Architecture composite signable

RAISE en action, le geste qui referme la série

AUTEUR	Jérôme Vetillard
PUBLICATION	27 mai 2026
MAJ	29 mai 2026
SOURCE	Twingital Institute
RAISE	Article V · Article VI · Refusal taxonomy · Unsigned agent
PLANCHES	PL.01 · PL.02 · PL.03

VALIDITY DOMAIN

Cours doctrinal de clôture. Synthèse architecturale articulée au cadre normatif RAISE et au cadre régulatoire 2026. Ne se substitue pas aux conseils juridiques spécialisés sur l'EU AI Act, le MDR, ou les autres régulations citées.

SOMMAIRE

Architecture composite signable

01 La question n’est pas la performance, c’est la signabilité 4

02 L’unité de conformité est l’architecture, pas l’agent 5

03 Architecture composite, déterminé signable et non déterminé non 6

04 Article V, le port de promotion 7

05 Article VI, l’écart de promotion 10

06 L’unsigned agent, quatre conditions de signabilité 12

07 Refusal taxonomy, six catégories 14

08 RAISE comme grille d’audit 16

09 Cadre régulatoire 2026 17

10 L’insuffisance des frameworks de marché 23

11 La clôture de la série 25

12 Limites assumées de la série 26

RESUME

Cours doctrinal de clôture. La question pertinente n'est pas est-ce que cet agent est performant, c'est est-ce que cet agent est signable. L'unité de conformité n'est pas l'agent, c'est l'architecture composite. Examine les quatre conditions de signabilité,

déploie l'Article V (port de promotion) et l'Article VI (écart de promotion), articule la refusal taxonomy à six catégories, et positionne le tout dans le cadre régulatoire 2026.

01 La question n’est pas la performance, c’est la signabilité

L’écosystème agentique de 2026 reste obsédé par la performance des modèles. Les classements évoluent au mois, les benchmarks prolifèrent, les pages techniques des frameworks rivalisent de chiffres. Cette obsession est compréhensible. Elle est inadéquate.

Dans un environnement régulé, la performance d’un agent est une condition nécessaire mais terriblement insuffisante. Ce que la régulation exige au premier rang n’est pas qu’un acte soit correct, c’est qu’il soit attribuable. Ce que l’auditeur cherche n’est pas la justesse statistique d’une décision, c’est la chaîne de responsabilité qui permet d’imputer cette décision à un signataire identifié, de la rejouer si nécessaire, et de la défendre devant une autorité de contrôle. La performance s’évalue ex post sur des indicateurs agrégés. La signabilité s’évalue ex ante sur l’architecture qui produit l’acte.

> La question pertinente n’est pas est-ce que cet agent est performant, c’est est-ce que cet agent est signable.

Cette substitution de question n’est pas un détail rhétorique. Elle déplace l’unité d’évaluation. Un agent qui produit en moyenne d’excellentes inférences mais dont aucune ne peut être attribuée à un signataire identifié, rejouée à l’identique, ni défendue devant un auditeur, n’est pas un système conforme. Inversement, un agent dont chaque acte est signé, traçable, et défendable peut très bien être en performance moyenne et rester pleinement opérable en environnement régulé, à la condition que ses performances soient mesurées et que son écart à l’attendu fasse l’objet d’une gouvernance explicite.

Le Twingital Institute pose comme principe que la signabilité est la primitive d’évaluation des systèmes agentiques en environnement réglementé. Les performances suivent. L’inverse ne fonctionne pas.

02 L’unité de conformité est l’architecture, pas l’agent

Conséquence directe du déplacement précédent. Si la signabilité prime, alors l’objet à qualifier n’est pas l’agent considéré comme une boîte noire douée d’une certaine compétence statistique. L’objet à qualifier est l’architecture dans laquelle l’agent prend place et qui rend ses sorties signables ou ne les rend pas signables.

Un agent isolé, même excellent, qui s’exécute dans un script sans workflow durable, sans port de promotion, sans audit immuable, sans refus typés, sans observabilité, est un agent non signable par construction. Aucune amélioration du modèle sous-jacent ne peut compenser cette absence d’architecture. Inversement, un agent moyen embarqué dans une architecture composite proprement conçue produit des actes signables que la régulation accepte sans difficulté.

Cette inversion porte une conséquence opérationnelle forte sur la conduite de programme. Les efforts d’ingénierie qui visent uniquement à améliorer les inférences (fine-tuning, sélection de modèle, prompt engineering) sans toucher à l’architecture qui produit l’acte signé ne déplacent pas le système vers la conformité. Ils déplacent ses performances internes sans déplacer son statut juridique. À l’inverse, le travail sur l’architecture (séparation déterministe/non-déterministe, instrumentation, workflows durables, ports de promotion) déplace immédiatement le système vers la conformité, indépendamment des performances internes.

> L’unité de conformité est l’architecture composite, pas l’agent.

Ce principe a également une conséquence sur la gouvernance des fournisseurs de modèles. Aucun fournisseur de LLM ne porte la signabilité de l’acte. Le fournisseur livre un composant. La signabilité est une propriété émergente de l’architecture qui mobilise ce composant. Cette répartition des responsabilités est centrale dans le cadre EU AI Act, où la responsabilité du déployeur de système d’IA à haut risque est explicite et distincte de celle du fournisseur de modèle.

03 Architecture composite, déterminé signable et non déterminé non

La thèse architecturale centrale du Twingital Institute, qui parcourt toute la série, peut se résumer en une asymétrie. Un système agentique en environnement régulé doit comporter deux couches articulées mais étanches dans leurs propriétés. Une couche déterministe qui constitue l'épine dorsale de l'architecture et qui porte les propriétés de signabilité, de rejouabilité, et d'auditabilité. Une couche non déterministe qui produit les inférences, les propositions, les choix de tools, et qui ne porte aucune de ces propriétés directement.

Le bon design ne consiste pas à rendre la couche non déterministe signable, ce qui est impossible par construction. Le bon design consiste à encapsuler chaque interaction avec la couche non déterministe dans une enveloppe (activity envelope) qui transforme une proposition non déterministe en un acte qui, lui, peut être signé. Cette transformation s'opère au moment du franchissement du port de promotion.

Cette asymétrie n'est pas une convention stylistique. Elle est le théorème fondamental de l'architecture agentique en régulé. Tenter de la contourner revient à exiger de la régulation qu'elle accepte des actes non rejouables et non attribuables, ce qu'elle n'a jamais accepté et ce qu'elle n'acceptera pas dans le futur prévisible. Réciproquement, accepter cette asymétrie revient à se donner les moyens techniques d'opérer des systèmes agentiques dans des environnements où la barrière réglementaire est la plus haute, c'est-à-dire les environnements où la valeur économique d'une opération conforme est la plus grande.

> Le workflow signe, l'activity propose.

Cette compression formule a une portée pratique immédiate. Elle pose ce qui doit vivre dans le workflow (la séquence d'étapes auditables, les états sauvegardés, les signatures), et ce qui doit rester dans l'activity (l'appel au LLM, l'invocation des tools, les choix non rejouables). Toute architecture qui mélange ces deux mondes (workflow qui appelle directement un LLM sans encapsulation, activity qui maintient un état durable) produit des actes hybrides dont le statut juridique est ambigu et l'audit problématique. Le théorème fondamental de la signabilité tient parce que ses deux conditions sont strictement disjointes.

04 Article V, le port de promotion

L'Article V du Framework RAISE, publié en mai 2026, formalise le geste qui permet à une proposition non déterministe de devenir un acte signé. Ce geste est ce que l'Institut nomme le port de promotion. Le terme est doctrinal. Un port désigne un point de passage contrôlé entre deux régimes. La promotion désigne l'élévation d'une proposition au statut d'acte signé. Le port de promotion est donc le point d'architecture par lequel transite chaque acte qui sortira du système et qui pourra être audité, défendu, et signé.

L'Article V définit le port de promotion par quatre conditions cumulatives. Les quatre doivent être satisfaites pour qu'une proposition devienne un acte. Aucune ne peut être substituée par une autre.

Ce qui peut être promu

Première condition. Tout ne peut pas être promu. Seules les propositions qui satisfont aux critères d'éligibilité du domaine peuvent être candidates à promotion. Ces critères ne sont pas définis par l'agent ni par le LLM. Ils sont définis par le métier et formalisés dans le contrat d'enveloppe (activity envelope) que le workflow durable garantit. Une proposition qui ne satisfait pas aux critères d'éligibilité ne franchit pas le port. Elle est rejetée par refus de port, catégorie spécifique de la refusal taxonomy traitée plus loin.

L'éligibilité n'est pas une affaire de probabilité ou de score. Elle est une affaire de type. Pour qu'un acte soit signable, sa proposition doit appartenir à une classe d'actes que le métier a reconnue comme signable. Cette classe est explicite, énumérée, et auditée. Dans le cas d'un programme de médecine prédictive territoriale comme PREDICARE, par exemple, les classes d'actes signables sont définies par l'autorité de santé compétente (référentiels HAS, recommandations professionnelles), et toute proposition issue du modèle qui ne tomberait pas dans ces classes serait refusée au port, indépendamment de la qualité interne de la proposition.

Ce qui accompagne, les pièces de promotion

Deuxième condition. Un acte signable ne se signe pas seul. Il est accompagné de pièces qui constituent son dossier de promotion. Ces pièces sont au minimum quatre. La trace du raisonnement qui a produit la proposition, sous forme exploitable par un humain qualifié. Les références aux sources métier qui ont été mobi-

lisées pour produire cette proposition. Les conditions de validité explicites qui doivent être satisfaites pour que l'acte conserve sa validité. La classification du risque associée à l'acte, dans la nomenclature du référentiel applicable.

Sans ces quatre pièces, un acte peut être produit mais il ne peut pas être promu. La distinction est centrale. Beaucoup de systèmes agentiques en production aujourd'hui produisent des actes qui ne sont pas accompagnés de leur dossier de promotion. Ces actes sont émis. Ils ne sont pas signés. Ils ne survivent pas à un audit sérieux.

Qui signe

Troisième condition. La signature d'un acte signable engage un signataire humain identifiable. L'agent ne signe pas. Le LLM ne signe pas. Le système d'orchestration ne signe pas. Le signataire est une personne physique ou morale qui assume la responsabilité de l'acte, qui est habilitée à le faire dans le cadre régulatoire applicable, et dont l'identité est cryptographiquement liée à l'acte au moment de la promotion.

La conformité au règlement eIDAS (règlement européen 910/2014 sur l'identification électronique et les services de confiance) impose pour les actes à enjeu une signature électronique qualifiée. Une signature qualifiée requiert un certificat émis par un prestataire de services de confiance qualifié et un dispositif de création de signature qualifié. Pour les actes à enjeu réduit, une signature électronique avancée peut suffire. Le choix entre les deux n'est pas un choix technique, c'est une décision juridique fondée sur l'analyse de risque.

L'humain qui signe peut être l'opérateur qui valide en temps réel chaque acte (signature synchrone, modèle clinique courant), ou un signataire institutionnel qui a délégué à un pipeline de signature automatisée la signature d'actes qui satisfont à un cahier des charges qu'il a explicitement validé (signature asynchrone, modèle de signature pipeline HSM). Les deux modèles coexistent. Aucun ne dispense de l'identification d'un signataire humain en bout de chaîne.

Ce qui est journalisé

Quatrième condition. L'acte de promotion lui-même est journalisé dans un audit immuable. Ce journal n'est pas un fichier de logs. Il n'est pas une base relationnelle classique. Il est une chaîne d'événements signés (event-sourced audit log) dont l'intégrité est garantie cryptographiquement et dont la rétention satisfait aux exigences de durée de conservation applicables au domaine.

Le journal contient au minimum, pour chaque acte promu, l'horodatage de la promotion, l'identité du signataire, le hash de la proposition initiale issue de l'activity, le hash du dossier de promotion, la signature elle-même, et les références aux versions des modèles, prompts, et politiques en vigueur au moment de la promotion. Ces six éléments forment la base minimale d'une rejouabilité auditable. Toute rejouabilité partielle qui omet l'un de ces éléments laisse une zone d'ombre dans l'audit, et cette zone d'ombre est précisément l'endroit où la conformité s'effondre devant l'auditeur.

05 Article VI, l’écart de promotion

L’Article VI du Framework RAISE introduit une métrique qui n’a pas d’équivalent dans les pratiques industrielles standard de 2026, mais qui devient la mesure de référence pour la dérive d’un système agentique en environnement régulé. Cette métrique s’appelle l’écart de promotion.

L’écart de promotion mesure la distance entre la proposition issue de l’activity et l’acte qui sort du port de promotion. Si la proposition est promue sans modification, l’écart est nul. Si la proposition est modifiée par le signataire avant promotion (correction d’une formulation, ajustement d’une dose, refus partiel d’une recommandation), l’écart est non nul et sa nature est typée. Si la proposition est refusée et qu’aucun acte n’est promu, l’écart est total et c’est un refus.

> La dérive d’un agent se voit d’abord dans la distribution de ses écarts de promotion, avant de se voir dans la qualité de ses outputs.

Cette distribution est le tableau de bord central de la gouvernance d’un système agentique en production. Trois régimes coexistent.

Écart nul

Un écart nul signifie que la proposition de l’agent a été promue telle quelle par le signataire. Sur un système bien conçu et bien calibré, le régime d’écart nul concerne la majorité des cas. Une proportion d’écart nul trop élevée (par exemple supérieure à 95% sur une fenêtre glissante) doit alerter, parce qu’elle suggère que les signataires deviennent des tampons et ne lisent plus les propositions. Le port de promotion devient alors formel sans être effectif, ce qui ruine la signabilité.

Écart contrôlé

Un écart contrôlé signifie que la proposition a été modifiée par le signataire dans un cadre prédéfini. La modification est typée selon une nomenclature métier (correction de dose, ajustement de fenêtre temporelle, substitution de molécule équivalente, etc.) et elle est journalisée comme telle. L’écart contrôlé est le régime normal d’un système agentique mature, dans lequel le signataire exerce son jugement professionnel sans pour autant désavouer le système. La distribution des

types d’écart contrôlé constitue le signal de dérive le plus précoce. Si une catégorie d’écart contrôlé monte sur une période courte, c’est qu’un biais s’installe quelque part dans la chaîne (modèle, prompt, données d’entrée, contexte métier modifié, version du LLM hosted déployée par le fournisseur), et qu’une investigation doit être engagée avant que la qualité des outputs visible aux indicateurs agrégés ne se dégrade.

Refus

Le troisième régime est le refus. Un refus n’est pas une dérive du système, c’est une décision du système. Il appartient à la refusal taxonomy traitée à la section suivante. Sa journalisation a la même importance que la journalisation d’une promotion. Un système qui ne refuse jamais visiblement n’est pas un système qui marche bien. C’est un système qui ne se confronte jamais aux limites de son périmètre, ce qui est statistiquement impossible et donc révèle un défaut de conception du port de promotion.

Le suivi de l’écart de promotion est en cohérence directe avec la doctrine FDA du Predetermined Change Control Plan, traitée à la section 9. Le PCCP impose à l’industriel qui déploie un dispositif médical à base d’IA de définir ex ante les modifications qu’il prévoit d’introduire dans le système après mise en marché. L’écart de promotion mesuré ex post permet de vérifier que les modifications observées en production tombent dans le périmètre prévu par le PCCP. Si elles en sortent, le système doit être requalifié.

06 L’unsigned agent, quatre conditions de signabilité

Le Framework RAISE introduit la notion d’unsigned agent. Le terme est volontairement paradoxal. Il désigne l’agent qui n’est pas signataire de ses propres actes (puisque la signature engage un humain), mais qui est néanmoins conçu de manière à ce que les actes qu’il propose soient signables par un humain en bout de chaîne. Un agent est unsigned au sens RAISE lorsqu’il satisfait à quatre conditions cumulatives.

Identité stable

Première condition. L’agent doit avoir une identité stable, c’est-à-dire un identifiant unique, persistant, et associé à un cahier des charges (modèle, prompts système, politiques applicables, périmètre de tools disponibles) versionné. Sans identité stable, deux actes successifs ne peuvent pas être attribués au même agent et l’audit s’effondre. L’identité stable est l’équivalent agentique de l’identifiant logiciel SaMD (Software as a Medical Device) au sens du MDR. Elle ne se contente pas de nommer l’agent, elle l’inscrit dans un registre versionnable.

Capacité bornée

Deuxième condition. L’agent doit avoir une capacité bornée et explicite. L’ensemble des actions qu’il peut entreprendre, l’ensemble des tools qu’il peut invoquer, l’ensemble des classes d’actes qu’il peut proposer en promotion, doivent être énumérés et limités. Un agent dont les capacités sont implicites ou découvertes en runtime n’est pas signable. La capacité bornée est ce qui distingue l’agent d’usage en régulé de l’assistant généraliste expérimental.

Audit immuable

Troisième condition. L’agent doit produire un audit immuable de toutes ses actions, qu’elles aboutissent à une promotion réussie, à un écart, ou à un refus. Cet audit n’est pas un fichier de logs applicatif. Il est l’événement source à partir duquel toutes les vues du système (mémoire, tableaux de bord, indicateurs) sont projetées. La discipline est inversée par rapport au logging classique. Les événements signés ne sont pas un effet de bord de l’opération. Ils sont la trace primordiale, et le reste est dérivé.

Refus auditable

Quatrième condition. L'agent doit pouvoir refuser visiblement, et ce refus doit être typé selon la refusal taxonomy. Un agent qui produit toujours une proposition, même improbable, n'est pas un agent signable. Un agent qui refuse silencieusement (par échec d'inférence, par timeout non géré, par hallucination de tool inexistant) n'est pas non plus signable. Le refus doit être un acte de premier ordre, typé, journalisé, et accessible à l'audit au même titre que la promotion.

> Un agent qui ne refuse jamais visiblement n'est pas un acteur, c'est un canal.

Ces quatre conditions sont nécessaires et suffisantes. Un système qui en satisfait trois sur quatre n'est pas unsigned au sens RAISE. Il est un système d'orchestration LLM en attente d'architecture.

07 Refusal taxonomy, six catégories

Le refus n'est pas un échec. Le refus est une décision du système, typée selon une taxonomie à six catégories qui couvre l'espace des situations dans lesquelles un acte ne peut pas être proposé en promotion. La taxonomie est doctrinale et doit être implémentée comme une nomenclature stricte dans le code du port de promotion. Chaque refus est journalisé avec sa catégorie, ses motifs précis, et la proposition qui a été refusée (pour rejouabilité).

Refus de capacité

L'agent n'a pas les capacités techniques requises pour traiter la demande. Le modèle ne supporte pas le type d'entrée, le tool nécessaire n'est pas accessible, le contexte est trop volumineux pour la fenêtre disponible. Le refus de capacité est typiquement résolu par une amélioration d'architecture ou de modèle, pas par une réinterprétation de la demande.

Refus de compétence

L'agent a les capacités techniques mais le cahier des charges du domaine ne lui accorde pas la compétence métier pour produire l'acte demandé. Cette distinction est subtile mais elle est centrale. Un agent peut être techniquement capable de produire une recommandation de dose médicamenteuse, mais ne pas être habilité par le cahier des charges du domaine à le faire dans la situation présente. Le refus de compétence est typiquement résolu par un escalation humain.

Refus de domaine

La demande sort du périmètre de domaine de l'agent. Un agent spécialisé en imagerie pulmonaire qui reçoit une question sur la cardiologie doit refuser pour sortie de domaine. Cette catégorie est la plus simple à implémenter et la plus négligée en pratique. Un agent qui répond hors domaine est un agent qui produit des actes que personne ne peut signer, parce que personne dans la chaîne d'autorité n'a délégué à l'agent la compétence sur le domaine demandé.

Refus éthique

La demande ou la proposition issue de l'activity rencontre une contrainte éthique formalisée dans les politiques de l'agent. Le refus éthique est typé selon une sous-taxonomie qui dépend du domaine (refus de produire un contenu discrimi-

natoire, refus d'assister à une situation de chantage ou de manipulation, refus de produire du contenu qui violerait la dignité d'un patient identifié). La sous-taxonomie est doctrinale et doit être documentée explicitement.

Refus de gouvernance

La demande ou la proposition rencontre une politique de gouvernance interne à l'organisation déployante (rétention de données, accès à des ressources, limitations de cas d'usage non autorisés par la direction). Le refus de gouvernance distingue ce que l'agent peut techniquement et éthiquement faire de ce que l'organisation lui permet de faire dans ce contexte précis. Cette distinction est ce qui permet à une même architecture agentique d'être déployée dans plusieurs organisations avec des politiques différentes sans modification du code de l'agent.

Refus de port

La proposition issue de l'activity ne satisfait pas aux critères d'éligibilité du port de promotion. C'est le refus le plus structurel, celui qui réalise l'Article V dans le négatif. Une proposition peut être techniquement valide, compétente, dans le domaine, éthiquement acceptable, et conforme aux politiques de gouvernance, et néanmoins ne pas franchir le port de promotion parce que son dossier de promotion est incomplet, ou parce que les conditions de validité explicites ne peuvent pas être garanties à l'instant de la promotion.

La distribution des refus par catégorie constitue, avec la distribution des écarts de promotion, le second tableau de bord de la gouvernance opérationnelle. Une montée du refus de domaine signale un défaut de filtrage en entrée. Une montée du refus de port signale un désalignement entre le modèle et le cahier des charges métier. Une montée du refus de gouvernance signale un durcissement des politiques internes, qui peut être délibéré ou subi. Chaque type de refus a sa signature opérationnelle, et la gouvernance les surveille distinctement.

08 RAISE comme grille d’audit

L’ensemble des principes ci-dessus, articulés dans le Framework RAISE Volume 1 publié en mai 2026, constitue une grille d’audit applicable à tout système agentique en environnement régulé. Cette grille est utilisable dans trois temps de la vie d’un programme.

Au moment de la conception, RAISE sert de spécification d’architecture. Les choix de conception (séparation déterministe/non déterministe, instrumentation, identification du signataire, refusal taxonomy) sont confrontés aux articles du Framework et leur conformité est qualifiée. Cette utilisation amont évite les refontes coûteuses qui surviendraient si la conformité n’était inspectée qu’à la mise en marché.

Au moment de la qualification réglementaire, RAISE sert de documentation structurée. Les autorités de contrôle (autorités notifiées en MDR, FDA en SaMD, autorités de protection des données en RGPD) accèdent à un dossier dont l’organisation suit les articles du Framework, ce qui réduit le temps d’instruction et améliore la lisibilité du système par des évaluateurs qui peuvent ne pas être eux-mêmes spécialistes de l’architecture agentique.

En opération continue, RAISE sert de référentiel pour la mesure de dérive. Les indicateurs de production (écart de promotion, distribution des refus, taux de promotion par catégorie d’acte) sont définis directement à partir des articles du Framework, et leur évolution dans le temps est rapportée selon le même langage.

Cette utilisation triple, amont, qualification, opération, fait de RAISE un cadre de gouvernance plutôt qu’une simple méthodologie de conception. Le passage du cadre à la documentation et de la documentation aux indicateurs opérationnels est continu, et c’est cette continuité qui rend la grille opératoire dans le quotidien des programmes.

09 Cadre régulatoire 2026

Le périmètre régulatoire applicable aux systèmes agentiques en environnement régulé est devenu en 2026 d'une complexité qui ne permet plus l'amateurisme. Trois cadres font l'essentiel de la charge et méritent un traitement détaillé. Plusieurs autres entrent en complément et sont mentionnés en notes.

EU AI Act

Le règlement (UE) 2024/1689 du 13 juin 2024, publié au Journal officiel de l'Union européenne le 12 juillet 2024 et entré en vigueur le 1er août 2024, est le texte structurant. Son calendrier de mise en application est progressif. Les pratiques prohibées (article 5) sont applicables depuis le 2 février 2025. Les obligations applicables aux modèles d'usage général (GPAI, chapitre V) sont entrées en application le 2 août 2025. Les obligations applicables aux systèmes d'IA à haut risque (chapitre III) deviennent applicables le 2 août 2026. Les obligations applicables aux systèmes d'IA à haut risque intégrés dans des dispositifs faisant l'objet d'une évaluation de conformité au titre d'autres textes (annexe I, qui inclut les dispositifs médicaux relevant du MDR et de l'IVDR) deviennent applicables le 2 août 2027.

L'article 14 sur la supervision humaine est central pour les architectures agentiques. Il impose que les systèmes à haut risque soient conçus de sorte qu'ils puissent être supervisés par des personnes physiques pendant leur utilisation. Cette supervision doit permettre, entre autres, de comprendre correctement les capacités et limites du système, de surveiller son fonctionnement pour identifier toute anomalie, de décider de ne pas utiliser le système ou d'interrompre son fonctionnement, et d'intervenir dans son fonctionnement ou d'arrêter le système. L'architecture composite signable instancie littéralement ces exigences. Le port de promotion est le point d'intervention humaine. Le journal d'audit immuable est le support de la surveillance. La refusal taxonomy est le mécanisme d'interruption typé.

L'article 17 sur le système de management des risques impose un processus continu et itératif de gestion des risques sur l'ensemble du cycle de vie du système. Cette continuité s'aligne directement sur la mesure d'écart de promotion en opération, qui constitue un indicateur de risque émergent surveillable.

L'article 15 sur l'exactitude, la robustesse et la cybersécurité impose un niveau approprié de performance et de résilience aux erreurs, défaillances, et incohérences. Pour une architecture composite signable, cet article a une lecture spéci-

fique. La robustesse n’est pas évaluée sur le LLM isolé, dont les défaillances sont par construction non éliminables, mais sur l’architecture composite dont la couche déterministe absorbe les défaillances de la couche non déterministe via la refusal taxonomy et la signature humaine. Cette lecture déplace la cible de robustesse du modèle vers le système. L’évaluation devient mesurable sur des indicateurs opérationnels (taux de refus de capacité, distribution des écarts contrôlés) plutôt que sur des benchmarks théoriques du modèle.

L’article 17 sur le système de management des risques impose un processus continu et itératif de gestion des risques sur l’ensemble du cycle de vie du système. Cette continuité s’aligne directement sur la mesure d’écart de promotion en opération, qui constitue un indicateur de risque émergent surveillable.

L’article 27 introduit la FRIA (fundamental rights impact assessment), évaluation d’impact sur les droits fondamentaux, applicable aux déployeurs publics ou aux déployeurs qui fournissent des services publics. La FRIA s’articule à la DPIA prévue par l’article 35 du RGPD lorsque les deux sont applicables.

L’article 50 sur la transparence impose à certains systèmes d’informer les utilisateurs qu’ils interagissent avec un système d’IA. Cette obligation, mineure en apparence, a une conséquence d’architecture. Le système doit savoir signaler ex ante sa nature, et cette signalisation est elle-même un acte qui doit être journalisé. L’article impose également l’étiquetage des contenus synthétiques produits par IA générative, ce qui s’étend aux outputs d’un système agentique mobilisant un LLM. La journalisation de l’étiquetage est, comme la promotion d’un acte, un événement signé qui doit être présent dans l’audit immuable.

Les sanctions, prévues à l’article 99, atteignent 35 millions d’euros ou 7% du chiffre d’affaires annuel mondial pour les manquements aux pratiques prohibées, et 15 millions d’euros ou 3% pour les manquements aux obligations applicables aux systèmes à haut risque. L’ordre de grandeur est suffisant pour justifier l’effort d’architecture demandé par RAISE, et il modifie les choix de portage en termes d’arbitrage économique. Le coût d’une architecture composite signable correctement implémentée est mesurable en personnes-mois d’ingénierie. Le coût d’une sanction pour un déploiement non conforme est de plusieurs ordres de grandeur supérieur, sans compter l’effet réputationnel et l’interdiction temporaire de mise en marché qui peuvent accompagner les sanctions les plus lourdes.

MDR et IVDR pour les dispositifs médicaux à base d'IA

Le règlement (UE) 2017/745 (MDR, medical devices regulation) et le règlement (UE) 2017/746 (IVDR, in vitro diagnostic medical devices regulation) couvrent les dispositifs médicaux mis sur le marché européen. Un logiciel qui répond à la définition d'un dispositif médical (logiciel destiné à être utilisé spécifiquement à des fins médicales) tombe dans leur périmètre. Un système agentique conçu pour assister une décision médicale entre dans cette catégorie dès que son output est qualifié de SaMD (Software as a Medical Device).

La classification du SaMD selon l'IMDRF (International Medical Device Regulators Forum) repose sur le croisement entre l'état de santé visé (non grave, sérieux, critique) et la nature du support apporté (information, traitement ou diagnostic informé, traitement ou diagnostic guidé). Cette grille produit quatre catégories de risque, et la majorité des systèmes agentiques d'aide à la décision clinique tombent dans les catégories II ou III qui imposent les exigences les plus élevées en matière d'évaluation clinique et de gestion du cycle de vie. La grille IMDRF est utilisée comme référence par les autorités européennes lors de l'évaluation des dossiers de marquage CE pour les SaMD.

Le MDR pose un principe de gestion du cycle de vie qui est exigeant pour les dispositifs apprenants. L'annexe I exige que les performances cliniques soient maintenues sur toute la durée de vie utile du dispositif. Cela impose une surveillance post-commercialisation (post-market surveillance) qui, pour un système agentique, prend la forme d'une mesure continue de l'écart de promotion et de la distribution des refus. Cette surveillance est documentée dans le PSUR (Periodic Safety Update Report) et alimente le PMCF (Post-Market Clinical Follow-up). L'architecture composite signable apporte la traçabilité nécessaire à ces deux documents sans surcoût d'instrumentation, parce que les indicateurs requis sont natifs de l'architecture.

L'article 117 du MDR, ajouté par le règlement de 2017, concerne les modifications du dispositif après mise en marché. Pour les dispositifs traditionnels, la qualification d'une modification est binaire : la modification est significative et requiert une nouvelle évaluation de conformité, ou elle ne l'est pas. Pour les dispositifs à base d'IA, cette binarité est inopérante, parce que le système évolue par construction lorsqu'il apprend ou lorsque ses dépendances logicielles (LLM hosted, par exemple) sont mises à jour par leur fournisseur. La doctrine européenne sur ce point reste en construction. Le PCCP de la FDA, traité à la section suivante, offre un précédent doctrinal que les autorités européennes observent de près.

L'évaluation clinique (annexe XIV) impose la production de preuves cliniques justifiant la sécurité et la performance du dispositif. Pour un système agentique, cette évaluation porte non seulement sur les performances de l'inférence, mais aussi sur la robustesse de l'architecture composite qui produit l'acte signable. L'évaluation doit démontrer que l'architecture maintient ses propriétés de signabilité dans les conditions réelles d'usage clinique, ce qui inclut la résilience aux entrées dégradées, aux pannes partielles d'infrastructure, et aux dérives lentes de la distribution des cas. C'est précisément ce que la grille RAISE permet de documenter de manière structurée.

FDA PCCP comme précédent doctrinal

La doctrine PCCP (Predetermined Change Control Plan) de la FDA américaine est, à fin 2024, le précédent doctrinal le plus avancé sur la question des dispositifs médicaux à base d'IA apprenants. La FDA a publié en décembre 2024 une guidance finale intitulée Marketing Submission Recommendations for a Predetermined Change Control Plan for Artificial Intelligence-Enabled Device Software Functions, qui succède à un discussion paper de 2019, à un action plan de 2021, et à un draft guidance de 2023.

Le principe central du PCCP est l'inversion temporelle de la qualification des modifications. Pour un dispositif traditionnel, les modifications sont qualifiées ex post lorsqu'elles surviennent. Pour un dispositif à base d'IA muni d'un PCCP, l'industriel définit ex ante, dans le dossier de mise en marché, les modifications qu'il prévoit de mettre en œuvre une fois le dispositif déployé. Ces modifications anticipées (par exemple, mise à jour du modèle avec de nouvelles données d'entraînement, élargissement de la cohorte cible, modification des seuils de décision) font partie intégrante du marquage initial et n'exigent pas de nouvelle évaluation tant qu'elles restent dans le périmètre prédéterminé.

Le PCCP éclaire le port de promotion sous un angle régulatoire. Le port de promotion est le mécanisme architectural qui rend opérante la doctrine PCCP. Sans port de promotion, la frontière entre une modification prévue et une modification non prévue est floue parce qu'aucune mesure quantitative ne sépare les deux régimes. Avec port de promotion et écart de promotion mesuré, la frontière devient mesurable. Tant que la distribution des écarts de promotion reste dans le périmètre prédéterminé, le système opère dans son PCCP. Lorsque la distribution sort de ce périmètre, la modification est non prévue et déclenche le processus de requalification.

L'industriel ToxTwin V3.0, qui développe une plateforme de prédiction de toxicité moléculaire à destination de l'industrie pharmaceutique, applique ce principe sur sa cohorte de holdout figée (SHA256[0:16] = 47d3927ebc18cb7f). Toute évolution du modèle qui dégraderait les performances mesurées sur cette cohorte au-delà d'un seuil prédéterminé est une modification non prévue au sens du PCCP, et déclenche une requalification avant déploiement.

L'instrumentation opérationnelle qui rend ce principe applicable est elle-même une architecture composite. Le modèle producteur de prédictions (couche non déterministe) propose une valeur de toxicité. Le workflow de qualification (couche déterministe) la confronte à la cohorte de holdout figée, mesure l'écart de promotion par rapport à la version de référence enregistrée dans le PCCP, et décide selon une politique explicite si la version proposée peut être promue ou si elle déclenche une investigation. Le signataire de la promotion est un responsable scientifique de l'industriel, identifié et habilité. Le journal de promotion est versionné dans une infrastructure conforme aux exigences réglementaires américaines (21 CFR Part 11) qui imposent l'horodatage, la signature électronique, et la non-répudiation des enregistrements.

L'apprentissage de ce cas pratique est généralisable. Toute architecture composite signable doit définir, dès la conception, une cohorte de référence figée à partir de laquelle l'écart de promotion sera mesuré en continu. Sans cohorte figée, la mesure d'écart n'a pas de référentiel stable et la doctrine PCCP devient inopérante. Avec cohorte figée, la mesure devient quantitative et permet à l'autorité de contrôle de qualifier ex ante les modifications acceptables, ce qui réduit considérablement le coût et la durée des cycles de qualification réglementaire en aval.

Notes complémentaires sur les autres cadres

Le RGPD (règlement (UE) 2016/679) reste structurant pour toute architecture agentique qui traite des données personnelles. L'article 22 (décisions individuelles automatisées produisant des effets juridiques ou affectant significativement la personne) impose dans la plupart des cas une intervention humaine, ce qui est nativement satisfait par l'architecture composite signable. L'article 35 (DPIA, data protection impact assessment) est obligatoire pour les traitements présentant un risque élevé pour les droits et libertés, ce qui inclut la majorité des systèmes agentiques en santé ou en services à la personne. L'article 17 (droit à l'oubli) impose une discipline de mémoire long-terme abordée en cours 3 sous l'angle du crypto-shredding.

La FRIA (article 27 de l'EU AI Act) et la DPIA (article 35 RGPD) s'articulent. La FRIA est plus large et couvre les droits fondamentaux au-delà de la protection des données. La DPIA est plus précise sur les traitements de données personnelles. Pour un système agentique déployé par une autorité publique et traitant des données personnelles, les deux évaluations sont conjointement obligatoires et leur intégration documentaire suit naturellement la grille RAISE.

L'IMDA Singapour (Infocomm Media Development Authority) a publié en 2023 et révisé en 2024 son Model AI Governance Framework, qui ne porte pas force de loi mais constitue un cadre de référence pour les déploiements régionaux. La AI Verify Foundation, soutenue par l'IMDA, met à disposition un toolkit open source d'évaluation des systèmes d'IA dont les principes recoupent l'esprit de RAISE sans en partager la doctrine architecturale.

Le Cloud Security Alliance (CSA) a publié en 2025 un Agentic Trust Framework qui aborde spécifiquement les enjeux d'identité, d'autorisation, et d'auditabilité des agents en environnement cloud. Son angle est complémentaire de RAISE sur les questions d'infrastructure et de sécurité périmétrique. Il ne couvre pas la signabilité doctrinale mais constitue une référence utile pour l'opérationnalisation cloud des architectures composite signables.

10 L’insuffisance des frameworks de marché

Aucun framework agentique disponible sur le marché en mai 2026 ne porte la signabilité au sens RAISE. Cette affirmation peut paraître brutale. Elle résulte d’un audit méthodique des douze frameworks examinés au cours 2 de la présente série.

Les frameworks vendor-first (Claude Agent SDK, OpenAI Agents SDK, Google ADK, Microsoft Agent Framework) excellent dans l’orchestration LLM et l’invocation de tools. Aucun ne propose nativement de port de promotion typé, d’audit immuable cryptographiquement scellé, ni de refusal taxonomy à six catégories. Le développeur qui souhaite atteindre la signabilité doit construire ces composants par-dessus le framework, ce qui est techniquement possible mais ce qui révèle que la signabilité n’est pas une fonctionnalité que le framework cherche à porter. Cette absence est doctrinalement explicable. Les fournisseurs de LLM ont intérêt à ne pas porter la signabilité, parce que la signabilité engage une responsabilité que les conditions générales de leurs services excluent par contrat. Ce qui est rationnel pour le fournisseur impose au déployeur de construire lui-même la signabilité.

Les frameworks indépendants (LangGraph, CrewAI, Pydantic AI, SmolAgents, Mastra) excellent dans l’orchestration multi-agents et la modularité. Pydantic AI se distingue par sa rigueur typologique stricte, qui constitue une fondation favorable pour l’implémentation de la signabilité. CrewAI offre le meilleur ratio simplicité/résultat sur des cas d’usage multi-agents simples. Aucun de ces frameworks ne porte non plus la signabilité comme primitive. Le développeur doit construire son port de promotion sur leur API, et la qualité de l’implémentation dépend de la discipline du développeur, ce qui est précisément le risque que la régulation refuse d’assumer.

Les moteurs durables (Temporal, DBOS, Restate, Inngest, Trigger.dev) apportent la couche déterministe qui rend la signabilité possible. Temporal est la référence du marché en termes de maturité, mais sa courbe d’apprentissage est exigeante et son couplage au payload LLM dans l’history pose des questions opérationnelles abordées au cours 5. DBOS a un avantage doctrinal majeur pour les environnements régulés santé. Sa durabilité repose sur un Postgres qui est déjà audité HDS, ce qui réduit le périmètre de qualification réglementaire de l’infrastructure de durabilité. Sans moteur durable, aucune signabilité sérieuse n’est atteignable. Le moteur durable est nécessaire. Il n’est pas suffisant.

Cette répartition des rôles a une conséquence sur la conduite de programme. Le bon design d'une architecture agentique en environnement régulé ne consiste pas à choisir le framework idéal, parce qu'aucun ne couvre l'ensemble des conditions de signabilité. Il consiste à composer plusieurs briques selon une grille fonctionnelle. Un framework agentique pour l'orchestration LLM et l'invocation de tools. Un moteur durable pour la couche déterministe. Une infrastructure d'événements pour l'audit immuable et la communication inter-composants. Un module de signature pour le port de promotion. Une couche d'observabilité OTel-first pour la mesure de l'écart de promotion. Une refusal taxonomy implémentée comme nomenclature stricte dans le code. Cette composition n'est pas accidentelle, elle est la traduction technique de la grille RAISE.

> La signabilité émerge de l'architecture composite.
> Aucun composant isolé ne la porte.

Le constat ne disqualifie aucun des frameworks examinés. Il établit qu'aucun ne peut être déployé en l'état dans un environnement régulé sans un effort d'architecture explicite. Cet effort est précisément ce que la présente série de cours documente. Le cours 2 cartographie les composants disponibles. Les cours 3 à 7 traitent leurs articulations. Le présent cours 8 articule l'ensemble au cadre normatif RAISE et au cadre régulatoire.

11 La clôture de la série

La série Architectures agentiques 2026 qui se referme avec ce cours a une thèse unique. Un système agentique en environnement régulé doit être conçu comme une architecture composite signable, articulant une couche déterministe qui porte la signabilité et une couche non déterministe qui porte les inférences. Cette architecture est instancée techniquement par un workflow durable qui orchestre des activities encapsulées, un port de promotion qui transforme les propositions en actes signés, une refusal taxonomy qui type les non-actes, un audit immuable qui constitue la trace primordiale, et une observabilité eval-driven qui mesure la dérive sur la métrique de l'écart de promotion.

Cette thèse n'est pas un point de vue de l'Institut parmi d'autres possibles. Elle est la conséquence logique du croisement entre les obligations posées par le cadre régulatoire 2026 et les contraintes techniques posées par la nature non déterministe des LLM. Toute architecture qui s'écarte de cette thèse devra, en environnement régulé, expliquer comment elle satisfait aux quatre conditions de l'unsigned agent sans recourir à l'asymétrie déterministe/non déterministe. À la connaissance de l'Institut, aucune telle architecture alternative n'a été démontrée à ce jour.

> Ce qui n'est pas signé n'a pas eu lieu.

Cette compression formule résume la doctrine en sept mots. Elle est applicable aux programmes en cours de conception comme aux programmes en production. Pour les programmes en conception, elle pose la spécification minimale. Pour les programmes en production, elle pose le critère d'audit que les autorités de contrôle appliqueront, que les équipes y soient préparées ou non.

12 Limites assumées de la série

Aucune doctrine n'est pérenne au-delà de la fenêtre temporelle dans laquelle son périmètre d'application reste stable. La présente série est datée de mai 2026. Quatre limites sont explicitement assumées.

Premièrement, le paysage des frameworks évolue à un rythme qui rend la cartographie du cours 2 obsolète à l'horizon de six à neuf mois. Cette obsolescence est intrinsèque à l'objet cartographié, pas à la méthode. Refaire l'exercice de cartographie tous les deux trimestres est une discipline, pas une option.

Deuxièmement, le cadre régulatoire 2026 lui-même évolue. L'EU AI Act entre dans sa phase d'application progressive jusqu'en 2027. Les guidances de la Commission européenne, les actes délégués, et les normes harmonisées qui en découlent vont préciser les obligations actuellement formulées à un niveau abstrait. La doctrine RAISE devra suivre ces précisions sans s'y enfermer.

Troisièmement, la thèse de l'architecture composite signable repose sur l'asymétrie déterministe/non déterministe. Cette asymétrie est valide dans le paysage technique de 2026. Elle pourrait évoluer si des avancées en interprétabilité des LLM, en raisonnement causal, ou en formal methods rendaient certaines inférences elles-mêmes auditables sans encapsulation. L'Institut surveille ces avancées sans anticiper leur calendrier d'industrialisation.

Quatrièmement, la série ne couvre pas certains périmètres connexes qui méritent leur propre série. La souveraineté infrastructure (SecNumCloud, HDS qualifié, eIDAS hébergement), la formation des opérateurs en bout de chaîne, et les enjeux organisationnels de gouvernance des programmes IA en santé sont identifiés comme points aveugles assumés. Ils feront l'objet de productions ultérieures.

Le validity domain de la série a été énoncé au cours 1 et il s'applique au présent cours sans modification. Cette série porte sur l'architecture des systèmes agentiques en environnement régulé. Elle ne porte pas sur l'éthique substantielle, ni sur la philosophie politique de l'IA, ni sur le droit substantiel des décisions automatisées. Elle constitue une grille d'analyse architecturale et un instantané daté, pas un manuel de référence permanent.

Le Twingital Institute remercie les lecteurs qui ont parcouru cette série jusqu'à sa clôture. Les retours, contradictions argumentées, et expériences de terrain sont accueillis avec intérêt et peuvent être adressés au Twingital Institute. La doctrine n'est pas figée. Elle progresse par confrontation à l'expérience accumulée des praticiens qui l'éprouvent en programme réel.

A propos de cette serie

Ce cours appartient a la serie Architectures agentiques 2026 publiee par le Twingital Institute. La serie aborde l'architecture des systemes agentiques en environnement regule a partir du Framework RAISE.

L'ensemble de la serie, les planches d'architecture interactives, et la matrice de decision frameworks sont accessibles sur le site de l'Institut : twingital-ventures.com/fr/cours/ https://twingital-ventures.com/fr/cours/

TITRE	Architecture composite signable
SOUS-TITRE	RAISE en action, le geste qui referme la série
AUTEUR	Jérôme Vetillard
SOURCE	Twingital Institute
SERIE	Architectures agentiques 2026 · Cours 8 / 8
PUBLICATION	27 mai 2026
MAJ	29 mai 2026
RAISE	Article V · Article VI · Refusal taxonomy · Unsigned agent
PLANCHES	PL.01 · PL.02 · PL.03
MOTS-CLES	composite · signabilité · RAISE · Article-V · Article-VI · refusal · EU-AI-Act
URL	https://twingital-ventures.com/fr/cours/2026-05-cours-8-architecture-composite-signable-fr/