

L'actif n'est plus le modèle

Ce qui a désormais une valeur patrimoniale, ce n'est pas le logiciel. C'est la capacité reproductible de l'opérer.

Le lecteur a changé

Depuis la première semaine de ce cycle, une phrase tient : un modèle performant sur un benchmark n'est pas exploitable pour autant. C'était une phrase d'ingénieur, défendue sur un terrain d'architecture.

Cette année, elle a changé de langue. Non parce que la thèse a évolué, mais parce qu'un nouveau lecteur est apparu au COMEX, et qu'il ne lit pas les scores de précision. Le directeur général, le General Counsel ou le Chief Risk Officer ne demandent plus si le modèle fonctionne en laboratoire. Ils demandent si l'organisation a la capacité soutenable d'en assumer les décisions, et à quel coût récurrent.

La déployabilité, entendue comme simple capacité technique de mise en production, ne suffit plus à décrire le problème. L'ancienne équation liait la performance à la mise en production. La nouvelle relie la performance à la capacité soutenable d'en exploiter les décisions. Le verrou n'est plus technologique. Il est institutionnel.

Trois niveaux, un ordre

Le cycle mobilise trois notions voisines qu'un lecteur pressé traite comme des synonymes, et qu'un contradicteur, lui, saura séparer. Elles ne vivent pas au même niveau, et l'ordre compte.

- Au premier niveau, la *capacité soutenable d'exploitation* : une propriété émergente de l'organisation, son aptitude à absorber, piloter et auditer le risque d'un système automatisé dans la durée, à coût maîtrisé.
- Au deuxième niveau, le *régime institutionnel* : l'ensemble des règles, formelles et informelles au sens où l'économie des institutions l'entend, qui rendent cette capacité possible, des lignes de responsabilité aux procédures d'arbitrage et aux obligations de preuve.
- Au troisième niveau, le *régime d'auditabilité* : l'implémentation technique et organisationnelle de ces règles dans un système donné, les dispositifs par lesquels il reste à tout instant contrôlable et justiciable.

Une propriété, les règles qui la fondent, leur incarnation dans l'architecture. Cet ordre n'est pas décoratif : il indiquera, en fin d'article, lequel de ces trois objets mérite le nom d'actif. Ce lexique est stable pour tout le cycle.

La migration de la contrainte et de la captation

Le marché a longtemps lu le coût de l'intelligence artificielle comme un coût d'acquisition. On achète une licence, on appelle une API, et le reste ne serait qu'une affaire de courbe de capacité : le prochain modèle, mieux entraîné, absorbera ce qui coince. Confondre le coût du modèle et le coût du régime qui l'entoure est l'erreur fondatrice des premiers déploiements.

Le coût d'acquisition est celui du modèle, payé une fois. Le coût d'exploitation est celui du régime d'auditabilité qui l'encadre, payé chaque mois : surveillance continue, ingénierie des données de contrôle, arbitrage humain, gestion des dérives, responsabilité juridique. On objectera que le delivery informatique n'a jamais rien eu de trivial, et c'est exact : intégrer un SAP, un cœur bancaire ou une plateforme hospitalière reste un art difficile. Ce qui s'est commoditisé n'est donc pas le delivery en général, mais un segment précis : l'accès aux modèles, l'intégration de base, le déploiement standard. Le delivery du modèle devient une commodité. Le delivery du régime, non.

De là une loi que les états-majors découvrent en 2026 : le coût d'exploitation ne baisse pas quand le modèle s'améliore. Souvent, il monte. Un modèle jugé plus performant est immédiatement appliqué à des processus plus critiques, capturant des flux métiers à haute responsabilité. Ce faisant, il étend la surface d'impact d'une anomalie et exige un durcissement des structures de contrôle.

Le facteur limitant a migré, et c'est de lui qu'il faut parler, non de la valeur. La valeur reste co-produite par le modèle, les données, les processus et les personnes : elle n'a pas changé de source. Ce qui se déplace, c'est le goulot d'étranglement, et la rente ne suit pas la valeur, elle suit le goulot. Hier, le goulot était la rareté technologique et la capacité brute du modèle, et la rente allait aux laboratoires fondateurs. Aujourd'hui, les capacités génériques se commoditisent, le prix de l'inférence s'effondre, et le goulot devient l'incapacité des organisations à absorber, piloter et auditer le risque opérationnel. La captation de la rente se déplace donc vers la maîtrise du régime institutionnel. Le déploiement massif de forces d'ingénierie de terrain, l'expansion des branches de conseil technologique des Big Four, ne sont que les incarnations provisoires de ce transfert : le marché cherche à acheter la capacité d'intégration que les structures internes ne savent pas encore produire.

Le symptôme institutionnel

Ce basculement ne se mesure pas au décompte brut des défaillances techniques. Qu'un rapport recense une hausse des incidents d'une année sur l'autre ne prouve pas une causalité stricte : cela peut refléter un volume croissant de systèmes en production, ou une transparence accrue des déclarations. Ce qui démontre la thèse, ce n'est pas la statistique de l'accident. C'est la réaction de l'institution.

Les signaux de 2026 sont organisationnels. L'AI Index de Stanford note cette année que la part d'entreprises opérant sans politique explicite de Responsable AI s'est effondrée, passant de 24 % à 11 %, tandis que les rôles dédiés à la gouvernance de l'IA progressent de 17 %. La gouvernance a cessé d'être un discours de conformité ou une posture éthique. Elle est devenue une ligne d'organigramme et un poste budgétaire au COMEX.

L'économie a tranché avant le régulateur. Lorsque de grands acteurs du conseil ou de la finance doivent assumer des préjudices financiers pour des livrables corrompus par des hallucinations ou des erreurs de raisonnement, l'origine se loge presque toujours au même endroit : l'absence de dispositifs capables de détecter ou de contenir la défaillance du modèle. Ce que le marché sanctionne, ce n'est pas le score du générateur. C'est la faillite du régime d'usage.

L'architecture comme condition de l'auditabilité

La capacité soutenable d'exploitation n'est pas un cadre de conformité surajouté par des auditeurs extérieurs. Elle s'incarne dans l'architecture technique, ou elle n'existe pas. Pour qu'un système soit institutionnellement exploitable, il doit répondre à une grammaire d'auditabilité stricte, structurée autour de trois invariants : un split rigoureux des flux de données, une calibration explicite des incertitudes, un domaine d'applicabilité dynamiquement contrôlé. La validation d'un système hautement automatisé cesse alors d'être un tampon obtenu une fois pour toutes à la mise en production.

Le terrain en fournit une instance, non une preuve générale. En santé connectée et en triage, PREDICARE traite le domaine d'applicabilité non comme une spécification passive, mais comme un objet d'architecture actif : le système re-mesure en continu ses propres frontières de confiance et suspend l'exécution dès que le contexte dévie de son espace de sécurité. L'auditabilité devient une propriété systémique plutôt qu'un document figé.

Ce qui devient l'actif

La thèse change de propriétaire, pas de contenu. Ce n'est pas une traduction de l'ingénierie en langage financier, avec la perte que toute traduction suppose. C'est la même assertion d'architecture, portant cette fois une responsabilité globale dont le COMEX doit répondre devant ses actionnaires, ses clients et ses régulateurs.

Devant un conseil d'administration, la direction ne répond plus d'avoir acheté le modèle le plus avancé du marché. Elle répond du régime sous lequel elle l'opère, et de sa capacité à le suspendre sans rompre la continuité d'activité. Le benchmark reste acheté par la DSI. Mais la captation de valeur ne se fait plus au niveau du modèle : elle se fait au niveau de l'organisation capable d'en répondre.

Reste à nommer précisément ce qui, dans cette organisation, tient lieu d'actif. Ce n'est pas le régime d'auditabilité, qui n'est qu'une implémentation. Ce n'est pas davantage le régime institutionnel, qui n'est qu'un ensemble de règles. C'est la capacité soutenable

d'exploitation elle-même, le premier niveau de notre hiérarchie, entendue comme capacité institutionnelle reproductible d'opérer durablement des systèmes autonomes.

Cette qualification n'est pas rhétorique. Un actif stratégique se reconnaît à cinq propriétés : il s'accumule, il s'imité mal, il ne se transfère que de manière imparfaite, il engendre des flux futurs, et il se valorise indépendamment de son support. La capacité soutenable d'exploitation les réunit. Elle se construit par apprentissage organisationnel et ne s'acquiert pas d'un bloc. Elle résiste à l'imitation parce qu'elle tient à des routines et à des responsabilités largement tacites. Elle ne se transfère qu'en se dégradant. Elle conditionne les flux que le système autorisera demain. Et elle conserve sa valeur quand on remplace le modèle sous-jacent. Le modèle, lui, a perdu ces propriétés une à une : interchangeable, imitable, transférable, commoditisé. Le support est devenu un intrant. La capacité de l'opérer est devenue l'actif.

Une théorie qui ne produit pas de prédiction reste une intuition. Celle-ci est falsifiable. À modèles comparables, deux organisations devraient présenter des écarts durables de performance économique corrélés à la qualité de leur régime d'auditabilité, non à la qualité intrinsèque des modèles employés ; et quand les modèles deviennent interchangeables, les écarts de valorisation devraient s'expliquer davantage par les capacités institutionnelles que par les actifs algorithmiques. Si l'inverse s'observait, si les écarts durables suivaient les modèles et non les régimes, la thèse serait fausse.

Il faut mesurer ce que ce constat déplace. Pendant trente ans, la technologie a été tenue pour l'actif stratégique. Puis les données. Les systèmes autonomes imposent une troisième bascule, où l'actif devient la capacité institutionnelle de les opérer. Ce n'est plus une migration de valeur à l'intérieur d'une définition stable de l'actif. C'est la définition de l'actif qui se déplace. Ce déplacement rejoint une question ancienne sur la nature du capital immatériel et organisationnel, que ce cycle n'a pas vocation à trancher ici.

Reste une question que ce cycle a ouverte sans la refermer. Si la gouvernance d'architecture, la conformité réglementaire et la gestion des risques financiers obéissent à la même grammaire d'auditabilité, comment les organisations vont-elles structurer leurs départements pour unifier ces référentiels concurrents ? C'est l'objet de la semaine qui vient.

[Série « les coûts cachés de déploiement de l'IA » : Synthèse dominicale, S11/12]