

À L'ÉPREUVE DU FLUX

Volume d'architecture de la trilogie
« Une ontologie processuelle de la donnée »

David Bories

À l'épreuve du flux — Volume d'architecture de la trilogie « Une ontologie processuelle de la donnée »

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

Juin 2026.
1.0.0-poche

contact@artisanat-de-la-donnee.fr

Table des matières

À l'épreuve du flux — Volume d'architecture	1
§0. Seuil — le statut de ce volume	1
§1. Les douze thèses	1
§2. Le schéma consolidé bouclé	4
§3. Les trois familles de constitution	5
§4. Les couches du datum	6
§5. L'ontologie minimale	6
§6. L'apparat formel	9
§7. La carte des voisins	11
§8. La formule prudente d'originalité	12
§9. La formule de synthèse	12

À l'épreuve du flux — Volume d'architecture

§0. Seuil — le statut de ce volume

Les tomes parlent la langue du canal ; ce volume parle la langue des catégories. Les deux disent le même fond.

Ce volume n'est ni un résumé, ni un quatrième tome. La synthèse de la première édition condensait le récit ; le volume d'architecture assemble le registre systématique : les thèses, le schéma, les familles, les couches, l'ontologie minimale, l'apparat formel, la carte des voisins, les deux formules.

Sa règle de circulation est celle de toute l'édition : un apport, un foyer. Chaque apport a son lieu unique de démonstration dans les tomes ; ici il s'énonce, se relie et, quand il le faut, se formalise. Les renvois indiquent où la démonstration vit.

L'apparat vit ici et nulle part ailleurs : le vecteur d'état, les temps opérationnels, le paquet de persistance, la formalisation du sur-débit. Dans les tomes : zéro occurrence pour le vecteur d'état et le paquet de persistance ; une mention unique, en prose, pour les temps opérationnels. La formule de synthèse longue clôt ce volume ; elle ne paraît jamais dans les tomes.

§1. Les douze thèses

L'épine dorsale de l'édition. Chaque thèse s'énonce ici ; sa démonstration vit au foyer indiqué. Trois mouvements suivent la dramaturgie de la trilogie — la source, le torrent, le lecteur.

Premier mouvement — la constitution

Thèse 1. *Rien n'est donnée sous toute description par nature. Le statut de donnée est constitué dans une pratique réglée.* Renvoi : A1 — T1 « Préface d'édition » + « Partie I — Les deux ontologies de la donnée » ; appui A5. Contre la donnée par nature, sans retomber dans le décret arbitraire — la constitution est réglée et imputable, portée par les trois intentions plutôt que par un vouloir à chaque inscription. L'entité est processuellement constituée, historiquement individualisée, relationnellement maintenue.

Thèse 2. *Toute donnée empirique est contrainte par un matériau qui lui résiste, mais ce matériau ne détermine jamais seul sa description.* Renvoi : B9 — T1 « Le flux et l'annotation » ; appui A3. Double garde-fou. Contre l'arbitraire : le matériau résiste, toute description n'est pas soutenable. Contre le réalisme naïf : annoter, c'est choisir — le réel pertinent pour une observation excède toujours les découpages produits par elle, et le non-prélevé fait partie du geste.

Thèse 3. *La trace n'est ni transparente ni nécessairement univoque. Elle résulte elle-même d'un processus de formation et de conservation.* Renvoi : A4 — T1 « La trace et le lecteur », pilier I. La traçabilité est différentielle — marquage causal, trace conservée, trace accessible, trace interprétable — et l'équivoque demeure possible : des processus différents laissent des marques semblables. « Aucune connaissance n'a d'autre voie que la trace » demeure ; « tout est lisible, partout » tombe.

Thèse 4. *L'inscription ne capture pas une valeur nue. Elle stabilise un résultat produit par une opération, avec un régime d'incertitude ou de preuve.* Renvoi : A3 — T1 « Partie IV — L'acte d'annotation » + « Les deux temps de la donnée ». Le résultat porte son opération — objet visé, règle, instrument, modèle, incertitude. De là le partage des temps : le fait documentaire (tel dispositif a produit tel résultat) est irréversible ; l'assertion sur le phénomène reste révisable.

Deuxième mouvement — l'identité et la persistance

Thèse 5. *Toute occurrence est matérielle, mais l'identité d'un datum n'est pas celle d'un support particulier.* Renvoi : A6 — T2 « Les propriétés du mouvement » ; écho T1, la duplication. Dépendance matérielle sans identité de support. Le datum d'origine est identifié par son acte — quel prélèvement, sur quel fragment du réel, sous quels critères — non par le nombre de ses porteurs. Les occurrences se comptent ; le datum demeure un : la copie sans division, vue du support.

Thèse 6. *L'identité est historique. Elle repose sur l'origine et la provenance ; la copie multiplie les occurrences, la transformation engendre une filiation.* Renvoi : A6 — T2 « Les propriétés du mouvement » ; appui A3 (inscrire avec son origine) ; modèle PROV en appareil. Le faisceau d'individuation — opération de constitution, objet visé, temps pertinent, résultat, provenance. Deux enregistrements de même valeur

peuvent être deux data ; deux fichiers différents peuvent réaliser le même datum. La provenance est constitutive, non décorative ; pour une dérivée, elle inclut le modèle et les données parentes.

Thèse 7. *Persister exige davantage que survivre matériellement. Il faut maintenir, selon l'objectif, intégrité, authenticité, représentation, contexte et possibilité d'accès.* Renvoi : B2 — T1 « Les deux temps de la donnée » pour les images ; démonstration formelle ici même, §4 (le paquet de persistance) et §6 (le vecteur d'état). Un datum durable n'est pas un couple valeur-support mais un paquet stratifié destiné à une communauté qui saura le lire ; la perte d'une couche ne détruit pas toujours le tout, elle réduit ce que la donnée peut soutenir. Les deux morts demeurent, dans les tomes, les images des effondrements typiques ; la mort universelle est un idéal-limite.

Thèse 8. *La fraîcheur n'est pas l'âge, ni un verdict d'usage. C'est la récence de l'acte : la proximité de l'acte d'inscription au présent du monde, qui ne se rattrape jamais.* Renvoi : B3 — T1 « Les deux temps de la donnée ». La péremption décomposée devient validité pour un usage — fonction de l'objet, de la question, de l'instant, de la tolérance. La même inscription peut être périmée pour l'alerte et exacte pour l'histoire ; l'énoncé qui porte sur le passé ne vieillit pas comme celui qui prétend dire le présent. La fraîcheur, elle, se constate : récence de l'acte, la même pour tous les usages — ce qui varie d'un usage à l'autre est le verdict de validité, fonction de l'objet, de la question, de l'instant et de la tolérance.

Troisième mouvement — la circulation, la lecture, le torrent, la boucle

Thèse 9. *La circulation peut conserver, altérer ou engendrer. Elle doit être évaluée par les invariants qu'elle prétend maintenir.* Renvoi : A6, B4, B8 — T2 « Les propriétés du mouvement » + « L'acte d'échange ». Les cinq identités — origine, contenu, encodage, binaire, filiation — donnent la grille : la copie fidèle maintient, la conversion déplace, la dérivation engendre. L'échange étend la disponibilité et la possibilité d'usage ; il n'accomplit rien de plus. Et la reprise à l'autre bout n'a pas un support — elle a des conditions.

Thèse 10. *La lecture actualise sans constituer de nouveau la donnée. Elle est située, sélective et coûteuse ; elle peut produire de nouvelles données dérivées.* Renvoi : A2 — T3, document 1 (+ document 4) ; appuis C2 « Le lecteur situé », B5. Le statut de donnée constituée ne dépend pas de la relecture ; l'actualisation en dépend — c'est elle, et elle seule, que la lecture ajoute. L'échelle remplace la retombée en trace partout sauf au cas-limite. Le lecteur est fini et situé — la fenêtre est multi-contraintes — et la machine qui lit est un lecteur situé de plus. Quand la lecture réinscrit, elle ouvre une filiation : le sens produit peut redevenir donnée.

Thèse 11. *Le torrent est un régime de désajustement entre production, sélection, traitement et échéance, susceptible d'être amplifié par des*

boucles sociotechniques. Renvoi : C1, B6, B7 — T2 « Le torrent » ; T3 « Le sur-plus » ; formalisation ici même, §6. Incompatibilité de cadence, non bonté des arrivées — distinguer débit brut, pertinent, prioritaire, traitable, lu. Les sept issues de la file ne sont pas ontologiquement équivalentes : une donnée en attente ne devient pas trace, une rejetée reste donnée, une expirée pour un usage peut servir un autre. Pas de classement exhaustif à coût fini — mais fenêtres, résumés, priorités, archives partielles : ce que cadrer et déléguer organisent. La règle qui transforme le brut en présenté est une décision — politique au sens propre.

Thèse 12. *Le processus est réflexif. Les données contribuent à transformer le réel, les catégories et les infrastructures qui produiront les données suivantes.* Renvoi : A8 — T3, clôture ; appuis T1 partie VI, T2 « L'auto-amplification ». La chaîne se ferme — action, réel modifié, nouvelles traces, nouveau régime d'observation — et la boucle peut stabiliser ou amplifier ; les délais peuvent rendre une décision correcte au calcul, nocive à l'application. Le passage de la lecture à l'action exige le pont observation / inférence / intervention : une prédiction n'est pas une donnée fraîche sur le futur. L'ontologie devient responsabilité de la chaîne — chaque maillon répond de son geste et transmet au suivant de quoi répondre ; la provenance en est la forme documentaire, la probité décomposée en donne les sept lieux.

Le mot *processuel* qualifie ainsi en priorité le mode de constitution et d'existence relationnelle du datum, non l'obligation de classer le datum lui-même comme événement. La priorité explicative revient au processus — sur l'identité, la persistance, la circulation, l'actualisation et la valeur de la donnée.

§2. Le schéma consolidé bouclé

Le mouvement directeur de la trilogie, avec ses médiations, et sa fermeture : de la trace à l'action, et de l'action aux nouvelles traces.

```

flowchart TD
    R["Processus ou événement du réel"] --> M["Marquage causal ou interaction"]
    M --> T["Trace conservée et accessible"]
    T --> O["Observation située : objet, règle, instrument, modèle"]
    O --> D["Datum d'origine : résultat, temps, incertitude, provenance"]
    D --> P["Paquet de persistance : occurrence, code, contexte, fixité"]
    P --> X["Copie, transport ou transformation"]
    X --> A["Occurrence accessible à un lecteur"]
    A --> L["Lecture et interprétation situées"]
    L --> I["Inférence, décision ou nouvelle inscription"]
    I --> U["Action sur le réel"]
    U --> R
    U -- "réel modifié, nouvelles traces" --> R

```

Le schéma indique des dépendances logiques, non une chronologie obligée. Une donnée institutionnelle est constituée par un acte qui

institue plutôt que par une mesure ; une dérivée part d'autres données plutôt que d'une trace primaire ; une copie peut circuler sans jamais être lue ; une lecture peut ne conduire à aucune action (§3).

La boucle peut être stabilisante ou amplificatrice. Les délais entre observation et action peuvent produire des oscillations, ou rendre une décision correcte au moment du calcul et nocive au moment de son application ; une politique fondée sur des données modifie parfois les comportements mêmes qui produisent ces données. Le passage de la lecture à l'action exige en outre le pont des trois rapports au réel — observation, inférence, intervention (thèse 12 ; foyer : T3, clôture).

Le torrent lui-même peut naître d'une boucle endogène : davantage de données rend possible davantage d'automatisation ; l'automatisation multiplie les points de captation et les décisions ; ces décisions modifient le réel et produisent de nouvelles traces ; ces traces deviennent de nouvelles données ; l'augmentation du flux justifie encore plus d'automatisation (foyer : T2, l'auto-amplification).

§3. Les trois familles de constitution

Trois manières, pour une donnée, de naître ; le prélèvement de mesure cesse d'être le prototype implicite de toute donnée. Foyer : A7 — T1, la constitution ; échos T2 et T3.

Observationnelle. Événement ou état → interaction → trace → mesure ou description → résultat inscrit. Régimes : mesure physique, observation biologique, témoignage, photographie, relevé. Une interaction laisse une trace ; l'acte la constitue en résultat sous des critères.

Institutionnelle. Acte social ou juridique → catégorie et règle constitutive → enregistrement. Le mariage inscrit, le vote compté, le titre enregistré, le solde arrêté : sous la règle constitutive, l'inscription ne décrit pas un phénomène indépendant — elle participe à faire exister le fait qu'elle enregistre. Sa vérité ne se vérifie pas contre le monde ; elle se conteste devant la règle et l'autorité qui l'ont instituée. Sa provenance nomme l'acte, la règle, l'institution.

Dérivée. Données parentes + modèle ou opération → résultat transformé → provenance nouvelle. Genre dont le corpus tient trois espèces — la consolidée, la vraisemblable, le récit au sens technique — et que la famille étend : correction, traduction, agrégation, annotation, classification, estimation, prédiction, donnée synthétique. Ni prélèvements directs, ni copies fidèles : toute dérivée ouvre une filiation, et sa provenance inclut le modèle et les données parentes ; la traduction n'en relève que si un invariant déclaré est altéré — préservés, elle est une conversion fidèle, occurrence nouvelle du même datum —, l'opération et le texte source étant inscrits dans la provenance dans les deux cas.

Les chaînes réelles tressent les trois familles : un relevé observationnel entre dans une consolidation dérivée qui fonde un acte institutionnel, dont l'enregistrement nourrira d'autres dérivées.

§4. Les couches du datum

Un datum durable dépend d'un ensemble stratifié :

1. **contenu** — valeur, symbole, proposition ou structure ;
2. **occurrence** — réalisation matérielle déterminée ;
3. **encodage** — règles qui font correspondre états matériels et symboles ;
4. **schéma** — organisation des champs et des relations ;
5. **provenance** — origine, agents, activités et transformations ;
6. **contexte** — objet visé, unités, population, méthode, finalité ;
7. **information de représentation** — ce qu'une communauté doit savoir pour interpréter ;
8. **conditions d'accès** — localisation, droits, interfaces, moyens techniques ;
9. **régime de validité** — question, instant, tolérance et usage auxquels le résultat peut répondre.

C'est ici que le **paquet de persistance** prend corps : ce qu'est un datum durable — non un couple valeur-support, mais un paquet de ces neuf couches, destiné à une communauté qui saura le lire. L'archivistique de la conservation numérique en a établi le principe ; le corpus le reprend à son compte.

La perte d'une couche ne détruit pas toujours le tout ; elle réduit ce que la donnée peut soutenir — et la fait descendre le long de l'échelle, d'accessible à intelligible, jusqu'à la muette et, à l'idéal-limite, la morte. Renvois : conditions de reprise interprétative, foyer T2 « L'acte d'échange » ; échelle, foyer T3, document 1.

§5. L'ontologie minimale

Rendre explicites les catégories et les relations sans lesquelles les énoncés restent ambigus. Registre sobre — prose et tableaux ; il ne s'agit pas de produire une ontologie informatique, mais de fixer ce que les tomes démontrent.

Primitives

Symbole	Catégorie	Définition minimale
<i>R</i>	Processus réel	événement, état ou activité auquel une donnée pourra être rapportée
<i>T</i>	Trace	modification ou occurrence causalement reliée à un processus antérieur
<i>O</i>	Opération de constitution	activité réglée qui constitue un résultat : elle sélectionne, transforme ou décrit une trace ou un phénomène (voie observationnelle), institue un fait sous une règle constitutive (voie institutionnelle), ou dérive un résultat d'autres données (voie dérivée)
<i>D</i>	Datum d'origine	résultat historiquement individualisé produit par une opération de constitution
<i>C</i>	Occurrence	réalisation matérielle d'un datum sur un support déterminé
<i>P</i>	Provenance	ensemble des relations d'origine, d'activité, d'agent et de dérivation
<i>X</i>	Transformation	opération qui copie, convertit ou dérive une occurrence ou un datum
<i>L</i>	Lecture	activité par laquelle un lecteur produit une interprétation à partir d'une occurrence
<i>Q</i>	Question ou usage	problème pour lequel une donnée peut être pertinente ou valide

Symbole	Catégorie	Définition minimale
U	Action	intervention, décision ou inscription nouvelle susceptible de modifier le réel

Relations nécessaires

- $causedBy(T, R)$ — la trace est causalement reliée au processus ;
- $uses(O, T)$ — l'observation mobilise une trace ;
- $generates(O, D)$ — l'opération constitue un datum d'origine ;
- $realizes(C, D)$ — une occurrence matérielle réalise ce datum ;
- $encodedBy(C, E)$ — l'occurrence utilise un encodage ;
- $derivedFrom(D_2, D_1)$ — un datum descend d'un autre ;
- $copyOf(C_2, C_1)$ — une occurrence a été produite comme copie d'une autre ;
- $accessibleTo(C, r, t)$ — une occurrence est accessible à un lecteur donné à un moment donné ;
- $interpretableBy(D, r, t)$ — le lecteur dispose des ressources pour l'interpréter ;
- $readAt(D, r, t)$ — une lecture a effectivement eu lieu ;
- $about(D, R)$ — le datum prétend porter sur un processus ou un état ;
- $validFor(D, Q, t, \epsilon)$ — le datum est acceptable pour une question, à un moment, selon une tolérance ;
- $affects(U, R)$ — une action modifie le processus observé ou son environnement.

Critères d'identité

Un datum d'origine est individualisé par un faisceau minimal : **opération de constitution + objet visé + temps pertinent + résultat + provenance**. Deux occurrences peuvent réaliser le même datum. Une transformation produit un datum nouveau si elle change de manière significative le contenu, le statut épistémique, l'objet visé ou le régime de validité ; la frontière entre conversion fidèle et dérivation n'est pas naturelle — elle se définit par les invariants que la pratique entend préserver.

Cinq identités départagent les cas : **identité d'origine** (même événement de constitution documenté) ; **identité de contenu** (même contenu logique selon un critère donné) ; **identité d'encodage** (même représentation structurée) ; **identité binaire** (même suite de bits) ; **continuité de filiation** (relation documentée de transformation). « Les deux copies sont la même donnée » devient défendable si la phrase signifie : deux occurrences réalisent le même datum d'origine ; elle devient fausse si elle prétend qu'il n'existe qu'un seul objet matériel.

Les cinq états

Relatifs à un lecteur et à un instant, cumulatifs. Foyer : A2 — T3, document 1.

- **constituée** — un résultat a été produit et stabilisé selon une règle ; son existence ne doit rien à la lecture ;
- **accessible** — une occurrence peut être atteinte par ce lecteur-ci ;
- **intelligible** — le lecteur ou sa communauté possède les ressources d'interprétation ; inintelligible pour ce lecteur-ci, la donnée est sa donnée muette ; inintelligible pour tous, à l'idéal-limite, la donnée morte ;
- **actualisée** — une lecture effective en a produit un sens dans une situation ;
- **opérante** — cette actualisation a contribué à une inférence, une coordination ou une action.

Une donnée peut être constituée sans être accessible, accessible sans être intelligible, intelligible sans être lue, lue sans rien nourrir — et utilisée sans être valide.

§6. L'apparat formel

Les formalismes vivent ici ; dans les tomes, la prose du canal.

Le vecteur d'état

L'état d'une donnée, pour un lecteur r , un instant t , un usage u , se décompose en sept composantes indépendantes :

$$S(D, r, t, u) = (E, F, I, A, X, G, V)$$

- E — existence d'au moins une occurrence ;
- F — fixité de l'occurrence examinée ;
- I — intégrité du contenu ;
- A — authenticité et continuité de provenance ;
- X — accessibilité pour le lecteur r au temps t ;
- G — intelligibilité pour ce lecteur ou sa communauté ;
- V — validité pour l'usage u .

Chaque composante évolue sans entraîner les autres. Les deux morts demeurent, dans les tomes, les images de ses effondrements typiques ; la mort universelle est un idéal-limite — on parle rigoureusement d'inintelligibilité actuelle pour une communauté définie, de perte documentée des informations de représentation, d'indécidabilité pratique de la reconstruction. Foyer des images : B2 — T1 « Les deux temps » ; T3, document 1.

Les temps opérationnels

Le déploiement opérationnel des deux régimes de temps. Foyer des régimes : B1 — T1 « Les deux temps ».

Temps de l'**événement** — quand ce qui est représenté s'est produit ; de l'**observation** — quand une interaction avec le phénomène a eu lieu ; de l'**inscription** — quand le résultat a été fixé ; du **traitement** — quand il a été transformé ou agrégé ; de la **transmission** — quand il a quitté un nœud ; de la **réception** — quand une occurrence est devenue disponible ailleurs ; de la **lecture** — quand un lecteur l'a interprétée ; de l'**action** — quand elle a contribué à modifier le réel.

Coïncidents dans le dispositif simple, divergents dans les infrastructures réelles : l'ordre causal ne se confond pas avec l'ordre des horloges, et le temps du processus se documente — il ne se suppose pas unique. La datation d'un temps opérationnel se rapporte à une source ou version de référence, à un instant et une horloge, à un niveau de cohérence attendu ; la fraîcheur est la récence de l'acte ainsi daté (thèse 8), et la validité pour un usage demeure fonction de l'objet, de la question, de l'instant et de la tolérance.

La formalisation du sur-débit

Le sur-débit se formalise à partir de : λ , taux d'arrivée des éléments susceptibles d'être traités ; μ , capacité effective de traitement ou de lecture ; B , stock en attente ; W , temps moyen d'attente ; d , échéance au-delà de laquelle un élément perd sa valeur pour l'usage ; p , politique de sélection, priorité, agrégation ou rejet. Quand les arrivées excèdent durablement le service — $\lambda > \mu$ — la file croît si aucun mécanisme de rejet, de synthèse ou de réduction n'intervient ; dans le domaine des files d'attente, la loi de Little, $L = \lambda W$, relie le stock moyen, le taux d'arrivée et le temps de séjour, moyennes prises sur une période longue, en régime stationnaire et sous conservation des unités ; son application au flux de données est une analogie de structure, la phase de sur-débit croissant, non stationnaire, étant hors de son domaine, et aucune valeur quantitative n'en est tirée.

Sept issues de la file, ontologiquement distinctes : **retard accumulé** ; **expiration en file** ; **rejet explicite** ; **perte silencieuse** ; **agrégation** ; **priorisation** ; **délégation**. Une donnée en attente ne devient pas trace ; une donnée rejetée reste donnée ; une donnée expirée pour un usage peut rester utile pour un autre.

Cinq débits à ne pas confondre : **brut**, **pertinent**, **prioritaire**, **traitable**, **effectivement lu**. La règle qui transforme le débit brut en débit présenté au lecteur est une décision — politique au sens propre. Foyers : C1, B6, B7 — T2 « Le torrent » ; T3 « Le sur-plus ».

§7. La carte des voisins

L'absence trouvée n'est pas une preuve d'absence : une recherche peut manquer un vocabulaire, une langue, une discipline voisine, un chapitre non indexé. Cinq registres d'originalité se distinguent donc : **terminologique** (l'expression est-elle nouvelle ?), **propositionnelle** (telle thèse précise a-t-elle déjà été formulée ?), **architecturale** (ces propositions ont-elles déjà été articulées dans ce schéma ?), **déductive** (les conséquences tirées de l'architecture sont-elles nouvelles ?), **d'application** (le cadre éclaire-t-il des objets autrement ?).

Six familles voisines, avec leur recoupement et leur différence ; le détail des références vit dans les bibliographies raisonnées des trois tomes — l'ensemble à celle du tome 1, les familles propres au mouvement et à la reprise aux tomes 2 et 3.

Donnée construite, non brute — Gitelman (dir.), Markham. Recoupement : non-préexistence et caractère construit. Différence : la trilogie en tire une architecture ontologique couvrant persistance, circulation et lecture.

Donnée relationnelle — Leonelli. Recoupement : dépendance au contexte, à l'usage et à la pratique. Tension féconde : un objet peut acquérir le statut de donnée par son rôle probatoire ultérieur, ce qui conteste un statut entièrement fixé à l'inscription ; l'édition y répond en distinguant constitution et actualisation, et par l'intention d'usage qui recrute.

Donnée comme processus — Jones. Recoupement : rejet de la donnée comme ressource objective stable. Différence : ce travail relève de la théorie des organisations et de la performance contingente ; il n'établit pas l'architecture d'ensemble — mais il interdit d'affirmer qu'aucun cadre existant ne pense la donnée en régime processuel.

Trajectoires et matérialité — les *data journeys* : Bates, Lin et Goodale ; Leonelli et Tempini. Recoupement : mouvement matériel, transformation, réutilisation, changement de contexte. Tension féconde : la circulation ne préserve pas toujours un datum invariant, elle peut fabriquer l'objet qui circule — ce que la thèse 9 intègre : conserver, altérer, engendrer.

De la trace aux données — Rheinberger. Recoupement : la médiation expérimentale de la trace à la donnée. Différence : la trilogie généralise cette chaîne au-delà des sciences expérimentales et la relie à la circulation, à la lecture et au torrent.

Philosophie des processus — Whitehead, Bergson, Simondon. Le socle métaphysique : priorité du devenir, individuation, durée, relations. Il empêche de présenter le renversement substance/processus comme inédit en lui-même.

Où l'originalité peut raisonnablement se situer : dans la **combinaison**. Huit traits la composent — la chaîne complète, du réel à la trace, de la constitution à la circulation, de la lecture à l'action et au réel modifié ; le

flux d'observation comme terme médian entre réel et données consolidées ; les deux régimes temporels, opérationnellement déployés ; l'articulation datum d'origine, copies, filiation ; la jonction entre micro-ontologie de la donnée et macro-dynamique du torrent ; l'intégration du lecteur situé et de sa responsabilité dans le même schéma ; la boucle par laquelle la production de données transforme le réel dont elle tirera les données suivantes ; la tentative de déduire une éthique minimale de la structure même du processus.

La revue de portée est différée, programmée après publication ; jusqu'à son achèvement vaut la formule prudente.

§8. La formule prudente d'originalité

À notre connaissance, aucun travail identifié ne réunit exactement cette architecture ; plusieurs traditions en ont établi indépendamment des composantes majeures.

La préface d'édition porte cette phrase ; les bibliographies raisonnées en portent le détail. La revendication se recentre sur l'**originalité de composition et de déduction** : la trilogie propose une synthèse architecturale de la donnée comme processus de constitution, de persistance, de circulation, de reprise et de rétroaction, et en déduit une théorie du torrent et de la responsabilité du lecteur.

Une thèse nouvelle devient durable non lorsqu'elle prétend n'avoir aucune dette, mais lorsqu'elle sait exactement ce qu'elle reprend, ce qu'elle transforme et ce qu'elle ajoute.

§9. La formule de synthèse

La version condensée du fond ; elle vit ici, jamais dans les tomes.

Une donnée est un résultat historiquement individualisé, produit sous une description par une opération réglée et située, réalisé sur un ou plusieurs supports, accompagné des conditions qui permettent d'en établir l'origine et de l'interpréter, susceptible d'être copié ou transformé, et dont l'accessibilité, l'intelligibilité, la validité et les effets dépendent des relations dans lesquelles il est repris.

Cette définition ne remplace pas les tomes ; elle en condense le fond dans la langue des catégories.

Nous ne vivons pas entourés de données.

—

Un jour de mai 2026, je demandai à l'IA :

«

Produis un documentaire étayé, complet, argumenté, propre, sourcé, sur les origines de la donnée dans l'humanité depuis les origines les plus anciennes qui ont été repérées :

Comment sont apparues les premières données collectées par l'homme, sous quelle forme ? Pour quels usages ? Dans quels buts ?

Puis détaille les évolutions majeures qui ont conduit la civilisation contemporaine à être dépendante des usages de la donnée à travers le numérique et son flot incessant de données.

Comment l'homme a pris conscience de la donnée, de son potentiel ?

Comment l'a-t-il collectée, produite et exploitée au fil des millénaires ? »

L'IA s'exécuta.

Et l'essai commença.

—

Dire que l'auteur, au moment de produire cet essai, ne l'avait pas en tête, c'est ne pas tenir compte de près de 20 ans d'observation de ce qui se trame en sous-bas : au début des câbles, un écran, un clavier. Puis rapidement, du portable, un réseau, du tactile, du mobile. Et maintenant, des applications, des utilisateurs, des codes et une vie numérique omniprésente dont le flot de données charrie autant d'intérêts pour le présent que d'interrogations sur le devenir.

Dans le fond, cet essai dresse le portrait du monde qui bouge.

Sur la forme, il en fige l'image d'une société contemporaine.