

Phronesis

*Aperçu d'une œuvre en cours — une posture, quatre outils, onze
normes et une chaîne qui en fait du code.*

AUTEUR Kaspar Brönnimann

DATE 14 août 2026

VERSION v1

À quoi sert cette note

Lors d'une formation, un outil montre à une chef de projet une formulation — lisse, plausible, à l'apparence vérifiée. Elle la reprend. Un an plus tard, il s'avère que la formulation contient une référence qui n'existe pas.

Ni l'outil ni la chef de projet n'ont menti. Tous deux ont fait de leur mieux. Le dommage n'en est pas moins réel.

Cette note ordonne une œuvre qui veut précisément éviter de tels cas — non par de meilleures instructions, mais par une autre posture fondamentale : l'énoncé assumé demeure réservé. Les outils peuvent calculer et interpréter ; ils ne doivent pas juger. Ce qui suit est l'ordre à partir duquel cette posture devient logiciel.

Six niveaux portent l'œuvre, une chaîne en fait du code, un atelier les produit tous. Qui veut savoir comment cet ordre est né et ce qu'il exige le trouvera dans les pages qui suivent.

S O M M A I R E

— Avant-propos — À quoi sert cette note

I Préface — Ce qui est ici réuni

II La posture — Phronesis

III Les quatre outils

IV Les instanciations

V L'atelier — Traverse : l'établi

VI Les onze normes de la suite

VII La chaîne — Comment les normes deviennent code

VIII L'infrastructure — Capabilities, skills, code

IX Le flanc — Science et présence publique

X L'état actuel

XI Conclusion — La figure unique

Préface — Ce qui est ici réuni

Cette note n'est pas un recueil. C'est une carte.

Cette note n'est pas un recueil. C'est une carte.

Cette note ordonne une œuvre qui, à travers quatre outils, leurs instanciations concrètes, onze normes et deux publications scientifiques, tient la même posture. Ce n'est pas une anthologie, mais une carte. Qui a la carte peut situer chaque artefact — et voir pourquoi aucun d'eux ne tient seul.

Ce qui se trouve ici est né en de nombreux endroits, à de nombreuses occasions et dans différentes langues — allemand, anglais, code, document, diagramme, dialogue. Un observateur extérieur pourrait avoir l'impression que quelqu'un travaille sur beaucoup de choses sans lien. L'impression trompe. Il n'y a qu'une seule chose ; la multiplicité en est l'exposition.

Cette chose s'appelle Phronesis : la sagesse pratique, la capacité de peser, de juger et d'assumer la responsabilité dans le cas particulier concret. C'est une posture, non un outil. Mais elle porte une famille d'outils qui rend cette posture applicable au quotidien — et elle porte un atelier qui non seulement construit ces outils, mais les fonde de manière vérifiable.

L'ensemble comporte six niveaux et, à côté d'eux, un flanc :

- Niveau 1 — Posture. Phronesis comme manifeste et principe directeur.
- Niveau 2 — Outils. Aletheia, Mnemosyne, Methodos, Kairon.
- Niveau 3 — Instanciations. Produits concrets : ALETHEIA Radar, couche de pseudonymisation, kennora, HERMES PIA, KAIRON Dashboard, DigiTwin.
- Niveau 4 — Normes. Onze corpus de règles qui se tiennent au-dessus de toutes les instanciations.

- Niveau 5 — Chaîne. La chaîne de développement d'agents qui transforme les normes en code.
- Niveau 6 — Infrastructure. Capabilities, skills, code — la fondation technique partagée.
- Flanc — Science et présence publique. Deux publications scientifiques et un dépôt silencieux.

La posture — Phronesis

Niveau 1. Ce qui ne peut être logiciel, parce qu'il doit d'abord être porté.

Phronesis est une famille d'outils ouverte pour des décisions réfléchies dans un monde façonné par l'intelligence artificielle. Mais la première et la plus importante phrase est celle-ci : ce n'est pas un logiciel. C'est une posture, à laquelle servent trois conditions : elle doit être énoncée, elle doit être partagée, et elle doit trouver sa forme dans des outils — sinon elle reste opinion.

La référence est Aristote. Dans l'Éthique à Nicomaque (livre VI), il distingue trois manières de connaître : *episteme* — le savoir de ce qui est nécessairement ainsi —, *techne* — l'art de produire une œuvre selon une règle — et *phronesis* — la sagesse pratique, faire ce qui convient dans le cas concret. Seule la *phronesis* connaît le cas particulier, le contexte, la personne qui répond. Seule elle est indépassablement liée à un humain qui agit et peut assumer ce qui découle de son action.

C'est plus qu'une note philologique. C'est la distinction fondamentale à laquelle toute la construction est suspendue. Car l'intelligence artificielle est au mieux *episteme* (elle sait ce qui, dans bien des cas, est ainsi) et *techne* (elle sait ce que beaucoup font souvent). Qu'un système artificiel puisse un jour juger *phronétiquement* est une question ouverte — peut-être même à laquelle on ne pourra jamais répondre définitivement. Cette note ne tranche pas cette question. Elle en tranche une autre, indépendante : dans une institution humaine, un humain doit assumer ce qu'elle fait. Le droit et le devoir de juger, dans une matière qui a des conséquences pour autrui, sont liés à des personnes qui peuvent être appelées à rendre compte — non à des systèmes dont la capacité de rendre compte demeure non clarifiée. Cette question d'imputation demeure même si la question ontologique venait un jour à être tranchée autrement.

Phronesis : la sagesse pratique demeure humaine.

Le manifeste formule cette posture en dix convictions. Elles ne sont pas traitées ici comme un catalogue, mais résumées en un bref discours — de la manière dont elles vont ensemble.

Nous pensons que la plus grande faiblesse de l'action humaine n'est pas la malveillance, mais l'inconscience. Là où quelque chose se passe sans que personne ne le remarque, l'action perd son lieu moral. Nous pensons que la transparence importe plus que la perfection ; qui promet la perfection doit cacher les erreurs, qui promet la transparence doit les montrer. Nous pensons que l'IA peut conseiller, mais que la responsabilité doit continuer d'incomber à l'être humain — socialement, politiquement, moralement. Nous pensons que l'incertitude est un champ, non une formulation : elle veut être nommée, non voilée par la rhétorique.

Nous pensons que la qualité du jugement est plus décisive pour l'avenir que la seule intelligence technique — car même des recommandations justes deviennent fausses lorsqu'il n'y a plus personne pour les examiner. Nous pensons que la culture et la construction du sens sont profondément humaines : elles sont le sol d'où le jugement peut naître. Et nous pensons, enfin, qu'il faut distinguer strictement l'observé du supposé.

L'observé et le supposé sont deux choses distinctes.

De cette posture découlent les quatre outils. Ils en sont l'exposition ouverte — ouverte, parce qu'ils ne sont pas des logiciels fermés, mais des rôles de pensée qui prennent des formes différentes dans des applications différentes. Et ils découlent de la posture, non l'inverse : c'est seulement une fois établi à quoi un outil doit servir qu'on peut utilement demander comment le construire. Un outil qui ne connaît pas sa posture égare ses utilisateurs.

C'est pourquoi ce chapitre vient en premier. Tout ce qui est décrit dans les chapitres suivants — les quatre outils, les instanciations concrètes, l'établi, les onze normes, la chaîne, l'infrastructure — est conséquence de ce qui est dit ici. Rien de tout cela ne tient pour soi ; tout va ensemble ; et le trait d'union entre toutes les parties est la seule posture qui ne peut être déléguée.

Les quatre outils

Niveau 2. L'exposition ouverte de la posture — non des logiciels, mais des rôles de pensée.

Les quatre outils sont l'exposition ouverte de la posture. Ils ne sont pas eux-mêmes des logiciels, mais des rôles de pensée : Aletheia (percevoir), Mnemosyne (se souvenir), Methodos (guider méthodiquement), Kairon (décider). Chaque application concrète — radar, couche de pseudonymisation, graphe de connaissances, assistant de projet, tableau de bord de décision — est une instanciation de l'un de ces quatre, souvent avec des traits de plusieurs. Les quatre outils lui donnent sa manière de penser ; l'instanciation lui donne sa forme.

OUTIL	OPÉRATION FONDAMENTALE	POINT DE BASCULE
<i>Aletheia · Or</i>	Perception du changement	là où rendre visible devient juger
<i>Mnemosyne · Bleu</i>	Mémoire ; séparer énoncé et personne	là où se souvenir devient livrer
<i>Methodos · Vert</i>	Du dialogue au résultat structuré	là où la forme règne sur l'humain
<i>Kairon · Terre cuite</i>	Décision : Kairos et Aion ensemble	là où un résultat se tient sans raison

Aletheia · Or

Aletheia — perception du changement

Aletheia est la philosophie de la Phronesis pour la perception. Le mot grec signifie le dévoilement, la sortie hors de la dissimulation. Aletheia rend les évolutions visibles, les classe et propose une évaluation — mais soumet le jugement à l'humain pour confirmation.

Rendre visible n'est pas encore juger. Amener quelque chose à la lumière n'est pas déjà l'avoir évalué.

Six convictions portent Aletheia : premièrement, la stricte séparation du rendre visible et du juger ; deuxièmement, le devoir de source (ce qui reste sans source est rumeur) ; troisièmement, la distinction selon le degré de certitude (événement confirmé, signal précoce, simple annonce) ; quatrièmement, l'attention comme sélection, non exhaustivité ; cinquièmement, l'épreuve du nouveau au passé ; sixièmement, l'orientation du dévoilement vers un dedans — vers la mémoire, la mise en forme et la décision. Une perception qui se prend pour un jugement n'est qu'un flux.

Mnemosyne · Bleu

Mnemosyne — mémoire qui protège

Mnemosyne — dans le souvenir grec la mère des Muses — représente une mémoire qui préserve sans livrer. Son opération fondamentale est la séparation de l'énoncé et de la personne : le contenu substantiel reste utilisable, l'identité reste protégée.

Un énoncé et la personne qu'il concerne sont deux choses distinctes.

Quatre convictions : un savoir peut être partagé sans exposer les personnes. Une couche de protection doit elle-même être visible — nul ne fait confiance à une boîte noire qui promet la confiance. Dans le doute, la protection prime : un nom oublié pèse plus lourd qu'un

remplacement de trop. Et enfin : se souvenir et protéger vont ensemble ; une mémoire qui ne protège pas n'est pas une mémoire, mais une livraison.

Methodos · Vert

Methodos — du dialogue au résultat structuré

Methodos conduit du dialogue au résultat, sans oublier le chemin. Il questionne, écoute, ordonne — et laisse advenir une version structurée dans laquelle ce qui a été dit reste reconnaissable.

Le savoir naît dans le dialogue, non dans le formulaire vide.

Une méthode est le jugement décanté de plusieurs.

La structure est un moyen, non une fin.

Quatre convictions : le savoir naît dans le dialogue, non dans des champs vides. Une méthode éprouvée n'est pas un corset bureaucratique, mais le savoir pratique rassemblé de ceux qui ont eu à traiter la même tâche avant nous. Penser avec, c'est remarquer ce qui manque — la complétude s'obtient par de bonnes questions, non par des champs obligatoires. Et la forme doit servir l'humain, non l'inverse.

Kairon · Terre cuite

Kairon — décision qui reste observée

Kairon porte deux mots grecs : Kairos, le juste instant, et Aion, la durée. Tous deux appartiennent à la décision : elle tombe en un instant, mais elle agit dans le temps. Sans Kairos, pas de décision ; sans Aion, pas d'apprentissage.

Les décisions ont besoin non seulement de résultats, mais de raisons.

Toute bonne décision doit demeurer observable.

Le dissensus peut être un signe de maturité.

Trois convictions : les décisions ont besoin de raisons, non seulement de résultats — un oui ou un non sans contexte d'options, d'hypothèses, de risques, de responsabilités et d'effets attendus n'est pas une décision, mais une note. Les bonnes décisions doivent demeurer observables, car ce n'est que dans la réalité qu'apparaît si les hypothèses tenaient. Et le dissensus n'est pas une perturbation, mais souvent le signe qu'une chose a été comprise dans sa profondeur.

Les instanciations

Niveau 3. Les rôles de pensée deviennent des produits qui se tiennent dans le monde.

À ce niveau, les quatre outils deviennent des produits concrets. Chaque instanciation est l'instanciation d'un outil — souvent avec des traits de plusieurs, car aucun produit réel ne se laisse réduire proprement à un seul rôle de pensée. Un entretien de projet est majoritairement Methodos, mais il prend des décisions ; un graphe de connaissances est majoritairement Methodos, mais il se souvient ; une couche de pseudonymisation est majoritairement Mnemosyne, mais elle perçoit aussi ce qui est à protéger.

INSTANCIATION	OUTIL	MATURITÉ	BRÈVE CARACTÉRISATION
<i>ALETHEIA Radar</i>	Aletheia	en construction	radar de technologie IA pour administrations, multi-locataire
<i>Couche de pseudonymisation v0.0.1</i>	Mnemosyne	première version, opérationnelle	local sur 127.0.0.1, entre application et fournisseur, 243 402 noms de famille + 66 916 prénoms (OFS)
<i>kennora</i>	Methodos	en construction	auditeur pour le savoir issu de dialogues, graphe de connaissances en arbre et en réseau
<i>HERMES PLA v0.35.0</i>	Methodos + Kairon	en service	mandat d'initialisation HERMES-2022 sous forme d'entretien, examen formel et substantiel séparés

<i>KAIRON</i> <i>Dashboard (Feature-</i> <i>V2, commit</i> <i>369db5a)</i>	Kairon	maturité fonctionnelle, en test	cycle de vie à dix états, 75 tests de régression, 4 507 lignes de code applicatif
<i>DigiTwin</i> <i>(simulateur de</i> <i>processus)</i>	apparenté · pas dans la famille	prix Hackathon eJustice.ch	figure de pensée apparentée, non centrale à la famille d'outils

ALETHEIA Radar — le regard vérifié vers l'extérieur

L'ALETHEIA Radar est l'instanciation conséquente d'Aletheia pour les administrations. Il rassemble et ordonne les observations sur les évolutions de l'IA, distingue entre événements confirmés, signaux précoces et simples annonces, et joint à chaque entrée sa source. Il est multi-locataire : chaque administration voit son propre radar, avec sa propre sélection d'attention, sa propre histoire, ses propres tampons de vérification. Il ne juge pas, mais propose un classement ; le jugement reste chez l'humain. La maturité est en construction — la structure tient, la pratique d'observation croît.

Couche de pseudonymisation v0.0.1 — l'étape silencieuse en amont

La couche de pseudonymisation est un petit service placé entre l'application et le fournisseur de modèle de langage. Elle remplace les noms par des pseudonymes stables avant que les contenus ne quittent la machine — et défait la substitution après la réponse. Le fournisseur voit l'énoncé substantiel ; les personnes restent chez elles. Les listes de noms proviennent de l'Office fédéral de la statistique (OFS) : 243 402 noms de famille et 66 916 prénoms. La couche tourne localement sur 127.0.0.1 et est ainsi vérifiable pour qui l'utilise : ce n'est pas une boîte noire, mais un service visible. Maturité : première version, opérationnelle.

kennora — l'auditeur pour le savoir

kennora écoute un dialogue et pose un graphe de connaissances. Le graphe a deux vues : un arbre, qui montre les hiérarchies et les appartenances, et un réseau, qui accueille les liens de

toute nature. Outre les arêtes habituelles (est un, appartient à, découle de), il y a la connexion particulière rime-avec. Elle marque des voisinages sémantiques qui naissent d'analogies et d'images — non de similarité formelle, mais du son entendu d'un rapport. Maturité : en construction.

HERMES PIA v0.35.0 — l'entretien avec examen

HERMES PIA accompagne la phase d'initialisation d'un projet HERMES-2022.

L'utilisateur raconte en langage courant ce qu'il projette ; le logiciel en forme un mandat d'initialisation. Deux examens s'exécutent alors séparément : un examen formel en code, sans modèle de langage — il compte si toutes les mentions obligatoires sont présentes, si des espaces réservés sont restés, si des chapitres se contredisent. Et un examen substantiel avec modèle de langage — il lit le mandat du point de vue du mandant et propose une recommandation d'approbation assortie de conditions. L'utilisateur se connecte avec son adresse e-mail professionnelle ; chaque unité organisationnelle ne voit que ses propres projets. Chiffres : 104 tests, 2 054 lignes de code substantiel, 42 fonctions publiques. Maturité : en service.

KAIRON Dashboard — le cycle de vie à dix états

Le KAIRON Dashboard est l'instanciation la plus conséquente de l'axe Kairos-Aion. Chaque décision traverse un cycle de vie à dix états qui ne laisse pas la décision se terminer avec l'approbation, mais la fait passer à l'observation : Draft → In Review → Simulated → Risk Reviewed → Governance Reviewed → Approved → Observed → Reassessment Needed → Reassessing → Archived. L'état Observed est celui qui compte : ici la décision entre dans le temps et est éprouvée pour savoir si ses hypothèses tiennent. Reassessment Needed et Reassessing ferment le cercle. Chiffres : 75 tests de régression, 4 507 lignes de code applicatif. Techniquement : Flask, SQLAlchemy, Jinja, migrations Alembic. Branche Feature-V2, commit 369db5a. Maturité : maturité fonctionnelle, en test.

DigiTwin — le voisin apparenté

DigiTwin est un simulateur de processus qui a été distingué au Hackathon eJustice.ch. Il n'appartient pas à la famille centrale des outils Phronesis, car ce n'est pas un outil au sens

des quatre rôles de pensée. Mais c'est une pensée apparentée — surtout à l'égard de l'axe Kairos-Aion : la simulation éprouve les hypothèses avant qu'une décision soit prise, et flanque ainsi le mode Kairon depuis l'autre côté. Là où Kairon observe l'effet après coup, DigiTwin l'éprouve à l'avance.

L'atelier — l'établi

Traverse. Non un outil de la famille, mais ce qui produit des outils.

L'établi a un rôle particulier dans l'ensemble. Il n'est lui-même pas un outil de la famille Phronesis — il n'est ni Aletheia ni Mnemosyne, ni Methodos ni Kairon. Il est au contraire l'atelier dans lequel les outils naissent. Et il ne les produit pas seulement, il les fonde de manière vérifiable. C'est là sa véritable prestation : il ne livre pas du code seul, mais le dossier qui montre pourquoi on peut faire confiance à ce code.

Quatre stades, sept produits

Un outil naît en quatre stades : image des exigences, architecture, spécification, mise en œuvre. Chaque stade a ses producteurs, ses examinateurs et ses portes. À la fin, sept produits sont là : image des exigences, architecture, spécification, cas de test, spécification technique, mise en service, et — comme véritable cœur de l'affaire — le protocole d'examen et de décision.

Le protocole d'examen et de décision n'est pas un ornement. C'est le produit qui tient ensemble les autres. Il montre qui a examiné quel énoncé, quels modèles étaient impliqués, où ils étaient d'accord, où ils se contredisaient, quels constats sont restés ouverts et ce qui, malgré l'approbation, n'est pas résolu. C'est le document par lequel l'établi rend justice à sa propre norme.

Producteurs et examinateurs, avec des modèles différents

À chaque stade travaille un couple producteur-examineur. Les deux sont des instances différentes ; plus important encore : ce sont des modèles différents. Dans l'exemple VOR-0580b0fa — le tuteur de mécanique quantique sur lequel l'établi est documenté dans sa version 1.0 — le producteur était xai/grok-4.3 et l'examineur openai/gpt-4.1. Deux philosophies d'entraînement différentes, deux régimes RLHF différents, deux profils d'hallucination différents. Des modèles différents ne garantissent pas des erreurs

indépendantes ; mais ils évitent au moins l'auto-examen trivial de la même instance de modèle et rendent visible la dépendance au modèle.

Le noyau calcule, le modèle interprète, l'humain juge et assume

Le noyau calcule, le modèle interprète, l'humain juge et assume.

C'est la décision architectonique fondamentale de l'établi, et c'est le jumeau artisanal de RA-01. Là où le déterminisme est exigé — valeurs de retour pytest, sonde de démarrage, si le service répond, si les cas de test passent, si les identifiants sont complets —, aucun modèle de langage ne doit décider. Là où l'interprétation est exigée — si un chapitre correspond substantiellement aux exigences, si l'architecture remplit son but —, aucun code de règles ne doit juger. Le noyau compte et mesure ; le modèle lit et évalue ; et là où il doit en résulter un jugement assumé, l'humain s'ajoute. Tous trois demeurent dans leur rôle.

La convergence n'est pas la correction

La convergence n'est pas la correction.

La seconde phrase porteuse de l'établi. Lorsqu'un examinateur se tait, cela peut signifier deux choses : le producteur l'a fait juste — ou l'examineur a abandonné. L'établi ne confond pas les deux. C'est pourquoi il consigne quand un examen se termine par accord et quand il se termine par épuisement. La seconde possibilité est la plus fréquente et la plus dangereuse ; elle n'est pas masquée.

*Une version d'approbation n'est jamais sans constat. Ce qui se tient ici est
décidé — non résolu.*

Et c'est pourquoi une approbation de l'établi ne peut jamais prétendre être sans erreur. Elle prétend seulement que les constats encore présents sont décidés : vus, nommés, situés, munis d'un traitement. Ce qui pouvait l'être a été résolu ; ce qui reste a été su.

Mode-noyau et mode-humain

L'établi connaît deux modes de décision. Le mode-noyau s'applique exactement lorsque examinateur et producteur concordent et que le constat est : « s'applique ». Tous les autres cas — chaque contradiction, chaque « ne s'applique pas », chaque « condition » — tombent dans le mode-humain. La machine ne peut que confirmer. Tout ce qui nie, restreint et conditionne demeure réservé à l'humain.

Un agent ne vaut que ce que vaut son critère d'acceptation.

C'est la mise en œuvre la plus conséquente de la condition qui s'appelle dans le research paper consequentiality : la position humaine détermine le résultat, à moins qu'un acte documenté ne la remplace. Dans l'établi, cette condition n'est pas philosophique mais opérationnelle : c'est la différence entre une ligne verte et une ligne rouge dans le protocole.

Ce qui s'ensuit

L'établi est ainsi la réponse opérationnelle à la question fondamentale de la posture : comment la sagesse pratique peut-elle demeurer chez un humain lorsque des machines écrivent avec, examinent avec, proposent avec ? La réponse est : par cela que la machine connaît son rôle, que deux machines différentes s'examinent l'une l'autre, que le noyau compte au lieu de juger, que toute négation revient à un humain.

Les onze normes de la suite

Niveau 4. Le corpus de règles commun au-dessus de toutes les instanciations.

Les onze normes forment le corpus de règles commun de la suite. Elles donnent à l'établi sa mesure, aux outils leur forme et aux instanciations leur vérifiabilité. Elles sont la raison pour laquelle deux outils qui diffèrent en substance — HERMES PIA et KAIRON — restent néanmoins reconnaissables comme parties de la même famille : leurs identifiants, leurs états, leurs régimes d'examen, leur langue suivent les mêmes normes.

N°	NORME	BUT
1	Manifestes	posture, but, limites
2	Identité visuelle	brand & logo styleguide, famille de logos
3	Utilisation	guide UX, guidage utilisateur
4	Architecture de référence v0.4	13 principes RA-01 à RA-13, 10 interdictions NR-01 à NR-10
5	Catalogue de capabilities	appel IA, persistance, workflow, STT, TTS, renderer, export
6	Forme contractuelle	skill-manifest.yaml
7	Norme de skill	SK-01 à SK-17 (norme de skill v0.2 ; la version précédente v0.1 s'arrêtait à SK-15), avec règle + motif + examen

8	Concept de test	examens déterministes et assistés par modèle
9	Directives de développement	CLAUDE.md par dépôt
10	Modèle de rôles	spécification HERMES : rôle métier ↔ méthode ↔ examineur ↔ gouvernance
11	Qualité substantielle	modèle de qualité avec critères D et F

Les treize principes de l'architecture de référence v0.4

L'architecture de référence est le noyau de la suite de normes. Ses treize principes décrivent comment un outil de la famille doit être construit pour pouvoir porter la posture. Ce ne sont pas des recommandations, mais des engagements : qui construit selon eux peut faire vérifier qu'il a construit selon eux.

RA-01 Le noyau calcule, le modèle interprète, l'humain juge et assume.

RA-02 Pas de faits autoritatifs issus du modèle.

RA-03 Chaque objet-noyau porte son gouvernance-mixin.

RA-04 Brouillon et confirmation sont des champs séparés.

RA-05 L'approbation est une porte ; un changement l'invalidé.

RA-06 Les changements sont des transitions, non des ruptures.

RA-07 Configuration avant programmation.

RA-08 La séparation des locataires est structure, non filtre.

RA-09 Pas de contact direct avec le modèle.

RA-10 Chaque résultat porte sa provenance.

RA-11 L'incertitude est un champ, non une formulation.

RA-12 Vocabulaire contrôlé.

RA-13 Un outil produit des preuves, non seulement des résultats.

Remarque : la version de RA-01 donnée ici est précisée par rapport à l'architecture de référence v0.4 (où elle indique : « Le noyau décide, le modèle propose, l'humain assume. »). La nouvelle version suit le triptyque du chapitre V et sera reprise avec la prochaine version de l'architecture de référence. Jusqu'à là, RA-01 dans cette version est une proposition d'ajustement, non la norme actuelle.

Les treize principes ne sont pas tous ancrés de la même manière. Dix reposent directement sur une phrase du manifeste. Deux — RA-07 « Configuration avant programmation » et RA-12 « Vocabulaire contrôlé » — sont des règles techniques sans ancrage direct dans des valeurs, mais légitimées par une pratique produite éprouvée. Un — RA-08 « La séparation des locataires est structure, non filtre » — est ancré indirectement ; sa phrase porteuse PH-11 sur le traitement de ce qui est confié est différée jusqu'à ce que l'opérationnalisation de « expressément destiné au partage » soit clarifiée. Cette différenciation n'est pas dissimulée mais déclarée : ne tient que ce qui est fondé ; ne peut être examiné que ce qui est déclaré.

Les dix interdictions

NR-01 Un outil ne doit pas : porter un jugement ou assumer une responsabilité qui revient à un humain.

NR-02 Un outil ne doit pas produire une référence, un chiffre ou une preuve qui ne provienne pas d'une source.

NR-03 Un outil ne doit pas s'adresser directement à un modèle.

NR-04 Un outil ne doit pas franchir une séparation de locataires.

NR-05 Un outil ne doit pas conserver un état approuvé pendant que le contenu change.

NR-06 Un outil ne doit pas étendre de sa propre autorité un vocabulaire contrôlé.

NR-07 Un outil ne doit pas émettre un résultat sans provenance.

NR-08 Un outil ne doit pas garantir une assurance uniquement par instruction à un modèle.

NR-09 Un outil ne doit pas modifier la signification de valeurs enregistrées sans consigner la transition.

NR-10 Un outil ne doit pas modifier ultérieurement une preuve délivrée.

Interdictions de l'établi (WB-01 à WB-10)

Ces interdictions sont dérivées de la pratique de l'établi et complètent l'architecture de référence ; elles portent leur propre préfixe afin d'exclure toute confusion avec les identifiants NR de l'architecture de référence.

WB-01 Pas de silent narrowing.

WB-02 Pas de routines de confirmation sans réel choix.

WB-03 Pas de choix de modèle sans protocole.

WB-04 Pas de confusion entre convergence et correction.

WB-05 Pas de jugements de modèle de langage dans des comptages déterministes.

WB-06 Pas de mélange producteur et examinateur dans la même instance de modèle.

WB-07 Pas d'approbation sans constats résiduels ouvertement déclarés.

WB-08 Pas d'énoncé sans identifiant de provenance dans les productions de l'établi.

WB-09 Pas de renommage de constats décidés en constats résolus.

WB-10 Pas de livraison d'identités sans vérifiabilité de la pseudonymisation.

Pourquoi ces onze normes ensemble

Les onze normes ne couvrent pas par hasard différents niveaux. De la posture (manifestes) à l'identité visuelle et à l'utilisation, en passant par l'architecture avec ses principes et interdictions, les capabilities, la forme contractuelle et la norme de skill, jusqu'au test, aux directives de développement, au modèle de rôles et au modèle de qualité — elles retracent le chemin qu'un outil parcourt avant de sortir dans le monde.

La chaîne — comment les normes deviennent code

Niveau 5. La chaîne de développement d'agents v0.7 — quatre stades, cinq portes.

La chaîne est la version opératoire de l'établi. Là où l'établi est l'atelier dans lequel une chose est effectivement fabriquée, la chaîne est le procédé qui rend cette fabrication descriptible. Elle a quatre stades et à chaque transition une porte que seul un humain peut ouvrir : image des exigences → architecture → spécification → mise en œuvre → preuve.

Les quatre stades et leurs portes

Au début se tient l'image des exigences. Elle naît dans le dialogue — non dans un formulaire vide, mais dans une narration guidée d'où les exigences sont extraites. À la fin de ce stade se tient la première porte humaine : le mandant lit ce qui est à faire, voit les points encore ouverts et les ferme, ou les confirme comme délibérément ouverts. Ce n'est qu'alors que l'image des exigences est approuvée.

Vient ensuite l'architecture. Elle répond à la question de savoir comment l'outil doit être construit pour honorer l'image des exigences. Ici travaille un couple producteur-examineur sur des modèles différents. À la fin se tient un constat de conformité par rapport à l'architecture de référence : lesquels des treize principes sont entièrement remplis, lesquels partiellement, lesquels restent ouverts. Ce n'est que lorsqu'un humain a examiné cette image et décidé des constats ouverts que l'architecture est approuvée.

Vient ensuite la spécification. Elle transpose l'architecture en une description assez précise pour qu'il puisse en résulter du code — objets-noyaux, interfaces, états, transitions, cas de test. Ici aussi un couple producteur-examineur, ici aussi une porte que seul un humain ouvre.

Enfin la mise en œuvre. Ici naît le code, et avec lui naissent les productions qui rendent le code vérifiable : cas de test, spécification technique, mise en service. À la fin se tient la preuve : le protocole d'examen et de décision qui rend toute la chaîne lisible en sens inverse.

Constat de conformité sur l'exemple HERMES PIA

Sur l'exemple HERMES PIA — examiné par rapport à l'architecture de référence dans sa version v0.2 — deux des principes sont entièrement remplis, neuf partiellement, deux ouverts. Le constat le plus critique était RA-09 (pas de contact direct avec le modèle) : via l'API navigateur Web Speech, il existait un chemin par lequel des données audio pouvaient parvenir à un fournisseur sans pseudonymisation en amont. Il a été nommé, décidé — et fermé dans la version suivante. C'est exactement le cas pour lequel la chaîne est faite : non pour que rien ne se passe mal, mais pour que ce qui pourrait se passer mal soit nommé et corrigé avant que l'approbation ne soit déclarée.

Pourquoi la chaîne a cette forme

La chaîne protège en trois endroits à la fois. Elle protège la posture, parce qu'à chaque porte un humain décide — non un bouton d'approbation, non une routine de confirmation. Elle protège les outils, parce que leurs productions restent vérifiables au regard des normes. Et elle protège les utilisateurs, parce qu'ils ne reprennent pas l'outil dans un état intransparent : ils le reçoivent avec sa propre histoire.

L'infrastructure — Capabilities, Skills, Code

Niveau 6. La fondation technique que tous les outils partagent.

À ce niveau se trouve l'infrastructure : les blocs techniques que tous les outils de la famille partagent. Ils sont la raison pour laquelle un nouvel outil ne recommence pas à zéro — et la raison pour laquelle une erreur dans un bloc partagé peut être corrigée à un seul endroit et tous les outils s'en rétablissent tout de même.

Capabilities — les blocs partagés

Les capabilities sont des blocs utilisés par plusieurs instanciations — et qui sont donc expliqués, construits et examinés exactement une fois. Ils ne sont pas réinventés dans chaque outil. La famille comprend actuellement huit capabilities :

- Appel IA — avec pseudonymisation en amont ; aucun outil n'appelle un fournisseur directement et sans protection (RA-09).
- Persistance — l'enregistrement réglé d'objets-noyaux avec leur gouvernance-mixin (RA-03).
- Workflow — la gestion des transitions entre états ; la machinerie de base derrière les cycles de vie à dix états.
- Reconnaissance vocale — parole vers texte ; guidée à travers la couche de pseudonymisation, là où des noms de personnes peuvent tomber.
- Synthèse vocale — texte vers parole ; pour la sortie vers l'utilisateur.
- Rendu de formules — notation mathématique, p. ex. pour des supports d'enseignement comme le tuteur de mécanique quantique.
- Rendu de diagrammes — diagrammes et schémas ; rendus de manière déterministe, non générés par modèle.

— Export de documents — sortie au format Word et PDF ; export natif, afin que les résultats restent modifiables.

Skills — la forme contractuelle

Les skills sont les plus petites unités fonctionnelles. Elles suivent une forme contractuelle, pour qu'un skill non seulement fonctionne, mais décrive sa fonction : un fichier frontmatter SKILL.md avec les champs name, description, scope, applies_to — et un skill-manifest.yaml comme véritable forme contractuelle. L'ensemble rend un skill vérifiable, échangeable et réutilisable.

Code — les décisions techniques

Le code des outils suit des décisions délibérément sobres. Pour KAIRON : Flask comme cadre web, SQLAlchemy comme ORM, Jinja comme moteur de gabarits, Alembic pour les migrations. C'est la combinaison classique de l'écosystème Python — non la plus récente, mais la plus éprouvée. Elle est choisie parce que ses comportements sont connus : une équipe qui la connaît sait où ce qui se passe mal se passe mal.

Pour les examens, pytest sert d'instance d'examen déterministe. Les résultats sont verts ou rouges ; ils ne sont pas discutables. C'est le mode-noyau que l'établi exige : le noyau calcule.

Pour l'intégration des skills, MCP, le Model Context Protocol, sert d'interface. Il découple les outils de fournisseurs de modèles particuliers et fait ainsi du changement de modèle une question de configuration, non de rupture. C'est RA-07 (configuration avant programmation) dans sa version technique.

Pris ensemble, l'infrastructure ne donne rien de spectaculaire — et c'est voulu. Elle est construite pour se tenir en retrait derrière les outils. Qui porte la posture et connaît les normes n'a plus à être surpris à ce niveau.

Le flanc — Science et présence publique

Ce à quoi l'œuvre se mesure vers l'extérieur et où elle devient visible.

À côté des six niveaux se trouve le flanc : ce qui examine l'œuvre vers l'extérieur et où elle se montre. Il se compose de deux publications scientifiques et d'un dépôt silencieux.

Research paper — Where Judgement Is Reserved

Le research paper au titre de travail *Where Judgement Is Reserved* est le noyau scientifique du flanc. Il suit la méthodologie du Design Science Research selon Peffers et introduit trois termes qui rendent la posture scientifiquement vérifiable : structural reservation — la réservation structurelle d'occasions de jugement comme propriété du système, non comme trait d'interaction ; judgement occasion rate (JOR) — une série temporelle qui mesure quelle part des occasions requérant un jugement revient effectivement à un humain ; et silent narrowing test — un critère d'examen qui demande si l'ensemble des cas réservés peut être silencieusement réduit sans que cette réduction ne soit elle-même une décision documentée et motivée.

L'article positionne la posture par rapport à Green (2022), s'appuie empiriquement sur Bainbridge (1983), Shen & Tamkin (2026), Parasuraman & Manzey (2010) et Buçinca et al. (2021), et trouve dans la FAA-SAFO-17007 (2017) le précédent historique d'une structural reservation imposée par une autorité de surveillance.

Note de position — les douze pages pour tous

La note de position est la lecture unifiante des cinq manifestes — Phronesis, Aletheia, Mnemosyne, Methodos, Kairon — en un unique texte de douze pages. Elle s'adresse à des interlocuteurs qui ne connaissent aucun des textes d'origine. Là où cette note de synthèse est une carte, la note de position est un panorama : le même sol, une autre perspective.

kaspair.ch — le dépôt silencieux

kaspair.ch est la présence de l'auteur. Elle est délibérément silencieuse : pas de marketing, pas de newsletters, pas d'analytique au premier plan. C'est un lieu où les textes reposent — pour tous ceux qui souhaitent les lire, sans qu'une relation ne soit engagée. Qui le trouve ne trouve que l'œuvre.

Le flanc n'est donc ni publicité ni vente. C'est le lieu où l'œuvre, dans sa figure actuelle, doit être examinée — par peer review, par le public, par des lectrices et des lecteurs qui la lisent dans leur propre quotidien.

L'état actuel

Ce qui est fini et ce qui est ouvert. Un bilan honnête.

L'état au 14 août 2026 se lit en deux blocs : ce qui est fini et ce qui est ouvert. L'honnêteté de ce bilan appartient à la posture elle-même — ce que je tais, je ne peux pas l'améliorer.

Ce qui est fini

- Architecture de référence v0.4. Les treize principes et les dix interdictions sont consolidés.
- Note de position. Douze pages, prêtes à publication ; réunit les cinq manifestes.
- Research paper v0.6. Version de travail avec arc narratif complet.
- HERMES PIA v0.35.0. En service ; 104 tests, 2 054 lignes de code substantiel, 42 fonctions publiques.
- KAIRON Feature-V2. Avec 75 tests de régression et 4 507 lignes de code applicatif ; cycle de vie à dix états fermé.
- Établi version 1.0. Quatre stades, sept productions, couples producteur-examineur sur des modèles différents — documenté sur l'exemple VOR-0580b0fa.

Ce qui est ouvert

- Q4 de l'agenda de recherche. La question de savoir si des occasions de jugement réservées préservent une capacité ou ne font que générer du travail est déclarée ouverte dans l'article. Elle est l'axe empirique proprement dit d'un projet successeur.
- Deux limites de la Contribution Boundary v0.5. Premièrement, le selection risk : les organisations préfèrent peut-être les conceptions moins efficaces parce qu'elles se ressentent mieux — Buçinca montre la corrélation négative entre efficacité et préférence. Deuxièmement, les intervention-generated inequalities : les occasions de

jugement réservées n'aident pas également tous les groupes. Les deux limites doivent être nommées dans l'article, non dissimulées.

- Extension de HERMES PIA du mono-locataire au multi-locataire. La structure est actuellement préparée, mais pas techniquement achevée (RA-08 en tant que « structure, non filtre » exige plus qu'une ligne).
- Extension de la famille d'outils. HERMES PIA ne couvre que des parties de l'initialisation. Les phases conception, réalisation, introduction, clôture et mise en œuvre se tiennent comme candidats ouverts à d'autres instanciations.

Les deux blocs montrent où en est l'œuvre : avec une fondation qui porte, et avec des questions plus grandes que la prochaine version. Rien de tout cela ne doit être vite écarté. Ce qui est ouvert reste ouvert — et sera répondu exactement à l'endroit où cela naît.

Conclusion — la figure unique

Retour au commencement : le cercle ouvert dans l'anneau.

Quatre niveaux de résolution portent cette œuvre, et ils disent tous la même chose.

PHILOSOPHIQUEMENT

La sagesse pratique demeure humaine.

ARCHITECTONIQUEMENT

Le noyau calcule, le modèle interprète, l'humain juge et assume.

OPÉRATIONNELLEMENT

*Tout ce qui signifie négation, restriction ou condition tombe dans le mode-
humain.*

MÉTROLOGIQUEMENT

Structural reservation, judgement occasion rate, silent narrowing test.

Ce ne sont pas quatre idées. Ce sont quatre niveaux de la même.

Au commencement se tenait la question de la figure unique qui tient tout ensemble. À la fin se tient la réponse, dans la forme d'un signe simple : le cercle ouvert dans l'anneau. Le tout reste fermé — les six niveaux s'imbriquent, l'établi les relie, le flanc les examine — ; et au milieu un cercle reste ouvert : l'humain qui juge et assume.

Tout ce qui est décrit dans cette note est exposition de ce signe unique. La posture est l'anneau. Les quatre outils sont ses quatre expositions comme rôles de pensée. Les instanciations sont leurs formes dans le monde. L'établi est la main qui les rend vérifiables. Les onze normes sont le corpus de règles qui laisse à l'anneau trouver sa forme. La chaîne

est le chemin. L'infrastructure est la terre sur laquelle il court. Et le flanc est le lieu où doit se montrer que le signe vaut quelque chose.

Ce qui peut manquer et ce qui ne le peut pas, ce qui doit s'ajouter et ce qui doit être écarté, se lit sur cette figure. Le cercle ouvert au milieu est la sauvegarde. Il ne doit pas se fermer — ni par commodité, ni par efficacité, ni par la tentation de faire un outil si bon qu'il n'ait plus besoin du jugement. Si ce cercle se ferme, le signe entier est altéré.

Cinq signes portent la même figure de base : l'anneau comme le tout, le cercle ouvert comme l'humain. Le signe Phronesis les montre dans leur forme pure. Les quatre signes-outils ajoutent chacun exactement un geste d'où la signification de l'outil est dérivée.

Phronesis

LA POSTURE

Le tout reste fermé, l'humain reste ouvert — la figure de base d'où les quatre outils sont dérivés.

Aletheia

PERCEPTION

Le noyau se lève comme un soleil au-dessus de l'horizon — ce qui était caché sort au jour.

Mnemosyne

MÉMOIRE

Le noyau repose en couches préservantes — le savoir demeure, sans livrer la personne.

Methodos

MÉTHODE

Le chemin monte par degrés vers le noyau — de ce qui est dit naît ce que la méthode exige.

Kairon

DÉCISION

L'aiguille trouve le signal au bord — le juste moment dans le temps tout entier.

« *La sagesse pratique demeure humaine.* »