on peut encore y revenir. Morte : il ne reste qu'une trace.

Devant le sur-débit, le lecteur n'est pas démun, mais aucune voie ne le sauve entièrement. Le troisième tome nomme **trois postures**, chacune avec son coût. **Cadrer** : déplacer ou restreindre sa fenêtre, choisir délibérément ce qu'on regarde et ce qu'on laisse passer. **Déléguer** : hériter de la situation d'un autre lecteur — humain ou machine — qui lit à notre place, en assumant qu'on hérite alors de sa fenêtre et de ses bornes. **Accepter** : consentir avec lucidité à l'emportement, pour la part qu'on ne peut ni cadrer ni déléguer sans reste. Ces trois postures ne sont pas exclusives ; elles se combinent, et leur combinaison engage la **probité du lecteur** — l'éthique minimale de qui sait ce qu'il lit, ce qu'il ne lit pas, et ce qu'il délègue de sa lecture.

Les trois temps tenus ensemble

On peut maintenant resserrer l'essentiel en une phrase par temps. La donnée se **fabrique** par un acte qui l'inscrit et la fixe — dans sa voie observationnelle, un prélèvement du réel — et dès lors elle vieillit (tome 1). Elle **circule** par un acte qui la met en mouvement — et ce mouvement multiplie ses porteurs, la partage comme un fait sur autant de supports qu'on veut, sans déposséder qui la transmet (tome 2). Elle se **reçoit** par un acte qui la lit — et sans cette lecture elle demeure, constituée mais non actualisée (tome 3). Trois actes, trois temps, une seule donnée.

Cette tripartition n'est pas un classement commode : c'est la charpente de tout l'ouvrage. Chaque chapitre qui suit décline un ou plusieurs de ces temps dans un terrain concret. L'atlas historique éprouvera les trois temps sur des dispositifs pré-numériques. Les études sectorielles parcourront les trois temps, secteur par secteur, dans une même grille. La boîte à outils proposera, pour chaque temps, des objets maniables. La dimension politique se déploiera surtout sur le deuxième temps, la circulation, où se logent les questions de partage, de propriété et de commun. Et le chapitre sur le LLM se tiendra au point de jonction du premier et du troisième temps — un saut systémique de la fabrication qui se rejoue dans la réception.

Le chapitre suivant prend cette charpente et en fait une grille de lecture maniable : un petit nombre de questions, applicables à n'importe quelle donnée rencontrée, qui font passer le lecteur des trois temps abstraits à un réflexe d'examen.

Chapitre 2 — Lire la donnée comme un flux : grille générale aux trois temps

Du concept au réflexe

Le chapitre précédent a rappelé la grille. Celui-ci la rend maniable. Car connaître les concepts ne suffit pas : devant un chiffre, un tableau, un indicateur, un graphique, il faut pouvoir *réagir* — se poser, vite et dans le bon ordre, les questions qui comptent. L'objet de ce chapitre est de constituer ce réflexe : une grille générale, valable pour n'importe quelle donnée rencontrée, qui parcourt les trois temps sans rien oublier d'essentiel.

Cette grille n'est pas une norme. Elle ne dit pas ce qu'il faut conclure ; elle dit ce qu'il faut *interroger*. C'est un déclencheur de questionnement, fidèle au principe d'outillage de l'ouvrage : un outil ouvre une question, une règle dispense de penser. Deux praticiens appliquant cette grille à la même donnée n'aboutiront pas nécessairement au même jugement — et c'est normal, car chaque usage a son optimum propre et chaque lecteur sa fenêtre propre. Ce qu'ils auront en commun, c'est de n'avoir pas omis une question décisive.

La grille s'organise en trois familles de questions, une par temps. On peut la parcourir intégralement pour un examen approfondi, ou n'activer qu'une famille selon ce qui est en jeu. Mais l'ordre a un sens : on ne peut pas bien interroger la réception d'une donnée sans avoir d'abord situé sa fabrication et sa circulation.

Première famille — interroger la fabrication

Ces questions reconstituent le canal. Devant une donnée, elles demandent : d'où vient-elle, qui l'a creusée, et selon quels choix ?

La première chose à reconstituer est le **flux du réel** dont la donnée est tirée. De quel fragment du monde s'agit-il ? Une population, un marché, l'air d'une ville, le corps d'un patient ? Cette question paraît triviale ; elle ne l'est pas. Confondre la donnée avec le réel qu'elle décrit est l'écueil premier que toute la trilogie travaille à dissiper. Une donnée n'est jamais le réel ; c'est un prélèvement opéré sur lui, selon des critères.

Vient ensuite le **flux d'observation** : qui a creusé ce canal, et dans quelle intention ? Toute donnée porte la marque de qui l'a prélevée et du but qu'il poursuivait. L'intention n'est pas un défaut à corriger — elle est constitutive du flux d'observation, qui n'existe pas sans critères, et pas de critères sans visée. La question n'est donc pas « cette donnée est-elle neutre ? » (aucune ne l'est) mais « quelle intention a présidé à sa fabrication, et cette intention est-elle compatible avec mon usage ? ».

On interroge alors les **critères** : la cadence du prélèvement, son périmètre, l'instrument employé, les unités et les normes implicites. Un critère non examiné est un biais possible. Et l'on ajoute la question du **non-prélevé** : qu'est-ce qui n'a *pas* été capté, et qui pourrait être pertinent ? Le non-prélevé est souvent plus instructif que le prélevé, car il révèle les angles morts de l'intention amont. Là où l'on n'a pas mis d'instrument, on ne verra jamais rien — et l'on conclura à tort qu'il n'y avait rien à voir.

Deux questions de fabrication touchent enfin au temps et à la boucle. L'**âge fonctionnel** : quand le prélèvement a-t-il eu lieu, et la donnée est-elle encore fraîche au regard de la vitesse à laquelle son réel change ? Une donnée n'affiche pas sa péremption ; rien à sa surface ne dit depuis combien de temps le réel l'a quittée. Et la **boucle** : ce flux d'observation modifie-t-il ce qu'il observe ? Un score qui change le comportement qu'il mesure, une mesure qui déclenche une action qui altère le mesuré — autant de boucles dont l'ignorance fausse l'interprétation.

Deuxième famille — interroger la circulation

Une donnée rencontrée a presque toujours voyagé avant de nous parvenir. Ces questions remontent ce voyage.

La question maîtresse est celle de la **chaîne**. Peut-on remonter le transport jusqu'à la source ? Entre le prélèvement initial et la donnée que j'ai sous

les yeux, combien de mains, de copies, de retraitements ? Le deuxième tome a montré que le transport peut altérer la fidélité ; chaque maillon est un point où la donnée a pu être tronquée, recadrée, mal recopiée. Une donnée dont la chaîne n'est pas traçable jusqu'à la source est une donnée dont on ignore ce qu'elle a subi en chemin.

On interroge ensuite la **temporalité du transport** : combien de temps la donnée a-t-elle mis à arriver, et qu'a-t-elle vieilli pendant ce trajet ? Une donnée fraîche à la source peut être périmée à l'arrivée, si le transport a été lent au regard de la vitesse de son réel. Et l'on n'oublie pas la **multiplication** : cette donnée a-t-elle été partagée comme un fait — dupliquée à l'identique en de multiples lieux — ou tient-on l'unique exemplaire ? La réponse change la manière de raisonner sur sa fiabilité, sa propriété, sa diffusion : on ne sécurise pas un fait partagé comme on garde un objet unique.

Une question particulière mérite d'être posée à part : la donnée relève-t-elle d'une **source stable** ? Si son réel coule lentement, elle peut sembler immobile et fiable longtemps. Mais cette stabilité est un rapport de vitesses, jamais une propriété acquise : il faut vérifier que le réel observé n'a pas, depuis le prélèvement, changé de rythme. Une source longtemps stable peut cesser de l'être quand son réel s'emballe.

Troisième famille — interroger la réception

Les questions précédentes concernaient la donnée. Celles-ci concernent le lecteur — c'est-à-dire, le plus souvent, nous-mêmes. Elles sont les plus faciles à oublier, parce qu'on s'interroge volontiers sur ce qu'on lit, rarement sur les conditions dans lesquelles on le lit.

La première est celle de la **fenêtre**. Ai-je le moyen et le temps de lire vraiment cette donnée, ou va-t-elle me submerger ? Une donnée qu'on n'a pas la capacité d'examiner — faute de code pour la déchiffrer, faute d'attention disponible dans le débit — n'est pas lue, quoi qu'on en croie. La reconnaître permet de choisir une posture plutôt que de subir : **cadrer** (restreindre ce qu'on examine pour le lire bien), **déléguer** (s'en remettre à un lecteur capable, en assumant d'hériter de sa fenêtre), ou **accepter** (consentir avec lucidité à ce qu'on ne pourra pas tout lire). La pire option est la quatrième, non écrite : croire qu'on a lu ce qu'on n'a fait que voir passer.

La deuxième question diagnostique, si la donnée n'est pas lue, *pourquoi* elle ne l'est pas — selon les trois chemins. Est-ce une **incapacité** (il me manque le code, le moyen, l'accès — chemin remédiable, qui appelle d'apprendre ou de recourir à un référent) ? Un **désintérêt** (je pourrais, mais le débit me détourne — chemin qui n'appelle pas un moyen mais une décision d'attention) ? Ou une **péremption rapide** (la donnée n'a pas attendu, sa fraîcheur était trop brève — chemin qui n'appelle aucune réponse, car il n'y a là nul manquement) ? Ce diagnostic évite l'erreur la plus coûteuse en pratique : traiter comme un échec une donnée qui

n'avait pas vocation à être lue, ou inversement laisser passer comme normale une non-lecture qui était, elle, un vrai manquement.

La dernière question est celle de la **probité**. Pour mon usage précis, cette donnée est-elle adaptée — ou inadaptée ? Et si je m'apprête à la transmettre à mon tour, ou à fonder une décision dessus, suis-je en mesure de répondre de ce que j'en ai compris ? La probité du lecteur n'est pas une vertu décorative : c'est la condition pour que l'actualisation de la donnée, qui n'a lieu que dans la lecture, se fasse en conscience plutôt qu'en aveugle.

La grille en un coup d'œil

On peut récapituler les onze questions, regroupées par temps. Elles ne se lisent pas comme une procédure à dérouler mécaniquement, mais comme un aide-mémoire dans lequel puiser selon ce qui est en jeu.

Fabrication (tome 1)

1. De quel flux du réel cette donnée est-elle tirée ? (*ne pas confondre la donnée et le réel*)
2. Qui a creusé ce flux d'observation, dans quelle intention ?
3. Quels sont les critères — cadence, périmètre, instrument, unités, normes implicites ?
4. Qu'est-ce qui n'a pas été prélevé, et pourrait être pertinent ?
5. Quel est l'âge fonctionnel de la donnée, au regard de la vitesse de son réel ?
6. Le flux est-il dans une boucle — modifie-t-il ce qu'il observe ?

Circulation (tome 2)

7. La chaîne est-elle traçable jusqu'à la source — qu'a pu subir la donnée en chemin ?
8. Combien de temps le transport a-t-il pris, et qu'a-t-elle vieilli pendant ce trajet ?
9. La donnée a-t-elle été dupliquée, partagée comme un fait — ou tient-on l'unique exemplaire ?
10. Relève-t-elle d'une source stable — et ce rapport de vitesses tient-il toujours ?

Réception (tome 3)

11. Ai-je le moyen et le temps de la lire — et, sinon, est-ce incapacité, désintérêt ou péremption rapide ? Pour mon usage, est-elle adaptée, et puis-je répondre de ce que j'en comprends ?

Cette onzième entrée condense plusieurs questions parce que, dans la pratique, elles se posent d'un seul mouvement : reconnaître sa fenêtre, diagnostiquer la non-lecture éventuelle, juger l'adéquation à l'usage et engager sa probité sont les faces d'un même geste — celui par lequel un lecteur situé prend la responsabilité de ce qu'il lit.

Ce que la grille a appris, et ses limites

En la déclinant, on aura remarqué que les trois familles ne pèsent pas également selon les contextes. Une donnée scientifique de référence appelle surtout les questions de fabrication ; une donnée qui a transité par de nombreux intermédiaires appelle d'abord celles de circulation ; un flux d'information continu, qui nous submerge, appelle avant tout celles de réception. La grille est complète, mais son usage est sélectif : la compétence consiste à savoir, devant chaque donnée, quelle famille mobiliser en priorité.

Une limite doit être nommée d'emblée, car elle commande tout l'usage de la grille. Celle-ci aide à *poser* les bonnes questions ; elle ne fournit aucune des réponses. Elle ne dit pas quel âge fonctionnel est acceptable, ni quelle intention amont est légitime, ni quelle posture adopter face au sur-débit. Ces réponses dépendent du contexte, de l'usage, du lecteur — et c'est exactement pourquoi l'ouvrage qui suit décline la grille terrain par terrain plutôt que de prétendre la résoudre une fois pour toutes. Le concept éclaire la zone où il faut regarder ; il ne dresse pas la carte des récifs. La grille générale est cette lampe. Les chapitres suivants apprennent à s'en servir, là où l'on navigue vraiment.

Fin de la Partie I. Les parties suivantes déclinent la grille : l'atlas historique (Partie II), les études sectorielles (Partie III), le LLM (Partie IV), la boîte à outils (Partie V), les formes économiques et politiques (Partie VI), et la clôture méthodologique.

Partie II — Cas historiques

Atlas d'illustrations historiques de l'ontologie processuelle

Texte rédigé en mai 2026. Les cas traités sont historiques et ne se périment pas ; les études savantes qui les éclairent, en revanche, évoluent, et les rapprochements conceptuels proposés ici restent de l'ordre de l'illustration révisable.

Chapitre 3 — Atlas d'illustrations historiques

Pourquoi remonter avant le numérique

Le moment numérique a une faiblesse : il se croit sans lignée. Ses outils se présentent comme inédits, ses défis comme inouïs, sa donnée comme une invention récente. C'est une illusion, et elle a un coût pratique — elle empêche de voir que les difficultés d'aujourd'hui sont, pour la plupart, des difficultés anciennes simplement re-massifiées. Qui croit son problème inédit le résout moins bien que celui qui sait qu'on l'a déjà rencontré.

Les cas historiques rassemblés ici fonctionnent comme des **laboratoires conceptuels purs**. Sans la séduction technologique du présent — sans capteurs, sans bases de données, sans modèles —, la structure du phénomène apparaît à nu. Un curé qui tient son registre de baptêmes, un astronome babylonien qui note la position d'une planète, un naturaliste qui remplit son carnet : tous construisent des flux d'observation, opèrent des choix de critères, affrontent la péremption et le transport. Les concepts de la trilogie y sont à l'œuvre, dégagés du vernis technique qui, dans les cas contemporains, masque souvent leur fonctionnement.

Ces cas illustrent ; ils ne démontrent pas. Conformément à la posture de l'ouvrage, ils sont de l'ordre de la forme : révisables, contingents, choisis pour leur valeur pédagogique. Seul le concept engage. Et une prudence s'impose, qui vaut pour tout le chapitre : parler de « flux d'observation » à propos des Babyloniens, de « transport » à propos d'un portulan, de « lecteur situé » à propos d'un registre paroissial, c'est appliquer une grille moderne à des dispositifs qui ne se pensaient pas ainsi. On le fait

délibérément, pour révéler une structure — mais en sachant qu'on projette une grille, et que cette grille n'épuise pas ce que ces dispositifs furent pour ceux qui les tinrent.

Une grille de traitement homogène

Chaque cas est examiné selon une grille en huit points, pour que la comparaison d'un cas à l'autre reste possible. Les sept premiers points relèvent surtout de la fabrication ; le huitième, activé seulement quand le cas s'y prête, ouvre sur la circulation et la réception.

1. Le dispositif, en fait : qui, où, quand, comment.
2. Le flux d'observation construit.
3. Les critères, et l'intention qui les a choisis.
4. La probité, la cadence, le corpus.
5. La péremption et les deux temps, pour les cas longs.
6. Les conséquences sociales et politiques de la mise en mémoire.
7. Une mise en perspective propre au cas.
8. (*optionnel*) La circulation et la réception : comment la donnée a voyagé, qui l'a lue, ce qu'elle est devenue faute de lecture.

On traite ici huit cas, du plus ancien au plus récent, en activant le huitième point pour les quatre qui l'appellent le plus nettement.

Cas 1 — Les journaux astronomiques babyloniens

Du VIII^e au I^e siècle avant notre ère, des observateurs mésopotamiens ont tenu, sur tablettes d'argile, un relevé systématique des phénomènes célestes : positions des planètes, éclipses, météores, mais aussi niveau de l'Euphrate, prix des denrées, événements notables. Ces *journaux astronomiques* constituent le plus long flux d'observation systématique connu de l'Antiquité — plus de six siècles de continuité.

Le **flux d'observation** y est exemplaire de continuité : une cadence quotidienne, tenue par des générations successives d'observateurs, selon des critères remarquablement stables. Cette stabilité tenait à la nature de l'objet. Le ciel est le cas-type des **sources stables** du deuxième tome : son réel coule, mais à une lenteur telle qu'à l'échelle humaine il paraît immobile. Un astronome babylonien pouvait comparer une position planétaire à celle relevée des siècles plus tôt et les traiter comme commensurables — non parce que le ciel ne changeait pas, mais parce qu'il changeait infiniment plus lentement que les besoins de la prédiction. On « navigue sur place » sur six siècles quand le réel observé est aussi lent.

Les **critères** étaient choisis dans une intention précise : la prévision, à des fins divinatoires et calendaires. On ne notait pas tout du ciel ; on notait ce qui servait à présager et à régler le calendrier. Le **non-prélevé** est ici instructif : des phénomènes célestes sans valeur présagée n'entraient pas dans le flux, non par cécité mais par absence d'intention de les capter.

La **péremption** ne menaçait pas la donnée individuelle — une position relevée restait juste pour toujours dans le temps de l'inscription — mais sa valeur changeait de nature avec le temps. Fraîche, elle servait la prédiction ; ancienne, elle devenait matériau d'un corpus séculaire dont la valeur n'était plus présagée mais cumulative. La donnée babylonienne illustre ainsi la distinction des deux temps : éternelle dans le temps de l'inscription, elle a changé d'usage à mesure que le temps du réel s'éloignait, sans jamais devenir fausse.

Conséquence sociale : ce flux a soutenu une caste de spécialistes et une institution scribale, et sa très longue durée a rendu possible la découverte de régularités — périodes planétaires, cycles d'éclipses — qu'aucune observation courte n'aurait permis d'établir. La longueur du corpus n'était pas un sous-produit ; c'était la condition de la connaissance qu'on en tirait.

Cas 2 — Les registres paroissiaux

Du Moyen Âge à la Révolution française, les curés ont tenu registre des baptêmes, mariages et sépultures de leur paroisse. Ce flux d'observation démographique, institutionnel et décentralisé, a fini par couvrir des territoires entiers sur des siècles.

Le **flux d'observation** est ici socialement situé jusque dans sa définition. Le baptême marque l'entrée dans le flux ; la sépulture, la sortie. Mais ces seuils sont des conventions religieuses, non des faits biologiques bruts : l'entrée n'est pas la naissance, c'est le baptême. D'où un **non-prélevé** lourd de sens — les nouveau-nés morts avant le baptême, les marginaux hors paroisse, les fugitifs, les non-chrétiens. Ce que le registre ne capte pas n'est pas un oubli technique : c'est le négatif exact de l'intention qui l'a institué. Lire un registre paroissial sans voir son non-prélevé, c'est prendre la population baptisée pour la population réelle.

La **dynamique inter-flux** est remarquable : les registres paroissiaux se recoupent avec les registres seigneuriaux, fiscaux, notariaux. Croiser ces flux permet de reconstituer ce qu'aucun ne donne seul — c'est une propagation au sens du deuxième tome, où des flux distincts s'alimentent mutuellement.

Le huitième point s'impose ici, car le destin de ces registres illustre la **donnée constituée mais non actualisée** du troisième tome. Tenus d'abord pour un usage vivant — administratif, pastoral —, ils ont connu, à mesure que leur réel s'éloignait, une **péremption douce** : passé l'usage immédiat, ils ont cessé d'être lus comme des données courantes. Ils ne sont pas morts pour autant. Ils sont devenus corpus historique, relus des siècles plus tard par les démographes et les historiens de l'École des Annales. C'est un cas de **donnée muette** plus que de donnée morte : illisible comme donnée courante, le registre restait lisible comme source historique pour un autre lecteur, situé autrement. La non-lecture immédiate n'était pas un échec — elle relevait d'une péremption d'usage

—, et la relecture tardive l’a actualisé de nouveau, pour un autre usage. Non une trace qui attendait : une donnée restée constituée, qu’un lecteur d’un autre temps a actualisée.

Cas 3 — Les carnets de terrain des naturalistes

Aux XVIII^e et XIX^e siècles, les naturalistes voyageurs — Humboldt parcourant l’Amérique équinoxiale, Darwin embarqué sur le *Beagle* — ont tenu des carnets de terrain qui sont le cas paradigmatique du **flux d’observation idiosyncrasique** : un homme, un carnet, des critères qui évoluent au fil de l’expérience.

Contrairement au flux institutionnel des registres ou des journaux babyloniens, le carnet de naturaliste n’a pas de critères fixés d’avance par une institution. Ils se forment en chemin. Ce qui paraissait digne d’être noté au début du voyage ne l’est plus à la fin ; l’observateur apprend ce qu’il faut observer en observant. Le flux d’observation se construit donc en même temps que se construit l’intention qui le guide — illustration fine de l’idée que critères et visée sont indissociables, et qu’ils peuvent se former conjointement.

La question de la **probité** y est centrale et délicate. Un carnet de terrain est un flux d’observation difficilement vérifiable : l’observateur est souvent le seul témoin. Sa valeur tient à la confiance qu’on accorde à sa rigueur — à sa probité d’annotateur. Le **voisinage référentiel** y joue un grand rôle : un carnet vaut d’autant plus qu’il s’interconnecte avec d’autres, qu’il peut être recoupé, confronté, situé dans un réseau d’observations concordantes. La science naturaliste s’est construite sur cette interconnexion de flux idiosyncrasiques.

Conséquence majeure : ces carnets, longtemps tenus pour des supports de travail privés, sont devenus des monuments — relus, édités, scrutés pour ce qu’ils révèlent de la formation d’une pensée. Le carnet de Darwin n’est plus lu aujourd’hui pour ses observations zoologiques, largement dépassées, mais comme témoignage de l’élaboration d’une théorie. Là encore, un changement de lecteur a de nouveau actualisé, pour un autre usage, des carnets que la péremption des observations elles-mêmes avait fait cesser de lire.

Cas 4 — Les recensements romains

Sous la République puis l’Empire, Rome a tenu un recensement périodique — le *census* — à finalité fiscale et militaire : dénombrer les citoyens, évaluer leurs biens, fonder l’impôt et la conscription. Sa cadence était pluri-annuelle, organisée autour du *lustrum* qui séparait deux opérations.

Le recensement romain est le cas classique de la **boucle**. Le dénombrement détermine l’imposition ; l’imposition modifie le comportement déclaratif des recensés, qui ont intérêt à minorer ce qu’ils déclarent ; et cette minoration altère à son tour le flux d’observation. L’observation

transforme ce qu'elle observe, parce que les observés savent qu'ils sont observés et que l'observation a des conséquences sur eux. Toute donnée prélevée sous le regard d'un enjeu — fiscal, ici — porte la marque de cet enjeu.

La cadence pluri-annuelle illustre un arbitrage d'**optimum observationnel** dicté par le coût. Recenser un empire coûtait cher ; on ne pouvait le faire en continu. L'intervalle entre deux *lustra* était le prix d'un compromis entre la fraîcheur souhaitable et le coût supportable. Entre deux recensements, la donnée vieillissait — des citoyens naissaient, mouraient, s'enrichissaient, s'appauvrirent sans que le flux les suive. L'optimum n'était pas la fraîcheur maximale, économiquement hors d'atteinte, mais le meilleur compromis disponible.

Cas 5 — Les cartes portolanes

Du XIII^e au XVI^e siècle, les marins de Méditerranée ont disposé de cartes — les portulans — d'une précision côtière remarquable, construites non par une autorité unique mais par accumulation progressive des relevés de navigateurs. Le portulan est le cas paradigmatique de la **donnée consolidée** par propagation, et le huitième point y est central.

Une carte portolane synthétise des dizaines de flux d'observation individuels : chaque marin qui a longé une côte, mesuré une distance, noté un cap, a contribué un fragment. La carte n'est la donnée d'aucun observateur particulier ; elle est la consolidation de tous. C'est une **propagation** au sens du deuxième tome : des flux élémentaires s'agrègent en un flux de niveau supérieur, plus fiable que chacun de ses contributeurs.

Le portulan illustre surtout la **multiplication** — cette propriété du transport qui partage la donnée comme un fait sans en déposséder personne. Quand un port améliorait sa carte, la copie pouvait essaimer vers d'autres ports sans que le premier perde rien. Le savoir cartographique se propageait par duplication, non par cession : chaque copie était un original, chaque port pouvait à la fois recevoir et continuer d'enrichir. Cette **dynamique inter-entités** — chaque port améliore la carte que les autres utilisent — est exactement ce que le deuxième tome décrit quand il dit qu'on ne transporte pas une donnée comme un objet, mais qu'on la partage comme un fait. Le portulan montre que cette propriété, qu'on croit propre au numérique, est aussi ancienne que la copie manuscrite.

Cas 6 — Les tables de l'Académie del Cimento

Au XVII^e siècle, l'Académie florentine du Cimento a produit l'un des premiers flux d'observation thermométrique systématique, en relevant des températures à l'aide d'instruments calibrés. Ce cas illustre l'**émergence d'un saut systémique**.

Avant le thermomètre, on parlait du chaud et du froid en termes qualitatifs et relatifs : il fait plus chaud qu’hier, l’eau est tiède. Après, on mesure : l’eau est à tant de degrés. Ce passage n’est pas un raffinement du même geste — c’est un changement de régime. Il fait apparaître un type de donnée qui n’existait pas, et il exige une condition nouvelle : la **calibration**. Pour que deux relevés soient comparables, il faut que les instruments soient étalonnés sur un référentiel partagé. Le Cimento a buté sur cette exigence : sans échelle universelle, ses relevés restaient difficilement commensurables entre eux et avec ceux d’ailleurs.

Le saut systémique du thermomètre montre ainsi qu’un nouveau régime de fabrication de la donnée ne se réduit jamais à un instrument : il appelle l’**invention d’un référentiel commun**, sans lequel les données produites ne circulent pas — faute de pouvoir être lues hors du lieu qui les a produites. La donnée non calibrée est muette pour le lecteur distant : il la perçoit, il la distingue, mais il ne peut la déchiffrer, faute du code partagé qu’est l’échelle. Le saut de fabrication ne s’achève qu’avec un accord de réception.

Cas 7 — L’Encyclopédie de Diderot et d’Alembert

Entre 1751 et 1772, l’*Encyclopédie* a constitué un vaste corpus de connaissances organisé : un système de classement, des renvois, des rôles distincts. Ce cas est retenu pour une raison précise — il offre l’ancrage historique de la grille qui servira au chapitre 10 sur le LLM, et le huitième point y est décisif.

L’*Encyclopédie* met en œuvre une **donnée consolidée structurée** au sens classique, avec trois rôles encore nettement séparés. Il y a ceux qui produisent le savoir — les auteurs d’articles. Il y a ceux qui l’organisent et le rendent retrouvable — l’appareil de classement, les renvois, l’architecture d’ensemble. Et il y a ceux qui le restituent et le transmettent — l’ouvrage lui-même comme médiateur entre le savoir et le lecteur. Producteur, organisateur, restituteur : trois fonctions distinctes, tenues par des instances différentes, et cette distinction même est ce qui permet au lecteur de remonter d’un article à sa source, de contester un classement, de juger un rédacteur.

C’est exactement cette séparation que le chapitre 10 verra disparaître dans le saut systémique du LLM, où les trois fonctions fusionnent en une seule instance. L’*Encyclopédie* est ainsi le terme de comparaison historique qui rend lisible ce que la fusion contemporaine supprime : non pas le savoir lui-même, mais la séparation des rôles qui en garantissait le contrôle. On mesurera mieux ce que le LLM transforme en l’ayant d’abord vu fonctionner, séparé, deux siècles et demi plus tôt.

Cas 8 — L'horaire ferroviaire et le télégraphe

Au XIX^e siècle, le chemin de fer a posé un problème de coordination sans précédent : faire circuler des convois nombreux, lourds et rapides sur des voies finies — souvent uniques — que nul ne pouvait plus tenir à vue. Deux dispositifs y ont répondu : l'horaire, donnée imprimée qui fixe l'avenir des circulations, puis le télégraphe, qui fit pour la première fois voyager la donnée plus vite que le réel qu'elle décrit. C'est, un demi-siècle avant que le mot ne prenne son sens moderne, la coordination par la donnée — et le huitième point de la grille y est central.

Le réel à observer présentait une nouveauté radicale : il était **plus rapide que tout messager existant**. Jusqu'au rail, une information voyageait au mieux à la vitesse d'un cheval ou d'un navire — c'est-à-dire à la vitesse des choses qu'elle décrivait, ou guère plus. Le train, lui, distançait tout porteur : la donnée sur le train ne pouvait voyager que *dans* le train. La position d'un convoi, donnée dont la **fraîcheur** se compte en minutes, était donc structurellement périmée pour tout lecteur distant. Le secteur naissant affrontait un flux du réel que les régimes de transport existants ne pouvaient plus suivre.

La première réponse fut l'**horaire** — dont le guide de Bradshaw, compilant dès 1839 les tableaux des compagnies britanniques, devint l'emblème. L'horaire est une **donnée consolidée** d'un genre singulier : elle ne décrit pas le réel, elle le prescrit. Faute de pouvoir observer les trains en route, on renverse le rapport — on fixe d'avance leur circulation, et le réel est sommé de se conformer à la donnée. C'est la **boucle** portée à son comble : l'inscription ne reflète plus l'observé, elle le commande. Le dispositif a une grandeur et une faiblesse symétriques. Sa grandeur : il coordonne sans observer, par la seule autorité d'une donnée partagée. Sa faiblesse : il ne sait rien des écarts. Qu'un convoi prenne du retard, et tout le système devient aveugle — l'horaire continue de décrire un réel qui n'existe plus, et les collisions du temps des intervalles ont payé le prix de cette cécité.

Le **télégraphe** opère alors un **saut systémique**, et il n'est pas anodin qu'il soit né le long des voies — les premières lignes commerciales de Cooke et Wheatstone, à la fin des années 1830, couraient au flanc du Great Western Railway. Pour la première fois dans l'histoire, la donnée se déplace plus vite que la chose : le **transport** — ce rapport entre deux déplacements que le deuxième tome a posé — se dissocie du mouvement physique, et une position de train peut atteindre son lecteur *avant* de se périmer. Tout le régime de coordination en est changé : le cantonnement télégraphique fait de l'occupation d'une section de voie une donnée à fenêtre de validité brève, transmise et lue dans le temps même de sa fraîcheur ; la régulation des circulations par dépêche, mise au point dans les années 1850, permet de reprendre la main quand le réel s'écarte de l'horaire. La donnée prescrite et la donnée observée travaillent désormais ensemble — l'une fixe le plan, l'autre rattrape l'écart.

Comme au Cimento, le saut ne s'acheva qu'avec un **référentiel commun**

— et il fallut, ici, calibrer le temps lui-même. Des horaires que chaque ville lisait dans son heure solaire locale étaient incommensurables entre eux ; les compagnies britanniques adoptèrent dès les années 1840 une heure unique, dite heure des chemins de fer, et c'est la pression du rail qui imposa en 1883 les fuseaux horaires au continent nord-américain. L'heure standard n'est pas née d'un besoin astronomique mais d'un besoin de circulation : pour que des données de mouvement se composent, il fallait que tous les lecteurs partagent le code qui les rend déchiffrables.

La circulation et la réception, enfin, donnent au cas sa leçon la plus durable. Le lecteur de ce système — l'aiguilleur, le régulateur — a une fenêtre étroite et un temps compté, et il lit des données dont la péremption est presque immédiate : une dépêche lue trop tard ne décrit plus rien, et la donnée périmée y est pire que l'absence de donnée, car elle autorise avec assurance un réel disparu. La probité de ce lecteur — dater ce qu'il transmet, ne tenir pour vrai que ce qui est frais — fut, très tôt, une affaire de vies humaines. Le chapitre 8 retrouvera cette structure intacte dans le transport physique d'aujourd'hui : le secteur où la coordination par la donnée a pris sa forme moderne est aussi celui où la fraîcheur n'a jamais cessé d'être le critère premier.

Ce que l'atlas a montré

Huit dispositifs, de l'argile babylonienne au rail du XIX^e siècle, et la même grille les traverse tous. Aucun n'avait de capteur, de base de données, de modèle ; tous construisaient des flux d'observation, opéraient des choix de critères, affrontaient la péremption, le transport et la lecture. Les concepts de la trilogie ne sont pas nés avec le numérique : le numérique les a remassifiés, accélérés, parfois reconfigurés par saut systémique — mais leur structure était là, lisible, dès qu'un humain a creusé un canal dans le fleuve.

Une limite, pour finir, qui est aussi un avertissement de méthode. Ces cas éclairent ; ils ne dictent rien. On ne peut pas conclure d'un registre paroissial à une règle pour les bases de données contemporaines, ni d'un portulan à une politique de partage. Ce que l'atlas offre, c'est la reconnaissance d'une structure commune sous des formes très éloignées — et donc l'assurance que les difficultés présentes ne sont pas sans précédent. C'est beaucoup, pour qui doit décider aujourd'hui. Ce n'est pas une carte des récifs : c'est la preuve que d'autres, avant nous, ont navigué les mêmes eaux.

Fin de la Partie II.

Partie III — Études de cas sectorielles

Texte rédigé en mai 2026. Les secteurs évoluent et les techniques aussi ; les états des lieux ci-dessous portent leur date et signalent leurs zones de péremption rapide. La grille qui les organise, elle, est stable.

Note liminaire — La structure-type des études sectorielles

Chaque secteur est traité selon une même grille en huit sections, qui suit l'ordre des trois temps de la donnée. Cette homogénéité est délibérée : elle permet de comparer deux secteurs en parcourant les mêmes sections dans chacun.

1. **Le secteur en bref.**
2. **Le flux du réel observé** (*fabrication*).
3. **Les flux d'observation tenus** (*fabrication*).
4. **L'optimum observationnel et les périls de fabrication** (*fabrication*).
5. **La circulation des données du secteur** (*circulation*).
6. **La réception : qui lit, qui ne lit pas, ce qui demeure non actualisé** (*réception*). Cette section renvoie au chapitre 10 pour le cas du LLM, sans le dupliquer.
7. **Les questions politiques propres** — pont vers la Partie VI.
8. **Pistes pour l'artisan de la donnée dans ce secteur.**

Six secteurs sont traités, choisis pour couvrir les trois temps et une diversité d'enjeux : la qualité de l'air (optimum clair, cas pédagogique), la santé (probité forte, éthique), la sécurité et la justice (biais structurels, boucles, non-lu institutionnel), le climat (long terme, sources stables), le transport physique (péremption rapide, boucles d'itinéraire, coordination d'acteurs hétérogènes), et l'éducation (cas où la réception est l'enjeu central).

Chapitre 4 — Qualité de l'air urbain

Le secteur en bref. La surveillance de la qualité de l'air mobilise des réseaux de capteurs urbains, des agences publiques et, de plus en plus, des dispositifs citoyens. L'enjeu est sanitaire et politique : informer, alerter, déclencher des mesures de restriction. C'est une bonne porte d'entrée, car l'optimum observationnel y est presque palpable.

Le flux du réel observé. Le réel est ici la composition chimique de l'air : dioxyde d'azote, particules fines, ozone, composés organiques volatils. Ce réel a des vitesses propres très variées : la concentration de particules peut changer en une heure, sous l'effet du trafic ou du vent, tandis que des moyennes saisonnières évoluent lentement. Le secteur affronte donc plusieurs échelles temporelles à la fois.

Les flux d'observation tenus. Les stations de mesure de référence, peu nombreuses et coûteuses, fournissent des données fiables et calibrées. Les micro-capteurs, bon marché et multipliables, fournissent un maillage dense mais moins fidèle. Deux flux d'observation coexistent donc, construits dans des intentions distinctes : l'un vise la fidélité réglementaire, l'autre la couverture spatiale. Les confondre serait une erreur — leurs critères diffèrent.

L'optimum observationnel et les périls de fabrication. L'arbitrage est explicite : densité du maillage contre coût d'instrumentation contre fidélité. Sur-instrumenter en capteurs bon marché donne une carte dense mais bruitée ; sous-instrumenter en stations de référence donne une carte fidèle mais trouée. Le péril propre est la **dérive de calibration** : un capteur non recalibré produit des données qui s'éloignent silencieusement du réel. Autre péril, une **boucle** caractéristique : un pic mesuré déclenche une restriction de circulation, qui modifie le réel observé — l'observation agit sur ce qu'elle observe.

La circulation des données du secteur. Les mesures circulent vers des plateformes publiques, des indices agrégés, des applications. La **propagation** y est forte : une station alimente des indices régionaux, nationaux, des modèles de prévision. Le risque de circulation tient à la perte de contexte : un indice agrégé, transporté loin de sa source, perd la trace des critères qui l'ont produit. L'utilisateur d'une application qui affiche « air de bonne qualité » ignore le plus souvent quel flux, à quelle fidélité, avec quel âge, a produit ce verdict. La chaîne n'est pas traçable pour lui.

La réception. Le lecteur final — un habitant, un asthmatique, un parent — a une fenêtre étroite : il veut un verdict simple, pas une série de mesures. La réception y est donc presque toujours **déléguée** à un indice qui résume. Cette délégation est légitime, mais elle a un coût : l'utilisateur hérite de la fenêtre de celui qui a construit l'indice, sans le savoir. Beaucoup de données fines passent en **péremption rapide** sans avoir été actualisées — pour l'alerte, un pic d'il y a trois jours n'intéresse plus personne — et c'est le sort normal de ces données, non un échec. L'usage des modèles

automatiques de prévision relève du chapitre 10.

Les questions politiques propres. Qui finance le maillage ? Qui valide les capteurs citoyens ? L'accès aux données est-il public ? La qualité de l'air est un cas typique de donnée d'intérêt commun, où la dimension politique (Partie VI) se pose nettement : un commun de la mesure environnementale.

Pistes pour l'artisan de la donnée. Dans ce secteur, l'artisan veille à la calibration, sait distinguer les deux flux (référence et maillage) et ne les agrège qu'en connaissance de cause, et documente l'âge et la fidélité de tout indice qu'il transmet. Sa probité tient à ne jamais laisser un verdict simple masquer la chaîne qui l'a produit.

Chapitre 5 — Santé hospitalière

Le secteur en bref. L'hôpital produit et consomme des flux d'observation intenses sur le corps des patients : signes vitaux, examens, antécédents, observations cliniques. L'enjeu est vital au sens propre, et l'exigence de probité y est maximale.

Le flux du réel observé. Le réel est l'état du patient, qui a des vitesses propres extrêmement variées : un rythme cardiaque change en secondes, un bilan biologique en heures, un antécédent est quasi stable. Un même patient est donc l'objet de flux d'observation aux cadences hétérogènes.

Les flux d'observation tenus. Le monitoring continu en réanimation, les examens ponctuels, le dossier d'antécédents, les observations infirmières et médicales : autant de flux distincts, tenus par des acteurs différents, avec des intentions différentes. L'optimum n'est pas le même selon le service : la réanimation exige une cadence et une densité que la consultation externe ne justifierait pas.

L'optimum observationnel et les périls de fabrication. L'arbitrage oppose la finesse de surveillance au coût — humain, matériel, et au confort du patient. Sur-surveiller fatigue le patient et noie le soignant sous des alertes. Sous-surveiller laisse passer des dégradations. Le péril majeur est la **boucle décisionnelle** des scores d'alerte : un score qui déclenche une intervention modifie l'état qu'il mesurait. Autre péril, le **non-prélevé** : ce qui n'est pas dans le protocole de surveillance n'est pas vu, et l'on peut conclure à tort à l'absence de ce qu'on n'a pas cherché.

La circulation des données du secteur. Le partage inter-hospitalier est une **propagation** à fort enjeu : un dossier qui suit le patient d'un établissement à l'autre améliore le soin, mais chaque transport est un point de perte de fidélité et de contexte. La **confidentialité** agit ici comme un **empêchement délibéré du transport** : on restreint volontairement la circulation pour protéger le patient. C'est une limitation choisie, qui a un coût (la donnée non partagée peut manquer là où elle serait utile) et une justification (la part du lecteur indu doit rester impraticable). Le secteur

illustre que tout transport n'est pas souhaitable, et que l'empêchement peut être un acte de probité.

La réception. Le lecteur est le soignant, dont la fenêtre est saturée. Le sur-débit clinique est réel : trop d'alertes tuent l'alerte. Le soignant doit en permanence **cadrer** (décider ce qu'il surveille), **déléguer** (aux scores, aux collègues) et **accepter** qu'il ne peut tout lire. La probité du lecteur est ici une question de vie ou de mort : répondre de ce qu'on a lu, et de ce qu'on a délégué, engage la responsabilité clinique. L'usage des modèles automatiques d'aide au diagnostic relève du chapitre 10.

Les questions politiques propres. Propriété du dossier, consentement au partage, secret médical, accès de la recherche aux données de soin : la santé est un champ dense de tensions entre l'utilité du flux et la protection de la personne. La Partie VI y revient.

Pistes pour l'artisan de la donnée. Dans le soin, l'artisan tient la probité comme exigence première : il sait quel flux il lit, à quelle fraîcheur, et il documente ce qu'il transmet. Il traite la confidentialité non comme une contrainte externe mais comme un critère de fabrication à part entière. Et il se méfie des boucles, en distinguant ce que le score mesure de ce que le score provoque.

Chapitre 6 — Sécurité et justice

Ce chapitre porte sur un terrain politiquement chargé. Il s'en tient à la grille processuelle et se garde de toute prise de position partisane : il cartographie une structure, il ne plaide pas.

Le secteur en bref. Police, justice, sécurité publique produisent et exploitent des flux d'observation sur la délinquance, les infractions, les personnes. L'enjeu est lourd : ces flux orientent des décisions qui touchent à la liberté.

Le flux du réel observé. Le réel serait la criminalité réelle — l'ensemble des infractions commises. Mais ce réel est, par nature, partiellement caché : beaucoup d'infractions ne sont jamais signalées. Le secteur affronte d'emblée un écart structurel entre le réel et ce qu'on peut en prélever.

Les flux d'observation tenus. Les statistiques policières mesurent les infractions *constatées*, non les infractions *commises*. C'est le cas paradigmatique du **biais de prélèvement structurel** : on prélève là où l'on a déjà déployé des moyens. Patrouiller un quartier y fait constater plus d'infractions, ce qui justifie d'y patrouiller davantage. Le flux d'observation mesure en partie l'intensité de sa propre activité, non seulement le réel qu'il prétend décrire.

L'optimum observationnel et les périls de fabrication. Le péril

dominant est la **boucle**, ici particulièrement vicieuse : la prédiction policière. Un modèle qui prédit la délinquance là où elle a été constatée y envoie des moyens, qui y constatent davantage, ce qui confirme le modèle. La boucle se referme sur elle-même et se prend pour une vérité. Le **non-prélevé** est ici politique : les violences conjugales, longtemps sous-signalées et sous-prélevées, ou la cybercriminalité, mal captée par des dispositifs conçus pour la délinquance de voie publique, sont massivement absentes des flux — non parce qu’elles seraient rares, mais parce que les instruments n’étaient pas tournés vers elles.

La circulation des données du secteur. Les statistiques circulent vers le débat public, les décisions politiques, l’allocation de moyens. La **propagation** y est à haut risque : un chiffre de criminalité constatée, transporté loin de sa source et de ses critères, est lu comme un chiffre de criminalité réelle. La perte de contexte en circulation transforme une mesure d’activité en verdict sur le monde.

La réception. Le lecteur institutionnel — décideur, élu, citoyen — lit ces chiffres à travers une fenêtre qui ne distingue pas toujours le constaté du commis. C’est un cas remarquable de **non-lu structurel** au sens du troisième tome : certaines réalités (violences conjugales hier, cybercriminalité aujourd’hui) ne sont pas lues parce que la fenêtre du lecteur institutionnel ne les distingue pas — désintéressé non pas individuel mais organisé par la forme des instruments. Reconnaître ce non-lu, c’est déjà commencer à le corriger. L’usage des modèles prédictifs relève du chapitre 10, où la position située de reprise ou de traitement qu’ils occupent est centrale.

Les questions politiques propres. Transparence des algorithmes publics, droit de contester une décision automatisée, audit citoyen des modèles : ce secteur est le plus exposé de tous aux questions de la Partie VI. La prudence éditoriale y est de mise — on décrit la structure du biais, on ne désigne pas de coupable.

Pistes pour l’artisan de la donnée. Dans ce secteur, l’artisan distingue systématiquement la criminalité constatée de la criminalité réelle, traque les boucles de prédiction, et nomme le non-prélevé comme une donnée à part entière. Sa probité consiste à refuser qu’une mesure d’activité soit lue comme un verdict sur le monde.

Chapitre 7 — Climat et environnement

Le secteur en bref. La science du climat repose sur des flux d’observation de très longue durée et sur des reconstitutions du passé. L’enjeu est planétaire, et la dimension politique de la donnée y est explicite.

Le flux du réel observé. Le réel est le système climatique, dont les vitesses propres s’évaluent de l’événement météorologique quotidien aux

cycles séculaires. La donnée climatique est le cas-type des **sources stables** au long cours : températures moyennes, niveaux des mers, composition atmosphérique évoluent lentement à l'échelle humaine — mais ce rapport de vitesses se dégrade à mesure que le réel s'emballe, ce qui est précisément l'objet d'inquiétude du secteur.

Les flux d'observation tenus. Réseaux de stations, satellites, bouées, séries instrumentales longues. Et, pour le passé pré-instrumental, les **reconstitutions proxy** : cernes d'arbres, carottes de glace, sédiments. Ces proxys sont un cas particulièrement intéressant au regard de la trilogie — ce sont des **données rejouées**, constituées aujourd'hui sur des traces du réel passé par un prélèvement rétrospectif. La glace n'a pas été prélevée comme donnée quand elle s'est formée ; elle est devenue donnée quand un observateur, bien plus tard, a su la prélever et en inscrire les mesures — par un acte d'inscription, non par un acte de lecture.

L'optimum observationnel et les périls de fabrication. L'arbitrage oppose la densité spatio-temporelle du réseau au coût considérable des dispositifs longs. Le péril propre est l'**hétérogénéité des séries** : raccorder des séries instrumentales produites par des dispositifs différents sur un siècle exige une calibration soignée, faute de quoi on prend une rupture d'instrument pour un signal climatique. Le **non-prélevé** historique — les régions et époques mal couvertes — pèse sur la reconstitution.

La circulation des données du secteur. La donnée climatique circule largement et alimente des **modèles** qui sont des données consolidées de second ordre. La question de la circulation y est aussi une question d'accès : qui peut consulter, vérifier, recalculer ? La fidélité du transport est cruciale, car une donnée climatique mal recopiée ou décontextualisée nourrit la contestation. La traçabilité jusqu'à la source est ici un enjeu de crédibilité scientifique autant que technique.

La réception. Le sur-débit climatique est massif : aucun lecteur humain ne lit l'intégralité des séries et des modèles. La **délégation** aux modèles et aux synthèses est inévitable et légitime, à condition d'être assumée. Le lecteur public, lui, reçoit des synthèses très cadrées, dont il hérite la fenêtre. Une part de la donnée brute demeure non actualisée faute de lecture — mais constituée, conservée, archivée, donc rouverable : donnée muette pour ce lecteur plutôt que morte, car un autre lecteur pourra la reprendre. L'usage des modèles automatiques relève du chapitre 10.

Les questions politiques propres. Qui finance les flux longs ? Qui les valide ? À qui sont-ils accessibles ? La donnée climatique est un commun scientifique mondial dont la gouvernance est un enjeu politique de premier plan. La Partie VI y revient.

Pistes pour l'artisan de la donnée. Au climat, l'artisan soigne la calibration des séries longues, documente les reconstitutions proxy comme des données rejouées et non comme des mesures directes, et tient

la traçabilité jusqu'à la source comme une exigence de crédibilité. Sa probité consiste à distinguer la mesure du modèle, et le rejoué du prélevé.

Chapitre 8 — Transport physique et mobilité

*Dans toute la trilogie, « transport » désigne le mouvement de la donnée. Ce chapitre traite du déplacement des marchandises et des personnes, qu'il ne désigne que sous le nom complet de **transport physique**, ou de **mobilité**. Quand le mot « transport » apparaît seul, c'est au sens du deuxième tome.*

Le secteur en bref. Le déplacement des marchandises et des personnes — fret, logistique, transports collectifs, route — est l'un des plus anciens terrains de la coordination par la donnée : avant même que le mot ne soit d'usage, l'horaire ferroviaire et le télégraphe coordonnaient des circulations que nul ne pouvait plus tenir à vue (l'atlas du chapitre 3 en traite au cas 8). Le secteur a été précurseur par nécessité : faire circuler des véhicules nombreux sur des infrastructures finies exige de savoir, à tout instant, où sont les choses et ce qui les attend — l'optimum de son fonctionnement est, depuis l'origine, un optimum de coordination. Le secteur a deux objets, qu'il faut tenir ensemble sans les confondre : la marchandise, qui ne parle pas mais se déclare, et la personne, qui se déplace librement et se capte.

Le flux du réel observé. Le réel est ici un réel mobile : positions, vitesses, charges, états de l'infrastructure et de ce qui y circule. Ses vitesses propres sont parmi les plus contrastées de tous les secteurs : la position d'un véhicule change en secondes, l'état d'un trafic en minutes, une infrastructure en années. Et le secteur présente une particularité que les précédents n'avaient pas à ce degré : l'observé se déplace pendant qu'on l'observe. L'écart entre le réel et son prélèvement, constitutif de toute donnée, s'y creuse à la vitesse même de ce qu'on suit.

Les flux d'observation tenus. Quatre familles de flux coexistent, d'intentions distinctes. Les flux embarqués : le véhicule instrumenté prélève sur lui-même et sur son environnement — position, vitesse, état de charge. Les flux d'infrastructure : boucles de comptage, péages, capteurs et caméras prélèvent sur ce qui passe devant eux. Les flux déclaratifs : manifestes, lettres de voiture, réservations, billets — la part contractuelle du secteur, où la donnée précède le déplacement qu'elle annonce. Les flux de plateformes, enfin : les positions des téléphones, captées à d'autres fins et réemployées pour estimer le trafic. Chacun a été conçu pour son usage — exploiter, facturer, contracter, servir — et leur agrégation hérite de ces intentions divergentes.

L'optimum observationnel et les périls de fabrication. L'arbitrage oppose la cadence de prélèvement au coût d'instrumentation et de transmission : suivre chaque véhicule à la seconde produit un débit que

nul usage ne justifie ; le suivre à l'heure ne sert plus la régulation. Le péril dominant est la **péremption rapide** portée à son extrême : pour la régulation, une position se périme en secondes, une estimation de trafic en minutes — la donnée du secteur est parmi les plus périssables qui soient, et son sort normal est de passer en péremption presque aussitôt produite, souvent sans avoir été lue, ce qui n'est pas un échec. L'autre péril est une **boucle** caractéristique : l'information de trafic modifie le trafic. L'itinéraire conseillé à tous déplace la congestion qu'il prétend éviter ; l'estimation publiée agit sur le réel qu'elle estimait. Le **non-prélevé** a sa géographie propre : les modes non instrumentés — le piéton, le vélo, le fret informel, le dernier kilomètre — sont absents des flux non parce qu'ils seraient marginaux, mais parce que les instruments ne sont pas tournés vers eux.

La circulation des données du secteur. La donnée de mobilité circule entre des acteurs hétérogènes — transporteurs, gestionnaires d'infrastructure, autorités organisatrices, plateformes — dont les flux n'ont pas été conçus pour s'assembler : l'interopérabilité y est un travail permanent, fait de conversions et d'accords sémantiques, et chaque relais est un point de perte de fidélité et de contexte. Un temps de parcours moyen, transporté loin de sa source, ne dit plus ni sa fenêtre de calcul ni ce qu'il a exclu. Surtout, le secteur est l'un des rares où le vieillissement de trajet du deuxième tome reste sensible malgré l'instantanéité : la donnée doit rejoindre un lecteur en mouvement avant de se périmer, et le délai de la chaîne — prélèvement, consolidation, diffusion — se mesure à l'aune d'une validité qui se compte en secondes. Enfin, la position d'une personne est une donnée sur la personne : la restriction de sa circulation est, comme en santé, un empêchement délibéré du transport — une limitation choisie, qui protège.

La réception. Trois lecteurs aux fenêtres très inégales. L'exploitant, en poste de régulation, a une fenêtre saturée : il **cadre** (il choisit ce qu'il surveille), **délègue** (aux seuils, aux alarmes, aux synthèses) et **accepte** qu'il ne lira pas tout. Le conducteur a la fenêtre la plus étroite du corpus : mobile, occupée par la conduite même — le sur-débit y est un danger physique, car lire en se déplaçant se paie sur l'attention au réel. Le voyageur ou le chargeur, enfin, ne veut qu'un verdict : une heure d'arrivée, un statut de livraison — délégation légitime, dont il hérite la fenêtre sans la voir. Une grande part de la donnée produite passe en péremption rapide sans avoir été actualisée, et c'est son sort normal. Quant au traitement embarqué qui agit sans qu'aucun humain ne lise — l'assistance, la conduite automatisée —, il relève du fond posé par le troisième tome : la machine y occupe une position située de reprise ou de traitement, un traitement délégué inséré dans une chaîne dont l'imputabilité demeure aux agents et aux institutions qui conçoivent ce traitement et en emploient les résultats. L'usage des modèles relève du chapitre 10.

Les questions politiques propres. La trace de mobilité est l'une des

plus révélatrices qui soient : savoir où va une personne, c'est savoir beaucoup d'elle. Qui détient ces traces — plateformes privées, autorités organisatrices, constructeurs ? À quelles conditions la puissance publique accède-t-elle aux flux des plateformes qui régulent de fait sa voirie ? La souveraineté des flux logistiques — savoir ce qui circule sur son territoire — rejoint les questions de la Partie VI.

Pistes pour l'artisan de la donnée. Dans ce secteur, l'artisan tient la fraîcheur comme critère premier : il date tout ce qu'il transmet, et distingue systématiquement la position observée de la position estimée, l'arrivée prédite de l'arrivée constatée. Il traque les boucles d'itinéraire — ce que la publication d'une estimation fait au réel estimé — et documente ce que chaque agrégat a exclu. Sa probité consiste à ne jamais laisser une prédiction se lire comme une observation, ni une moyenne masquer la fenêtre qui l'a produite.

Chapitre 9 — Éducation et apprentissage

Le secteur en bref. L'éducation produit des flux d'observation sur les élèves, et elle est aussi le lieu où se forme — ou ne se forme pas — la capacité à lire les flux. C'est le secteur où la **réception** est l'enjeu central, ce qui en fait le complément nécessaire des quatre précédents.

Le flux du réel observé. Le réel est l'apprentissage : ce qui se passe dans la tête d'un élève, par nature inobservable directement. On ne prélève jamais l'apprentissage ; on prélève des indices — réponses, comportements, présences — dont on infère l'apprentissage. L'écart entre le réel et le prélevé est ici constitutif et profond.

Les flux d'observation tenus. Notes, présences, évaluations, comportements observés ; et, de plus en plus, les flux d'observation des contenus consommés par les élèves — algorithmes de recommandation, temps passé, parcours numériques. Deux familles de flux coexistent : ceux que l'institution tient sur l'élève, et ceux que des plateformes tiennent sur l'élève à son insu ou presque.

L'optimum observationnel et les périls de fabrication. Le péril dominant est la **confusion de l'indice et du réel** : prendre la note pour l'apprentissage, le temps passé pour l'attention, la présence pour l'engagement. Sur-prélever sur l'élève — le mesurer en continu — produit une donnée abondante mais qui peut dégrader ce qu'elle mesure, en transformant l'apprentissage en performance observable. La **boucle** y est forte : ce qui est évalué oriente ce qui est appris.

La circulation des données du secteur. Les données scolaires circulent vers les familles, les institutions, parfois des prestataires privés. La circulation des flux d'observation des plateformes est massive et opaque : l'élève est observé par des dispositifs dont ni lui ni ses maîtres

ne tracent la chaîne. La perte de contexte et l'empêchement de remonter à la source y sont la règle plutôt que l'exception.

La réception. C'est ici que le secteur prend tout son sens. L'éducation est le lieu où l'on devrait apprendre à **lire le flux** — à reconnaître un flux d'observation, ses critères, son intention, son non-prélevé. Or les élèves sont massivement exposés à des flux (recommandations, fils d'actualité) qu'ils subissent sans les lire, par **sur-débit** : ils voient passer sans déchiffrer. Former à la réception, c'est apprendre à choisir une posture — **cadrer** son attention, **déléguer** en connaissance de cause, **accepter** lucidement le débordement — plutôt que de subir. La compétence de lecture du flux est l'enjeu pédagogique propre du moment numérique, et elle conditionne tous les autres secteurs : un citoyen qui ne sait pas lire un flux ne saura juger ni l'air qu'il respire, ni les chiffres de la sécurité, ni les synthèses climatiques, ni l'heure d'arrivée qu'on lui promet. L'usage scolaire des LLM, central, relève du chapitre 10.

Les questions politiques propres. Protection des données des mineurs, encadrement des plateformes éducatives, place de la littérature du flux dans les programmes : autant de questions que la Partie VI prolonge.

Pistes pour l'artisan de la donnée. En éducation, l'artisan distingue l'indice de l'apprentissage, refuse de prendre la mesure pour le réel, et — surtout — œuvre à transmettre la lecture du flux comme une compétence. Sa probité est de second ordre : non seulement bien lire, mais apprendre à lire.

Ce que la traversée sectorielle a montré

Six secteurs, une seule grille. La comparaison transversale fait apparaître des constantes. Le péril de confondre la donnée avec le réel qu'elle prélève — l'indice avec l'apprentissage, le constaté avec le commis, l'indice d'air avec l'air, la prédiction d'arrivée avec l'arrivée — traverse tous les secteurs. La boucle, où l'observation modifie l'observé, est partout, du recensement au score clinique et à l'itinéraire conseillé. Le non-prélevé est partout politiquement chargé. Et la réception, longtemps négligée au profit de la fabrication, apparaît comme l'enjeu montant : le sur-débit, la délégation, le non-lu structurel se retrouvent d'un secteur à l'autre.

La limite, ici, est la même que pour l'atlas historique : ces études éclairent une structure, elles ne dictent pas de conduite. Chaque secteur a son optimum propre, que seul un praticien situé peut établir. La grille processuelle ne remplace pas l'expertise sectorielle ; elle lui donne un langage commun pour nommer ce qu'elle rencontre. C'est ce langage que la boîte à outils de la Partie V rend maniable.

Fin de la Partie III.

Partie IV — Vivre avec le LLM

Texte rédigé en mai 2026. C'est le chapitre le plus exposé à la péremption : les capacités techniques des grands modèles de langage évoluent vite, et les pratiques d'atténuation devront être révisées à chaque saut technique majeur — au minimum une fois l'an. Les principes, eux, tiennent ; les outils nommés datent.

Chapitre 10 — Le LLM comme saut systémique de réception, et les pratiques de restauration

Le statut de ce chapitre : une illustration, non un concept

Avant tout, une mise au point qui commande la lecture. Le grand modèle de langage n'est pas un concept de la trilogie. C'est une **illustration** — majeure, urgente, datée — de concepts qui, eux, sont génériques et stables : le saut systémique posé par le premier tome, et le lecteur situé, humain ou automatique, posé par le troisième. Tout ce qui suit décline ces concepts sur un cas du présent ; rien n'y ajoute au socle conceptuel. Cette discipline a une raison de fond et une raison de méthode. De fond : faire du LLM un concept autonome contredirait l'architecture de la trilogie, qui le range explicitement parmi les exemples. De méthode : un concept de fond ne se périme pas, alors qu'une illustration technique se périme en quelques années — et ce chapitre porte d'ailleurs la date de péremption la plus courte de tout l'ouvrage.

On traite donc le LLM comme ce qu'il est : le cas le plus pressant où une transformation technique rejoue, dans la réception, un saut systémique qui touche aussi la fabrication. À ce double titre, il se tient au point de jonction du premier et du troisième temps.

Rappel du socle : la fusion des trois rôles

Le matériau conceptuel de ce chapitre est posé ailleurs — dans la note de référence sur le saut systémique du stockage, et dans le chapitre du

troisième tome sur la lecture automatique. On en rappelle ici l'essentiel, en renvoyant à ces textes pour le détail.

Historiquement, l'accès au savoir stocké reposait sur trois fonctions séparées. Le **bibliothèque** : le stock, l'ensemble de ce qui est conservé. Le **bibliothécaire** : l'orientation, celui qui sait où chercher, qui ordonne et oriente. Le **conteur** : la restitution, celui qui met en forme et en sens ce que le stock recèle. Stock, orientation, restitution — trois fonctions tenues par des dispositifs et des personnes distincts. *L'Encyclopédie* de Diderot, vue au chapitre 3, en était une incarnation : auteurs, appareil de classement, ouvrage médiateur, séparés.

Cette séparation n'était pas qu'une commodité d'organisation. Elle était un **dispositif de contrôle**. Parce que les trois rôles étaient distincts, le lecteur pouvait remonter d'une restitution à sa source, contester un classement, juger un orienteur. La séparation rendait possible le retour à la source, la traçabilité du chemin, la lisibilité et la contestation du rangement — et, par-dessus tout, elle maintenait un contre-pouvoir : aucun acteur ne tenait à lui seul le savoir, du stock à sa mise en forme.

Le grand modèle de langage opère un **saut systémique** : il fusionne les trois fonctions en une seule instance. La même opération tient le stock, l'oriente et le restitue, en un geste unique et opaque. Ce n'est pas un perfectionnement de la bibliothèque ; c'est un changement de régime, qui supprime la séparation — et avec elle, ce que la séparation garantissait. Le modèle ne donne pas accès à un stock qu'on pourrait inspecter ; il restitue sans que la source soit séparable de la restitution. Il oriente sans qu'on puisse lire le principe de son orientation. Le contre-pouvoir que la séparation maintenait n'est plus tenu par l'architecture.

Et — c'est l'apport du troisième tome — ce modèle qui restitue occupe une **position située de reprise ou de traitement**. Quand il « lit » le corpus qui l'a formé, il le fait à travers une fenêtre qui est la sienne, avec ses bornes, ses angles morts, sa manière de distinguer et de déchiffrer. Déléguer sa lecture à un LLM, ce n'est pas accéder à une lecture sans point de vue : c'est hériter de la fenêtre d'un traitement automatique dont on ne connaît ni la position ni les limites.

Une typologie des usages, croisée avec les trois postures

Tous les usages d'un LLM ne mobilisent pas également la fusion des trois rôles. Avant d'agir, le praticien gagne à **positionner son propre usage**. On distingue trois classes, qu'on croise avec les trois postures du lecteur (cadrer, déléguer, accepter) vues au chapitre 1.

Les **usages neutres à la fusion**. Reformuler un texte qu'on a déjà sous les yeux, traduire un fragment, corriger un style, générer des variations. L'humain garde la source ; le modèle ne fait office que d'assistant linguistique. La fusion des trois rôles est peu mobilisée, les pertes sont

faibles. En termes de posture, l'humain **cadre** : sa fenêtre reste sur un matériau qu'il maîtrise, le modèle n'y ajoute pas de stock caché.

Les **usages risqués par la fusion**. Synthétiser un dossier qu'on n'a pas lu en entier, s'informer sur un sujet qu'on connaît mal, rédiger un argumentaire dans un domaine technique qu'on ne maîtrise pas. Ici, l'humain dépend du modèle pour le stock, l'orientation et la restitution à la fois : il ne peut ni vérifier la source, ni juger l'orientation, ni reformuler en connaissance de cause. Les pertes sont maximales. En termes de posture, l'humain **délègue sans le savoir** — il croit lire alors qu'il hérite d'une fenêtre dont il ignore les bornes.

Les **usages exigeant restauration**. Décider sur la foi d'un modèle, produire un contenu destiné à circuler, enseigner. Ces usages ne sont pas interdits, mais ils exigent qu'on restaure *activement* la séparation des rôles. Ne pas le faire, c'est ce que le troisième tome nomme une **délégation non assumée** : déléguer sa lecture à un lecteur automatique sans en répondre. Le défaut de restauration, dans ces cas, est un **défaut de probité du lecteur**. La fiche source le disait crûment : l'absence de restauration équivaut à une démission.

Cette typologie n'enferme pas un usage dans une case définitive ; le même outil sert tour à tour aux trois. Elle apprend à reconnaître, à chaque fois, dans quelle classe on se trouve — et donc quelle posture s'impose.

Cinq pratiques de restauration

Pour chaque rôle que la fusion supprime, une ou plusieurs pratiques le restaurent partiellement. Ce sont des **palliatifs**, non des solutions : ils n'annulent pas le saut systémique, ils en atténuent les pertes. La fusion demeure ; ce que des pratiques humaines peuvent restaurer, c'est une part de ce que l'architecture ne tient plus. Chacune s'ancre dans un concept de la trilogie.

Restaurer la bibliothèque — le retour à la source. Quand le contenu compte, exiger les sources sous-jacentes ; recourir aux dispositifs qui adossent la restitution à des documents identifiables ; vérifier au moins une source par sortie utilisée. C'est restaurer la possibilité de remonter le **transport jusqu'à la source** — la chaîne étendue du deuxième tome. Ce que cette pratique restaure : la traçabilité de l'origine. Ce qu'elle ne restaure pas : la garantie que la source citée a réellement fondé la restitution, le lien pouvant être reconstruit après coup.

Restaurer le bibliothécaire — la traçabilité du chemin. Tenir un journal de consultation : ce qu'on a demandé, à quel modèle, à quelle date, avec quelle reformulation. Pratique simple, qui restaure une part de la **traçabilité** que l'architecture supprime — le transport comme rapport que l'on peut documenter. Ce qu'elle restaure : la mémoire de son propre cheminement. Ce qu'elle ne restaure pas : la lisibilité du chemin interne du modèle, qui reste opaque.

Restaurer le conteur — la synthèse responsable. Ne jamais reprendre une formulation du modèle sans la reformuler avec ses propres mots. Reformuler oblige à percevoir, distinguer, déchiffrer — à accomplir l'**acte de lecture** du troisième tome au lieu de le déléguer. C'est par cet acte qu'on redevient capable de répondre de ce qu'on transmet. Ce qu'elle restaure : la compréhension et la responsabilité. Ce qu'elle ne restaure pas : ce que le modèle a omis et qu'on ne peut reformuler faute de le voir.

Restaurer la critique — la lisibilité du rangement. Confronter la sortie d'un modèle à celle d'un autre, ou à une source non issue d'un modèle. Le désaccord entre deux **lecteurs situés** automatiques révèle les zones où le rangement est incertain — là où leurs fenêtres divergent, il y a matière à se méfier. Ce qu'elle restaure : un signal sur les zones douteuses. Ce qu'elle ne restaure pas : les erreurs sur lesquelles deux modèles concordent par construction semblable.

Restaurer le contre-pouvoir — la séparation organisée par les humains. Ne jamais s'en remettre à un seul modèle pour une décision à fort enjeu. La séparation des rôles, que l'architecture ne tient plus, doit être réorganisée *autour* du modèle par les humains qui l'utilisent. C'est un acte de **sagesse collective** au sens du deuxième tome : décider ensemble de ne pas concentrer en une instance ce qui doit rester réparti. Ce qu'elle restaure : un contre-pouvoir externe. Ce qu'elle ne restaure pas : le contre-pouvoir interne à l'outil, structurellement absent.

Déclinaisons sectorielles

La grille précédente prend une couleur particulière selon les secteurs de la Partie III. On en esquisse les lignes ; chacune pourrait faire l'objet d'une fiche pratique brève.

Éducation. L'enjeu est de former à un usage discerné dès le secondaire, et de garder la lecture intacte comme compétence. Évaluer un travail produit avec un modèle suppose de distinguer ce que l'élève a déchiffré de ce qu'il a délégué. Le risque propre : qu'une génération apprenne à recevoir des restitutions sans jamais avoir construit la capacité de retour à la source.

Journalisme. Utiliser un modèle dans la rédaction sans déléguer l'enquête. La traçabilité doit être marquée dans la publication ; le retour à la source y est une exigence déontologique avant d'être une bonne pratique.

Justice. Quels usages sont compatibles avec l'exigence de motivation des décisions ? Une décision qui doit être motivée ne peut reposer sur une restitution opaque : l'usage qui compromet la traçabilité du raisonnement compromet la motivation elle-même.

Médecine. Utiliser un modèle en appui sans déléguer le diagnostic. La responsabilité du clinicien — répondre de ce qu'il a lu — interdit la délégation non assumée. L'aide est légitime ; l'abdication ne l'est pas.

Transport physique. Utiliser un modèle pour planifier, synthétiser des états de réseau ou assister la régulation, sans déléguer la responsabilité d'exploitation. Le risque propre tient à la fraîcheur : un modèle est une sauvegarde qui diverge, et dans un secteur où la donnée se périme en secondes, sa restitution décrit un réseau qui n'existe déjà plus — toute restitution doit porter son âge. La compétence-maître y est de distinguer la donnée observée de la donnée vraisemblable : une prédiction d'arrivée n'est pas une arrivée.

Administration et recherche. Dans l'administration, quelle traçabilité offrir au citoyen sur les documents produits avec un modèle ? Dans la recherche, quel usage reste compatible avec la vérifiabilité des sources ? Un cas limite mérite mention : quand un modèle est formé sur un corpus dont la trace d'origine a disparu ou est devenue illisible, ce qu'il « lit » n'est plus tout à fait une donnée mais une trace dont l'interprétation s'est perdue — proche de la **donnée morte** du troisième tome. Reconnaître la nature et l'âge des corpus qui alimentent un modèle est, à cet égard, la compétence-maître du praticien lucide.

Ce que ce chapitre a montré, et ses limites

Le LLM ne supprime pas le savoir : il supprime la séparation des rôles qui en garantissait le contrôle. Telle est la transformation, et les pratiques de restauration ne la renversent pas — elles la rendent vivable, partiellement, pour qui veut continuer d'exercer son métier en lucidité. C'est tout ce qu'elles peuvent, et c'est déjà beaucoup.

Les limites sont de deux ordres. La première est de fond : aucune pratique humaine ne restaure ce qu'une architecture a supprimé ; les palliatifs atténuent, ils ne réparent pas. La seconde est de forme, donc de date : ce chapitre vieillira vite. Les dispositifs techniques évoqués évolueront, certains termes employés ici seront bientôt désuets, et la grille devra être réinstruite à chaque saut. C'est la rançon assumée d'un chapitre qui traite du plus mouvant des objets avec les concepts les plus stables. La trousse à outils, ici plus qu'ailleurs, est datée — et le dire fait partie de sa probité.

Fin de la Partie IV.

Partie V — Boîte à outils

Texte rédigé en mai 2026. Les outils opérationnels sont une forme datable ; ils valent environ cinq ans avant révision. Le concept qui les fonde, lui, ne se périmé pas. Chaque outil renvoie au concept de la trilogie qui le porte et, le cas échéant, à la compétence du référentiel de l'artisan et à l'exercice du manuel d'apprentissage qui l'approfondissent.

Chapitre 11 — Checklists, grilles, fiches de signature, exercices

Des outils, pas des normes

Un outil n'est pas une règle. La règle dispense de penser ; l'outil ouvre une question. Tous les objets de ce chapitre sont des **déclencheurs de questionnement**, jamais des normes à appliquer mécaniquement. Une checklist qu'on coche sans réfléchir ne vaut rien ; une grille qui rend un verdict automatique trahit l'esprit de l'ontologie processuelle, où chaque flux a son optimum propre et chaque lecteur sa fenêtre. Ces outils servent à *ne rien oublier d'essentiel*, pas à remplacer le jugement situé.

Ils dérivent de l'écosystème de l'artisanat de la donnée — référentiel de compétences, manuel d'apprentissage — dont ils offrent un aperçu opérationnel plus court et plus accessible. Ils ne remplacent pas la formation d'artisan ; ils en donnent un avant-goût utilisable par tout praticien.

Cinq outils sont proposés, couvrant les trois temps de la donnée.

Outil 1 — Checklist du flux d'observation rencontré

Onze questions, une page. À se poser quand on vous présente un jeu de données, un indicateur, un chiffre.

Cette checklist est la version cochable de la grille générale du chapitre 2. Elle parcourt les trois temps.

Fabrication 1. De quel flux du réel cette donnée provient-elle ? 2. Qui l’a creusée, dans quelle intention ? 3. Quels sont les critères — cadence, périmètre, instrument ? 4. Quelles unités, normes, références implicites ? 5. Qu’est-ce qui n’a *pas* été prélevé et pourrait être pertinent ? 6. Quelle est la date du prélèvement ? Quel est son âge fonctionnel ? 7. Le flux est-il dans une boucle décisionnelle — modifie-t-il ce qu’il observe ? 8. Quels biais structurels la classe d’annotateur engendre-t-elle ?

Circulation 9. La chaîne amont est-elle traçable jusqu’à la source ? Par combien de mains, de copies, de retraitements la donnée est-elle passée — et qu’a-t-elle pu y perdre ou y subir ?

Réception 10. Pour mon usage, ce flux est-il adapté — ou inadapté ? 11. Ai-je le moyen et le temps de le lire vraiment, au regard de ma fenêtre de lecteur — ou va-t-il me submerger ? Si je ne le lis pas, est-ce incapacité, désintérêt, ou péremption rapide ?

Fondements : tome 1 (questions 1-8), tome 2 (question 9, la chaîne étendue), tome 3 (questions 10-11, le lecteur situé et les trois chemins). La question 9 reformulée rattache la traçabilité au transport ; la question 11 est l’ajout de réception qui complète la version d’origine.

Outil 2 — Grille de péremption

Tableau à remplir pour chaque donnée. Déclencheur de question, non verdict.

<i>Donnée</i>	<i>Fragment du réel observé</i>	<i>Vitesse de changement</i>	<i>Âge fonctionnel</i>	<i>Chemin probable de non-actualisation</i>
Donnée démographique	Population	Lente	Plusieurs années	—
Indicateur économique	Marché	Rapide	Jours à semaines	Péremption rapide
Mesure environnementale	Air, eau	Variable	Heures à mois	Selon usage
Score comportemental	Personne	Rapide	Jours	Péremption rapide

La dernière colonne est l’ajout de réception : elle relie l’âge fonctionnel aux **trois chemins** du troisième tome. Une donnée à fraîcheur très brève relève de la **péremption rapide** — son sort normal est de passer en péremption pour l’usage visé sans avoir été actualisée, et ce n’est pas un échec. La grille rappelle ainsi qu’elle ne juge pas la *qualité* d’une donnée, mais sa *temporalité d’usage* : une donnée périssable n’est pas une mauvaise donnée, c’est une donnée dont la fenêtre de validité est courte.

Fondements : tome 1 (deux temps, péremption), tome 3 (les trois chemins).

Outil 3 — Fiche de signature d'œuvre

Modèle d'une page, à apposer en tête de toute livraison d'un flux ou d'une donnée consolidée.

SIGNATURE D'ŒUVRE

- Artisan : [nom]
- Date de livraison : [date]
- Type d'œuvre :
 - Flux d'observation entretenu
 - Donnée consolidée
 - Diagnostic
- Question d'usage : [phrase courte]
- Critères majeurs :
 - Périmètre : [...]
 - Cadence : [...]
 - Fidélité visée : [...]
 - Corpus mobilisé : [...]
- Optimum observationnel : [justification de l'arbitrage]
- Transport / fidélité attendue à la réception : [ce que le destinataire doit pouvoir faire pour que la part du lecteur soit praticable ; altérations possibles en chemin]
- Fenêtre de lecteur visée : [à qui c'est lisible, à quelles conditions]
- Limites connues : [...]
- Validité estimée : [...]
- Conditions de maintenance / re-prélèvement : [...]

Deux champs complètent le modèle d'origine. **Transport / fidélité attendue à la réception** ancre la signature dans le deuxième tome : le transport est un rapport entre deux déplacements, et signer une œuvre, c'est rendre praticable la part du lecteur — pas seulement déposer la donnée quelque part. **Fenêtre de lecteur visée** l'ancre dans le troisième : une œuvre est faite pour être lue par quelqu'un de situé, et le dire fait partie de la livraison. Sans ces deux champs, on signerait une donnée comme si elle n'avait pas à être reçue — ce qui contredit la thèse d'ouverture du deuxième tome, qu'une donnée que nul ne reçoit n'est pas encore tout à fait une donnée.

Fondements : tome 1 (critères, optimum, deux morts via les conditions de maintenance), tome 2 (transport, fidélité), tome 3 (fenêtre de lecteur).

Outil 4 — Exercices d'éveil

Exercices courts, une demi-heure, pour le collège, le lycée et le grand public. Ils font passer des trois temps abstraits au réflexe.

Le chiffre orphelin. Devant un chiffre cité dans la presse — « 23 % des Français... » —, reconstituer le flux d'observation qui l'a produit : quel réel, quel prélèvement, quels critères, quelle intention. (*Fabrication.*)

Le carnet d'observation. Observer un phénomène simple — une fourmière, un carrefour — pendant quinze minutes, puis formaliser ses

propres choix de critères et leurs limites. On éprouve de l'intérieur que prélever, c'est choisir. (*Fabrication.*)

Qu'est-ce qui n'a pas été dit. Devant un article, lister les paramètres non cités et leur pertinence possible. On apprend à voir le non-prélevé — et, déjà, le non-lu. (*Fabrication et réception.*)

La date manquante. Devant un graphique sans date, estimer l'âge probable des données et ce qu'il change à l'interprétation. (*Fabrication, deux temps.*)

Le fil qui déborde. Devant un flux d'information continu — un fil d'actualité, une boîte de notifications —, identifier sa propre fenêtre de lecteur et choisir explicitement une posture : cadrer (restreindre ce qu'on suit), déléguer (à un tiers, à un filtre, en sachant qu'on hérite de sa fenêtre), ou accepter lucidement de ne pas tout lire. On éprouve le sur-débit et la responsabilité de la posture. (*Réception — exercice ajouté pour ancrer le troisième tome dans la boîte à outils.*)

Fondements : tome 1 (quatre premiers exercices), tome 3 (le cinquième, et la dimension réception du troisième).

Outil 5 — Modèle de revue de canal

Ordre du jour-type pour une équipe qui maintient un flux d'observation. Périodicité fixe, mensuelle ou trimestrielle.

1. Volume et qualité des prélèvements depuis la dernière revue.
2. Calibration des instruments — dérive constatée ?
3. Pertinence des critères au regard de l'usage actuel.
4. Corpus de référence — toujours suffisant ?
5. Boucles éventuelles — effets observés ?
6. Signaux de péremption — quoi rafraîchir ?
7. Charge de maintenance — soutenable ?
8. Décisions et arbitrages — qui répond ?
9. **Réception effective** — le flux est-il lu par ceux à qui il est destiné, ou demeure-t-il non actualisé par incapacité, désintérêt, ou péremption rapide ?

Le point 9 est l'ajout de réception. Une revue qui ne regarde que l'amont — calibration, corpus, péremption — oublie que le canal n'a de sens que s'il aboutit à un lecteur. Un flux parfaitement entretenu que personne ne lit est un flux constitué mais non actualisé ; la revue doit le voir, et distinguer si cette non-lecture est un manquement à corriger (incapacité, désintérêt) ou le sort normal d'une donnée à fraîcheur brève (péremption rapide).

Fondements : tome 1 (calibration, corpus, péremption, boucle), tome 3 (réception effective, trois chemins).

Comment se servir de la boîte

Ces cinq outils ne s'emploient pas tous à la fois ni dans tous les contextes. La checklist sert face à une donnée reçue ; la grille de péremption, quand l'âge est en jeu ; la fiche de signature, au moment de livrer ; les exercices, pour former ; la revue de canal, pour entretenir. Le bon usage est sélectif : prendre l'outil que la situation appelle, et le tenir pour ce qu'il est — un déclencheur, pas un verdict.

Une limite, pour finir, qui vaut pour toute la boîte. Ces outils aident à poser les bonnes questions ; ils ne fournissent jamais les réponses. Aucun ne dit quel âge est acceptable, quelle fidélité suffit, quelle posture adopter. Ces réponses sont situées, et c'est pourquoi l'ouvrage les décline secteur par secteur plutôt que de prétendre les trancher. La boîte à outils est l'incarnation maniable de la grille processuelle — une lampe qu'on peut tenir en main, non une carte des récifs.

Fin de la Partie V.

Partie VI — Formes économiques et politiques

Texte rédigé en mai 2026. C'est, avec le chapitre sur le LLM, la partie la plus datée de l'ouvrage : les projets, institutions et régulations cités peuvent évoluer ou disparaître en trois à cinq ans. Chaque exemple porte son statut. Le fond conceptuel sur lequel s'appuie cette cartographie — la circulation, le partage du fait, la sagesse collective — ne se périme pas ; les formes qui l'incarnent, oui.

Note liminaire — Cartographe, non militaire

Cette partie traite des formes économiques et politiques que prend, aujourd'hui, la gouvernance des données. Elle obéit strictement au principe d'évenhandedness (qualité d'être juste et impartial) de l'ouvrage : elle **cartographie**, elle ne **plaide pas**. Trois grands modèles coexistent en 2026 — le modèle privé concentré des grandes plateformes, qui domine ; le modèle étatique souverain, réactif et en construction ; le modèle des communs, émergent et fragmenté. L'ontologie processuelle offre un cadre conceptuel à ces trois modèles ; elle n'en élit aucun. Faire de la trilogie un manifeste pour les coopératives, pour la souveraineté étatique ou pour le militantisme des données serait confondre le fond avec une préférence politique.

Un point demande une vigilance constante. Le deuxième tome porte un fil — la sagesse collective, et sa formule « l'abstention n'est pas neutre ». Ce fil *ressemble* à une prise de position en faveur de la retenue et des communs. Il ne l'est pas. Le tome établit un **constat descriptif** : s'abstenir de prélever ou de circuler est un acte qui engage, et ce constat vaut en tout temps. Il ne prescrit aucun modèle institutionnel. Dans cette partie, la sagesse collective est mobilisée comme **grille de lecture** des pratiques — pour comprendre ce que fait, ou ne fait pas, tel mouvement — jamais comme **endossement** de ces pratiques. La distinction entre le constat

et la prescription est l'écueil principal de toute la partie ; on la tiendra fermement.

Chaque volet est rattaché au temps de la donnée qu'il gouverne, car la dimension politique se loge surtout dans la circulation — le deuxième temps.

Chapitre 12 — Le droit en évolution

Le premier tome et le deuxième ont esquissé trois directions d'évolution du droit. Ce chapitre les confronte au droit positif de 2026, sans plaider pour ni contre : il **cartographie l'écart** entre les concepts de la trilogie et le droit existant, à un instant donné. Cet écart se déplacera à mesure que le droit évolue — ce qui est cohérent avec le statut daté de la partie.

Les trois directions agissent sur des temps différents de la donnée, et les distinguer ainsi clarifie leur portée.

Le droit au non-prélèvement agit sur la **fabrication** (premier temps). Il viserait à reconnaître un droit à ne pas être prélevé — à ce que certains flux d'observation ne soient pas creusés du tout. Le droit positif en offre un écho partiel : le RGPD reconnaît un droit d'opposition (article 21) et fonde la plupart des traitements sur une base légale, parfois le consentement (article 6). Mais l'opposition s'exerce sur un traitement déjà en cours, et elle se heurte à des limites — l'intérêt légitime, l'obligation légale. Un droit au non-prélèvement serait plus amont : il porterait sur le creusement même du canal, non sur l'usage de l'eau déjà tirée. L'écart entre le concept et le droit positif est ici considérable, et le concept ne dit pas comment l'instituer — il pointe seulement la zone.

La transparence sur le geste d'annotation agit sur la **circulation** (deuxième temps). Elle viserait à rendre lisible, tout au long de la chaîne, qui a prélevé une donnée, selon quels critères, dans quelle intention — à restaurer la traçabilité du transport que le deuxième tome a montrée si fragile. Le droit positif s'en approche par un biais : l'AI Act européen fait émerger des obligations de documentation des corpus d'entraînement. Mais cette exigence reste cantonnée aux systèmes d'intelligence artificielle ; l'étendre à tout flux d'observation serait une autre affaire. L'écart, là encore, est cartographié, non comblé.

La péremption automatique se tient au **croisement de la fabrication et de la circulation**. Elle viserait à attacher à une donnée une durée de validité liée à sa fraîcheur, au-delà de laquelle elle s'effacerait. Le droit positif connaît la limitation de la conservation (article 5) et le droit à l'effacement (article 17), mais ces obligations sont conditionnées à la *finalité* du traitement, non à la *fraîcheur* de la donnée. Passer d'une obligation liée à la finalité à une obligation liée à la péremption serait un changement de logique. Un cas illustre joliment la difficulté : la

jurisprudence sur le droit à l'oubli a eu à connaître des registres de baptême — précisément le cas historique du chapitre 3 — et a jugé que l'intérêt à leur conservation prévalait. Une donnée peut être à la fois périmée pour un usage et conservée légitimement pour un autre : preuve que la péremption fonctionnelle et l'obligation juridique ne se recouvrent pas.

Cette cartographie n'est pas un plaidoyer pour réformer le droit dans tel sens. Elle montre où le droit positif et les concepts de la trilogie coïncident, et où ils divergent. Le volet pourrait, à terme, justifier une étude juridique autonome menée avec un juriste — hors du périmètre de cet ouvrage.

Chapitre 13 — Les coopératives de données

Le modèle coopératif est souvent présenté par analogie avec les coopératives agricoles : mutualisation des moyens, gouvernance démocratique, propriété collective. L'analogie est éclairante, mais elle masque une différence que le deuxième tome permet de nommer — et cette différence est le vrai fondement conceptuel des coopératives de données.

Une coopérative agricole mutualise des **biens rivaux** : un tracteur, un silo, une chaîne de conditionnement. Ce que l'un utilise, l'autre ne l'utilise pas au même moment. Une coopérative de données mutualise tout autre chose : un **fait partagé**. Le deuxième tome l'a établi — transporter une donnée ne la déplace pas, il en **multiplie les porteurs** ; elle ne se cède pas comme un objet, elle se partage comme un fait, sans que celui qui la transmet s'en dépossède. C'est cette propriété, et non l'analogie agricole, qui fonde la possibilité même d'une coopérative de données : on peut mettre en commun une donnée sans la perdre, la partager mille fois sans l'épuiser. Le modèle agricole donne la forme de gouvernance ; le partage du fait donne la matière.

Des formes existent en 2026, qu'il faut citer avec leur statut, car elles évoluent vite. MiData en Suisse, dans le domaine de la santé ; Salus.coop en Espagne ; des expérimentations de *data trusts* au Royaume-Uni ; diverses coopératives sectorielles ailleurs. Leur viabilité dépend de facteurs connus : un seuil critique de membres, une gouvernance qui tienne, un modèle économique, et la sortie du dilemme entre captation marchande et altruisme pur. Les secteurs propices sont ceux où la mutualisation du fait crée une valeur que nul membre seul ne pourrait produire : santé, agriculture, énergie, mobilité.

Les limites doivent être nommées autant que les promesses, sous peine de plaider au lieu de cartographier. Une coopérative peut devenir un nouveau monopole ; elle dépend souvent d'une technologie qu'elle ne maîtrise pas ; son financement initial est fragile ; et la traçabilité de la

chaîne, à mesure que les membres se multiplient, peut se dégrader. Le modèle est prometteur ; il n'est pas sans écueils.

Chapitre 14 — Souveraineté numérique et communs institutionnels

Ce chapitre traite des formes où une autorité — étatique, fondation, consortium — prend en charge un flux d'observation comme bien d'intérêt collectif. La question qui le traverse est celle que le troisième tome reformule en termes de **probité** : *qui répond du flux ?*

Les **projets de souveraineté** numérique cherchent à soustraire des flux critiques au modèle privé concentré. Gaia-X, en Europe, en est l'exemple le plus cité — avec un bilan, à ce jour, mitigé. Mention datée s'impose : ce qui se dit de Gaia-X en 2026 sera très probablement révisable dans trois à cinq ans, projet abouti, transformé ou abandonné. D'autres approches existent ailleurs, selon des logiques nationales ou régionales différentes. Le concept ne tranche pas entre elles ; il éclaire ce qu'elles ont en commun : déplacer la réponse au flux d'un acteur privé vers une instance publique ou souveraine.

Les **communs institutionnels** offrent un autre modèle de réponse. Des projets comme les grandes bases de connaissances ouvertes, les fondations de données scientifiques, les cartographies collaboratives, reposent sur une gouvernance par équipes éditoriales, comités scientifiques ou modération communautaire, et sur un financement par subventions, fondations ou micro-contributions. Ils incarnent une réponse répartie à la question « qui répond du flux ? » : non un propriétaire, mais une communauté organisée.

La question maîtresse demeure : *qui répond du flux ?* L'État, la coopérative, la fondation, le consortium — chaque réponse engage une organisation différente, des garanties différentes, des fragilités différentes. La probité du troisième tome, qui était l'éthique du lecteur individuel, trouve ici son pendant collectif : une institution est probe quand elle sait de quel flux elle répond, à quelle fraîcheur, devant qui. Le concept pose la question ; il laisse les formes y répondre diversement.

Chapitre 15 — Data activism et démocratisation

Ce dernier volet traite des mouvements citoyens qui construisent des contre-flux ou auditent les flux existants. Il est le plus exposé au risque de glisser de la cartographie au militantisme ; on le tiendra donc avec une vigilance particulière, en mobilisant la sagesse collective comme grille, non comme endossement.

Le fil conceptuel est ici le rapport entre **sagesse collective** et **auto-amplification**, tous deux du deuxième tome. Le torrent grossit par auto-amplification : la boucle creuse son propre lit, la circulation appelle la circulation. Les pratiques d'audit citoyen, de demande d'accès aux documents administratifs, de construction de contre-flux par la société civile, peuvent se lire comme des gestes qui s'opposent à cette auto-amplification — des actes par lesquels un collectif décide de ne pas laisser le torrent tout emporter. C'est précisément ce que la sagesse collective nomme : l'abstention, ou l'inflexion, n'est pas neutre, et elle ne devient une position pesable que là où le choix est rendu visible et délibéré.

Mais — et c'est la limite à tenir — dire cela, c'est *décrire la structure* de ces gestes, non les *approuver*. Le même cadre éclairerait aussi bien les pratiques inverses. La sagesse collective n'est pas un drapeau du militantisme des données ; c'est une grille qui permet de comprendre ce que fait n'importe quel acteur qui adopte ou refuse un régime de circulation. On l'applique ici aux mouvements citoyens parce que c'est l'objet du chapitre, non parce qu'on les endosse.

Des mouvements existent, à citer avec prudence et statut : des organisations de promotion de l'ouverture des données, des collectifs d'audit algorithmique, des mouvements de journalisme de données. Leurs pratiques émergentes — audit citoyen des algorithmes publics, contre-flux d'observation, formation à grande échelle à la lecture du flux — sont décrites, non recommandées. Quant à la **formation** à la littératie du flux dès l'école, elle rejoint directement le chapitre 9 : un citoyen qui sait lire un flux est mieux armé devant le torrent, quel que soit le modèle de gouvernance qu'il préfère par ailleurs.

Une mention savante, enfin, à situer sans s'y attacher. La critique du *dataïsme* — la dénonciation d'une pente où tout devient donnée — est conceptuellement voisine de l'ontologie processuelle, mais distincte : elle dénonce, là où la trilogie propose un cadre. Des travaux récents sur la gouvernance démocratique des données prolongent ces débats ; leur cadre est encore en formation, et on les mentionne comme chantier ouvert, non comme doctrine établie.

Ce que la cartographie a montré

Quatre volets, un même refus : celui de trancher. Le droit, les coopératives, la souveraineté, le militantisme sont autant de formes que prend, en 2026, la question politique de la donnée — et l'ontologie processuelle leur offre à toutes un langage, sans en couronner aucune. Ce que la traversée fait apparaître, c'est que la dimension politique se loge presque entièrement dans le deuxième temps, la circulation : c'est parce que la donnée se duplique et se partage comme un fait que sa propriété, sa gouvernance, son partage posent des questions sans équivalent dans le monde des biens

rivaux.

La limite de cette partie est inscrite dans sa date. Tout ce qui y est nommé — projets, institutions, régulations — peut avoir changé quand on la lira. C'est assumé : une mise en pratique qui voudrait parler des formes politiques sans se périmer ne parlerait de rien. Le fond, lui, tient : la circulation pose des questions politiques propres, le partage du fait fonde des modèles inédits, et la sagesse collective éclaire les choix sans les dicter. Les formes passeront ; la grille restera pour lire celles qui leur succéderont.

Fin de la Partie VI.

Clôture — La carte révisable

Texte rédigé en mai 2026.

Chapitre 16 — Statut de cet ouvrage et conditions de sa révision

Une trousse, pas une lampe

La trilogie était une lampe : elle éclairait la zone où regarder, sans dessiner la carte des récifs. Cet ouvrage est une trousse à outils. Il prend la lampe et fabrique, à partir d'elle, des objets maniables — des illustrations historiques pour reconnaître la structure, des études sectorielles pour la décliner sur un terrain, un chapitre sur le LLM pour affronter le saut le plus pressant, une boîte à outils pour agir, une cartographie politique pour situer les choix collectifs. Ces objets servent à faire ; or ce qui sert à faire s'use.

C'est la différence de statut qu'il faut tenir jusqu'au bout. La trilogie dit ce qui est vrai de la donnée pour longtemps. Cet ouvrage dit ce que cette vérité implique, en cette année, dans ces secteurs, avec ces outils. Le premier vieillit lentement ; le second vite. Cette dissymétrie n'est pas un défaut de l'ouvrage : c'est sa fonction même. Une mise en pratique qui ne se périmerait pas ne serait pas une mise en pratique — elle aurait renoncé à descendre au contact du présent.

Ce qui se périme et ce qui tient

On peut, dans chaque partie, séparer ce qui tient de ce qui passera.

Tient : la grille des trois temps, parce qu'elle vient de la trilogie. Les concepts — flux d'observation, deux temps, transport comme rapport, acte de lecture, trois chemins, sagesse collective — ne se périment pas, par construction. Tient aussi la structure de l'ouvrage : les quatre dimensions, la grille de traitement des cas, la structure-type des études sectorielles. Ce sont des formes stables, parce qu'elles épousent un fond stable.

Passera : tout le reste, à des rythmes divers. Les états des lieux sectoriels (cinq à dix ans). Les outils opérationnels (cinq ans). Les formes économiques et politiques, projets et régulations (trois à cinq ans). Le chapitre sur le LLM, le plus mouvant (deux à quatre ans, révision annuelle). Les cas historiques, eux, ne se périment pas — mais les études savantes qui les éclairent, oui, et les rapprochements conceptuels qu'on en tire restent révisables.

Le lecteur est ainsi prévenu : il tient un document daté, dont chaque partie porte sa propre échéance. La datation systématique, présente en tête de chaque partie, n'est pas une coquetterie ; c'est la probité de l'ouvrage envers son lecteur — lui dire de quand date ce qu'il lit, pour qu'il juge lui-même de ce qui a pu changer.

Comment cet ouvrage devra être révisé

Réviser cet ouvrage ne demande pas de toucher au fond. Le fond est dans la trilogie, et il ne bouge pas. Réviser consiste à reprendre les formes datées à mesure qu'elles vieillissent, sans jamais altérer la grille qui les organise.

Trois règles pour cette révision. La première : ne jamais introduire, en révisant, un concept absent de la trilogie. Si un fait nouveau du présent semble exiger un concept neuf, c'est presque toujours qu'on a mal mobilisé un concept existant — ou que le fait relève d'un autre cadre. La grille conceptuelle reste close ; seules les illustrations changent. La deuxième : réviser partie par partie, selon les échéances propres à chacune, plutôt que l'ouvrage en bloc. Le chapitre sur le LLM appellera plusieurs révisions quand l'atlas historique n'en demandera aucune. La troisième : à chaque révision, redater. Une forme révisée qui ne porte pas sa nouvelle date ment à son lecteur.

La carte et les cartographes

Cet ouvrage est une carte. Il dessine, à ce moment, ce que l'ontologie processuelle implique pour le geste. Mais une carte ne remplace pas le voyage. Devant une donnée réelle, dans un secteur réel, à une date réelle, c'est au praticien situé de juger — quel âge est acceptable, quelle fidélité suffit, quelle posture adopter face au sur-débit. L'ouvrage ne tranche pas ces questions ; il apprend à les poser, et il rappelle qu'elles ont une structure commune, déjà rencontrée ailleurs, déjà nommée par la grille.

C'est tout ce qu'une mise en pratique peut faire, et c'est sa juste mesure. Elle ne décide pas à la place du praticien — ce serait retomber dans l'ontologie objectale, où la donnée et sa bonne gestion seraient des choses qu'on possède une fois pour toutes. Elle outille un jugement qui reste, lui, irréductiblement situé. Le concept éclaire la zone où il faut regarder ; la trousse aide à y travailler ; mais la carte des récifs, dans

chaque eau particulière, c'est encore et toujours au navigateur de la dresser.

Reste à le faire. L'ouvrage est écrit pour que ce travail ne parte pas de zéro.

Fin de l'ouvrage.

Du fleuve qu'on observe aux ateliers où l'on travaille

—

Trois tomes ont suivi la donnée le long de sa vie :

Sa fabrication, dans *Le fleuve et le canal* ; sa circulation, dans *De la source au torrent* ; sa réception, dans *La trace ou la donnée*. Ensemble, ils ont posé une grille – le flux d'observation, les deux temps, le transport, l'acte de lecture, la péremption – et défendu une posture : dire ce qui est vrai de la donnée pour longtemps, sans descendre aux outils du moment, qui se périment.

Cette retenue laissait le praticien sur sa faim. L'enseignant qui veut former à la lecture du flux, le soignant qui partage un dossier, le décideur qui évalue un score automatisé, celui qui voudrait utiliser un grand modèle de langage en pleine lucidité : aucun ne trouvait dans la trilogie les outils de son métier.

—

Le canal et les ateliers comble ce que la trilogie avait laissé vide, sans jamais y toucher. Il ne raconte pas : il outille. Cas historiques où la grille se reconnaît à nu, études sectorielles déclinées aux trois temps de la donnée, pratiques pour vivre avec le LLM, boîte à outils maniable, cartographie des formes politiques de la donnée — sans plaidoyer.

Aucun de ces outils ne dicte une conduite : chacun est un déclencheur de questions, daté et révisable. Car le concept éclaire la zone où regarder ; il ne dresse pas la carte des récifs. Cette carte, dans chaque eau particulière, c'est au navigateur de la tracer. Cet ouvrage est la trousse qu'il emporte à bord.