
Plus le monde fluctue, plus nos systèmes doivent être robustes

Et nous faisons exactement l'inverse. La qualité interne, quatrième variable du pilotage de nos organisations, sacrifiée au moment même où elle devenait vitale.

Clément Roques

Version 1 — 17/09/2026

Docteur en informatique (1994) — développeur et architecte technique de 1987 à 2015, aujourd'hui coach agile.

Diffusé sous licence Creative Commons Attribution 4.0 International (CC BY 4.0).

Sommaire

§1. Résumé exécutif

La thèse, en quatre lignes
Trois entrées, une seule démonstration
Ce que nous soutenons
Ce qui rend ce diagnostic difficile à voir
Ce que nous ne prétendons pas
Ce que nous proposons

§2. Ce que nous avons regardé pendant que ça se défaisait

§3. La quatrième dimension

La dimension qui n'a pas de place sur le schéma
Quatre inconnues, trois équations
Le cordon
Ce que l'agilité a réellement renversé

§4. Le basculement silencieux

Ce qu'on démonte quand on démonte un coût
Pourquoi personne ne le voit
La fuite mémoire
Le nom que les directions financières comprendraient

§5. Ce que la dette coûte vraiment

L'ordre de grandeur
Le marché de dupes
Parler de qualité sans dire le mot

§6. Performance et robustesse

Palier 1 — Le dilemme, hors du monde logiciel
Palier 2 — Le renfort, et la sortie de l'entre-soi
Palier 3 — Le retournement
Palier 4 — Fermeture, et ce que la section suivante devra porter

§7. Le motif du dévoiement

Une histoire que vous avez déjà entendue
Ce qui se conserve, et ce qui s'évapore
Le nom du mécanisme
Rendre les méthodes à elles-mêmes

§8. L'IA générative, amplificateur de régime

Le paradoxe que personne n'avait prévu dans ce sens
Ce que montrent les mesures de code
Ce que ces chiffres ne prouvent pas
L'IA n'a pas de direction

§9. À quoi ressemble une organisation robuste

Ce chapitre ne vous donnera pas de méthode, et c'est délibéré

Condition 1 — Quelqu'un répond de la qualité interne

Condition 2 — Quelqu'un peut arrêter la chaîne de production, et l'arrêt tient

Condition 3 — Les projets clos ont ajusté leur périmètre, et leurs documents le disent

Condition 4 — Les temps de respiration tiennent même sous tension

Condition 5 — Ceux qui disent non ont un endroit où reprendre leurs arbitrages

Ce que ces cinq conditions ne disent pas

§10. Les cinq questions sur votre variable d'ajustement

1. Qui, nommément, répond de la qualité interne ?

2. À quel niveau la dette technique est-elle arbitrée ?

3. Quand quelqu'un a-t-il dit « ce n'est pas prêt » pour la dernière fois, et que s'est-il passé ?

4. Sur les douze derniers mois, quels périmètres avez-vous réduits ?

5. Combien de capacité d'ingénierie n'est affectée à aucun projet ?

Comment lire vos réponses

§11. Ce que nous n'avons pas prouvé

§12. Ouverture — qui valide ce qu'on ne voit plus

Bibliographie

Ouvrages et articles

Rapports et enquêtes

§1. Résumé exécutif

Factuel : entre 2022 et 2024, la plupart des grandes organisations ont mis fin à leurs programmes de transformation agile. Aucune n'a décidé de ce qui en découlerait nécessairement. C'est ce qu'elles n'ont pas décidé qui fait l'objet de cette étude.

La thèse, en quatre lignes

Plus un environnement fluctue, plus les systèmes qui s'y déploient ont besoin de robustesse. La robustesse se paie en marges. Et la qualité interne est la première marge que nos organisations liquident — sans arbitrage, sans trace, et au moment de leur histoire où elle leur serait le plus utile.

Trois entrées, une seule démonstration

Ce document se lit par trois portes. Elles ne mènent pas à trois sujets : elles mènent à la même démonstration, par trois vocabulaires.

Si vous venez de la robustesse. Olivier Hamant oppose deux stratégies. La performance — optimisation, spécialisation, élimination des marges — est la bonne réponse à un environnement stable. La robustesse — redondances, marges, sous-optimalité assumée — est la condition de survie d'un environnement fluctuant. Cette étude applique cette opposition à un objet qu'il ne traite pas : la qualité interne des systèmes techniques, et le dispositif humain qui la tenait. Elle montre où la marge a été prise, et ce qu'elle protégeait.

Si vous venez du pilotage. Le triangle périmètre-coût-délai en cache une quatrième dimension, la qualité interne. C'est la seule des quatre qui ne remonte à aucun comité de direction, donc la seule qui absorbe l'écart lorsque les trois autres sont tenues. Cette étude établit par quel mécanisme elle redevient la variable d'ajustement dès que disparaît le dispositif qui la protégeait, et ce que ce basculement coûte — chiffres publics, sources primaires, régime de preuve explicité dans le dossier de méthode publié à côté de cette étude.

Si vous venez de la complexité. Ross Ashby a formulé en 1956 la loi de la variété requise : un système de régulation doit disposer d'au moins autant de variété que le système qu'il régule. Une organisation qui simplifie ses instances pendant que son environnement se complexifie perd sa capacité à le réguler — méca-

niquement, sans faute de personne. Cette étude documente ce que produit cet écart, et pourquoi l'effondrement, le jour où il vient, paraît soudain à des gens qui n'ont rien décidé de tel.

Ces trois chemins se rejoignent au chapitre 6, où la performance et la robustesse cessent d'être des valeurs pour redevenir ce qu'elles sont : deux stratégies, chacune adaptée à un régime d'environnement, et une seule des deux adaptée au nôtre.

Ce que nous soutenons

La qualité interne est la quatrième dimension du pilotage, et c'est la seule que personne ne pilote. Le triangle classique — périmètre, coût, délai — en cache une quatrième. Dans la gestion de projet traditionnelle, le périmètre et le délai sont fixés ; la qualité devient mécaniquement la variable qui absorbe l'écart, sans que quiconque l'ait choisi. L'agilité véritable inversait précisément ce dispositif : coût et délai tenus par les itérations et des équipes stables, qualité déclarée non négociable, et **le périmètre fonctionnel comme unique variable d'ajustement assumée**. Ce renversement n'était pas un raffinement de méthode. C'était le cœur économique de l'affaire, et il tenait par un dispositif — des instances, un mandat, quelqu'un dont le métier était de dire non.

Quand la transformation s'arrête, l'invariant bascule sans que personne ne le décide. Déclarer l'agilité mature et dissoudre les dispositifs qui la portaient ne supprime pas la pression sur le périmètre : les besoins des clients, eux, ne s'arrêtent pas. La gouvernance revient au pilotage par le budget et les jalons ; le périmètre reste exigé, le coût et le délai restent verrouillés — et la qualité redevient la variable d'ajustement. Silencieusement, parce qu'elle est la seule des quatre dimensions qui ne remonte à aucun comité de direction. C'est ainsi qu'une organisation peut se dégrader pendant trois ans sans qu'aucune décision identifiable n'ait été prise. Aucun renoncement pris isolément ne fait tomber le système, et l'effondrement, le jour venu, paraîtra soudain.

Ce basculement survient au pire moment possible. Les travaux du biologiste Olivier Hamant posent une distinction que les directions n'ont pas : la performance — optimisation, spécialisation, élimination des marges — est la bonne stratégie d'un environnement stable ; la robustesse — marges, redondances, sous-optimalité assumée — est la condition de survie d'un environnement fluctuant. Or l'environnement des directions informatiques depuis 2024 est le plus instable observé depuis des décennies : irruption de l'IA générative, pression cyber, chocs budgétaires et géopolitiques. C'est exactement le moment que nos organi-

sations ont choisi pour liquider leurs marges. L'injonction est parfaitement rationnelle localement, ligne budgétaire par ligne budgétaire, et fragilisante globalement.

Ce qui rend ce diagnostic difficile à voir

Personne, dans cette histoire, n'agit par malveillance. C'est le propre d'un système complexe.

Les décisions qui produisent ce résultat sont prises par des gens rationnels et de bonne foi, et souvent à partir d'indicateurs bien construits. Lorsqu'en janvier 2023 Capital One supprime plus de mille cent postes de sa filière agile, le communiqué ne dit pas que l'agilité a échoué : il dit qu'elle a **réussi**, qu'elle est désormais intégrée aux pratiques d'ingénierie. C'est une déclaration de maturité, et elle est en partie vraie. C'est précisément ce qui la rend redoutable.

Le mécanisme porte un nom, et il n'est pas nouveau. Une méthode conçue autour de la qualité est adoptée pour ses promesses, industrialisée, puis convertie en instrument de suivi des cadres budgétaires. Le lean l'a subi avant l'agilité — le jidoka, qui était le droit d'arrêt, est devenu du reporting. Le change management l'a subi aussi. L'agilité à son tour : la vélocité comme mesure de productivité, les taux d'adoption trimestriels. Trois générations de méthodes, un seul motif, et la même leçon jamais tirée. Nous ne critiquons aucune de ces méthodes dans ce document ; nous nommons ce qui les capture toutes. Et nous observons que le même mécanisme est en train de se refermer sur l'accompagnement de l'adoption de l'IA, sous nos yeux.

Sur ce dernier point, les données publiques disent quelque chose d'inconfortable. Le rapport DORA 2024 mesure que pour vingt-cinq pour cent d'adoption supplémentaire de l'IA, le débit de livraison **recule** de 1,5 % et la stabilité de 7,2 %. Les auteurs signalent eux-mêmes que ce résultat est contraire à leurs attentes. L'IA ne crée pas le régime décrit ici : elle accélère la dégradation dans un régime se dégradant déjà, et elle accélère les gains dans un régime sain.

Ce que nous ne prétendons pas

Nous ne prétendons pas savoir quand. Nous soutenons qu'un régime produit une trajectoire ; nous ne prétendons pas connaître la vitesse à laquelle il la parcourt. Les données relatives à l'IA sont corrélationnelles. L'observation de terrain qui ouvre ce document est un témoignage direct, et elle sert à faire reconnaître un enchaînement, jamais à l'établir.

Ce que nous proposons

Cinq conditions, décrites au chapitre 9, qui distinguent une organisation robuste d'une organisation qui se croit telle. Ces conditions ne forment pas une méthode et ne se déploient pas : elles se vérifient. L'une d'elles se vérifie par une seule question, et vous pouvez vous la poser dès maintenant : **quand, pour la dernière fois, quelqu'un a-t-il arrêté quelque chose, et que lui est-il arrivé ?**

Et si vous le souhaitez, pour commencer par le plus facile à expérimenter, nous vous proposons cinq questions à vous poser, au chapitre 10. Elles tiennent sur une page, aucune ne demande d'outil, de budget ni de programme, et voici la première : **qui, dans votre organisation, répond de la qualité interne du code — et à quel étage de la hiérarchie siège-t-il ?**

§2. Ce que nous avons regardé pendant que ça se défaisait

Il y a cinq ans, la direction des systèmes d'information dans laquelle je travaillais a lancé un grand programme d'amélioration. Plusieurs domaines y figuraient ; l'agilité était l'un d'eux. Deux indicateurs avaient été retenus pour en suivre les progrès, et ils étaient élégants : la **largeur** de l'agilité — combien d'équipes étaient concernées — et sa **profondeur** — jusqu'où les pratiques étaient réellement adoptées dans chacune.

Personne, ne faisait mal son travail. Le programme était sérieux, les indicateurs étaient mesurables, et ils avaient ce mérite rare de ne pas se contenter de compter des équipes : quelqu'un avait pris la peine de penser qu'une adoption superficielle n'en était pas une.

Nous étions plusieurs coachs agile à trouver cela bizarre, et nous savions dire pourquoi.

La largeur traitait l'agilité comme une fin en soi, alors que toutes les équipes n'ont pas un environnement qui s'y prête durablement : une équipe de deux personnes, un produit qui part en production une fois par an, et la question de l'agilité ne se pose pas dans les mêmes termes — ou ne se pose pas du tout. Mesurer combien d'équipes sont passées revient à demander à celles-là de passer quand même, pour faire bien...

La profondeur, elle, partait d'une intention juste : refuser qu'une adoption de façade compte autant qu'une pratique réelle. Mais descendue d'un cran de hiérarchie à l'autre, elle s'est transformée en injonction — faites mieux — et une injonction de faire mieux ne produit pas ce qu'elle demande. L'indicateur se dénaturait de lui-même en descendant.

Nous n'avons pas gardé ces objections pour nous ; elles n'ont simplement jamais eu d'endroit où être entendues. Personne n'est venu nous les demander, et il n'existait aucune instance où elles auraient compté.

Et il y avait une troisième chose à dire, et celle-là, je ne l'ai comprise que bien plus tard. Ces deux indicateurs mesuraient le déploiement de l'agilité. Ni l'un ni l'autre ne mesurait ce que l'agilité était censée produire — la **qualité** de ce que nous livrions, notre capacité à absorber un **imprévu**, la **santé** des équipes qui

portaient tout cela. Or un indicateur de déploiement possède une propriété implicite : **il se remplit**. Il arrive nécessairement un jour où il est plein, et ce jour-là, il n'y a plus rien à faire.

Ce jour est arrivé environ deux ans plus tard. Un cabinet de conseil — des experts, un travail sérieux, des livrables que je ne songerais pas à critiquer — a été chargé de redéfinir le modèle opérationnel de la direction. Dans ce nouveau modèle, l'agilité n'occupait plus une place de transformation permanente, et le constat qui le justifiait était imparable : la transformation était achevée pour les équipes qui en avaient besoin. Désormais, l'agilité serait considérée « en mode run ».

Il faut s'arrêter sur ce mot. Dans le vocabulaire des directions informatiques, le run désigne ce qui tourne : ce qu'on exploite, ce qu'on maintient, ce pour quoi on ne prévoit plus de projet ni de budget de transformation. Faire passer une chose en run, c'est déclarer qu'elle est acquise. C'est affirmer qu'elle ne se dégradera pas toute seule.

Là encore, nous étions plusieurs coachs à juger que la formule dénaturait ce que nous avions passé des années à installer, et nous n'avons eu aucun mal à dire pourquoi. Là encore, personne n'est venu nous le demander. Je n'écris pas cela pour m'en plaindre. Je l'écris parce que c'est simplement factuel, et parce qu'il m'a fallu longtemps pour comprendre qu'il n'était pas anodin : la décision se prenait à un endroit où l'information dont nous disposions n'était pas attendue.

La suite s'est jouée sans réunion et sans décision. Dans les mois qui ont suivi, les demandes d'accompagnement se sont espacées, puis taries. Les grandes directions qui faisaient appel à nous avaient une réponse toute prête, et elle était polie : nous pouvons nous débrouiller par nous-mêmes, nous n'avons plus besoin de vous. Ce message avait même tendance à nous plaire, car il signifiait que nous avions bien tenu notre rôle pédagogique.

Aucune de ces décisions n'était contestable prise séparément. Chacune a été prise par des responsables compétents, avec de bonnes raisons, et aucune n'avait besoin de notre avis pour être prise. C'est précisément ce qui rend la séquence intéressante, et c'est le sujet de cette étude.

Alors je vous pose la question, celle que nous n'avons pas été en mesure de poser à l'époque : dans votre organisation, qu'est-ce qui a été déclaré acquis ? Et qui aurait le mandat d'aller vérifier que l'acquis est toujours stable ?

§3. La quatrième dimension

Où nous en sommes. La section précédente a raconté une séquence en trois temps : deux indicateurs élégants qui ne mesuraient pas ce qu'ils croyaient mesurer, un mode de fonctionnement déclaré acquis, et des demandes d'accompagnement qui se sont taries sans qu'aucune réunion ne le décide. Trois décisions dont aucune n'était contestable prise séparément. Nous avons posé le symptôme sans le nommer.

Ce que nous allons voir ici. Le mécanisme qui le produit. Nous verrons d'abord qu'il manque une dimension au triangle que tout le monde connaît, ensuite pourquoi c'est toujours celle-là qui cède, puis ce que trois traditions de management sans lien entre elles ont inventé pour l'empêcher — et enfin pourquoi ce qu'on prend pour une préférence de méthode est en réalité un arbitrage économique. À la sortie de cette section, la thèse de l'étude sera posée. Tout ce qui suit en découle.

La dimension qui n'a pas de place sur le schéma

Vous l'avez probablement vu cent fois, sur un tableau blanc, en ouverture d'un séminaire ou dans la troisième diapositive d'un comité de pilotage. Un triangle . Le présentateur l'entoure d'un cercle, explique qu'on ne peut pas tenir les trois à la fois, qu'il faut choisir — et tout le monde acquiesce, parce que c'est vrai.

Regardez la diapositive suivante. Neuf fois sur dix, la qualité y est apparue. Mais regardez où : au centre du triangle dans certaines versions, à l'une des pointes dans d'autres, en légende sous le schéma dans les plus honnêtes. Sa place flotte d'une école à l'autre, et personne ne s'en émeut. Cette instabilité graphique n'est pas une coquetterie de consultant : **une dimension dont la place flotte est une dimension qui n'a pas de titulaire.** L'argument est d'usage, nous ne le durcirons pas davantage — mais il est un bon indicateur de ce qui va suivre.

Figure 3.1 — Le triangle de fer, et la place flottante de la qualité

Trois sommets nommés, trois emplacements possibles pour un quatrième terme — et aucun qui fasse consensus

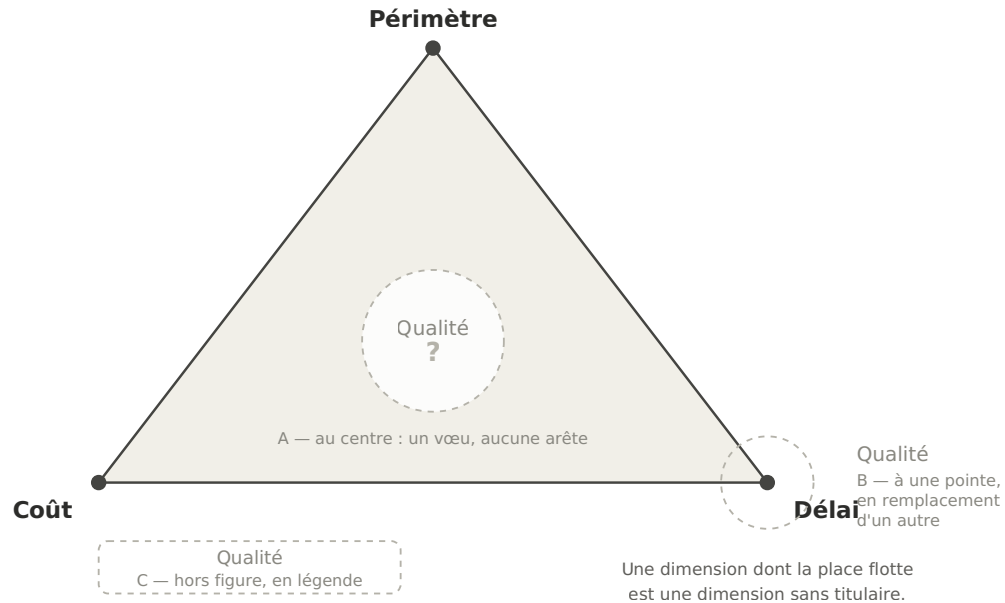


Figure 3.1 — Trois sommets nommés, et trois emplacements possibles pour un quatrième terme. Aucun ne fait consensus, et c'est le premier symptôme.

Disons tout de suite de quelle qualité nous parlons, parce que le mot est chargé de sémantiques diverses. Nous entendons ici la **qualité interne** : la structure du code, sa testabilité, sa lisibilité par quelqu'un qui arrive demain, la réversibilité des choix d'architecture, la capacité à modifier le système dans dix-huit mois sans le reconstruire. Ce n'est pas la satisfaction client, ce n'est pas la conformité fonctionnelle, ce n'est pas l'absence de bugs visibles — tout cela relève du périmètre, et se négocie en comité de pilotage. La qualité interne, elle, ne se voit pas depuis l'extérieur du système. C'est même sa définition la plus opératoire.

L'autre mot qui accompagne cette dimension mérite le même soin, car nous l'emploierons souvent : celui de **dette technique**. Il a un auteur et une date — Ward Cunningham, qui le forge au début des années 1990 pour expliquer à sa direction pourquoi le remaniement du code n'était pas un luxe. Et la définition d'origine est plus exigeante que l'usage qu'on en fait aujourd'hui.

Cunningham y est revenu lui-même en 2009, et il commence par écarter le contresens le plus répandu : « je ne suis jamais favorable à ce qu'on écrive du code de mauvaise qualité, mais je suis favorable à ce qu'on écrive du code qui reflète la compréhension qu'on a du problème aujourd'hui ». La dette n'est donc pas du code mal écrit. C'est **l'écart entre ce que le code exprime et ce que l'on a fini par comprendre du problème** — écart normal, inévitable, et même

souhaitable, puisqu'il permet de livrer avant d'avoir tout compris. Ce qui la rend coûteuse est de ne jamais y revenir : « si vous développez un programme pendant longtemps en n'y ajoutant que des fonctionnalités, sans jamais le réorganiser pour refléter votre compréhension, [...] à la fin ce programme ne contient plus aucune compréhension ».

Retenons ce que cette définition suppose, parce que c'est là que tout se joue : elle suppose un emprunt, donc une décision — et elle suppose que celui qui emprunte pourra rembourser.

Or l'emprunt se contracte au niveau où l'on écrit le code : un développeur, un responsable technique, à la fin d'une semaine où il faut livrer. Le remboursement, lui, se décide au niveau où l'on arbitre un budget ou un plan de charge — c'est-à-dire ailleurs, par quelqu'un à qui l'emprunt n'a jamais été signalé, et le plus souvent un ou deux exercices plus tard. La métaphore n'a pas été mal comprise par de mauvais praticiens. Elle a été conçue pour une échelle où l'emprunteur et le rembourseur sont la même personne, et elle est employée à une échelle où ils ne le sont plus.

C'est exactement l'écart de rang que nous venons de décrire à propos de la qualité interne, et la coïncidence n'en est pas une : une décision prise à un étage, une contrepartie qui se règle à un autre, et rien qui relie les deux. **C'est le mécanisme central de cette étude**, et nous le retrouverons à chaque section.

Et elle possède une propriété singulière dont toute la suite découle : **elle se dégrade sans action explicite**.

Le périmètre ne s'étend pas pendant la nuit. Le budget ne se consomme pas tout seul. Le retard, au pire, se constate — quelqu'un finit toujours par ouvrir le planning initial. La qualité interne, elle, se dégrade par le simple passage du temps et l'accumulation des modifications, sans que quiconque ait rien décidé. Manny Lehman, qui dirigea le département d'informatique de l'Imperial College après avoir étudié chez IBM la façon dont les grands programmes vieillissent, l'observait déjà sur les grands systèmes : à mesure qu'un logiciel évolue, sa complexité croît et sa qualité perçue décline, sauf si un travail actif est mené pour l'en empêcher.

Retenons la conséquence, parce qu'elle absout d'avance tous les acteurs de cette étude, et qu'il faut l'avoir en tête pour lire les quarante pages qui suivent sans y chercher un coupable : **il n'est besoin de personne pour dégrader la qualité : il suffit que l'effort ait cessé**.

Il y a donc une dimension qui cède. Reste à comprendre pourquoi c'est toujours elle.

Quatre inconnues, trois équations

Posons-le en termes que tout ingénieur reconnaîtra. Quatre dimensions, et aucun système de production réel ne les fixe toutes les quatre — parce qu'un système qui les fixerait toutes n'aurait plus aucun degré de liberté, donc plus aucune capacité d'absorber le moindre aléa. Quatre inconnues, trois équations : il reste un degré de liberté. La question n'est donc jamais s'il existe une variable d'ajustement. La question est de savoir **laquelle**, et surtout **qui l'a choisie**.

Passons alors les quatre dimensions en revue, non pas selon leur importance, mais selon quelque chose de bien plus prosaïque : qui, dans l'organisation, en répond.

Le périmètre a un comité de pilotage, un product owner, et au bout de la chaîne un client qui téléphone. Le coût a un contrôle de gestion, une ligne budgétaire, un directeur qui signe. Le délai a un jalon, un plan, un métier qui attend, et souvent un engagement contractuel derrière. Trois dimensions, trois instances, trois rôles portant explicitement un mandat de surveillance.

La qualité interne, elle, a des titulaires — et c'est justement ce qui rend l'asymétrie difficile à voir. Les architectures existent, et elles ne datent pas d'hier : architecture technique, logicielle, fonctionnelle, architecture d'entreprise. Le framework SAFe est même allé plus loin en instituant un rôle dont c'est explicitement le mandat, l'architecte système, chargé de porter un contrepoids aux décisions fonctionnelles et budgétaires.

L'écart n'est donc pas qu'il n'y ait personne. **C'est que personne ne siège au rang où l'arbitrage se tranche.** Le périmètre a un directeur métier (un Business Owner), le coût un directeur financier qui signe, le délai un engagement contractuel et un comité qui le surveille. L'architecte, lui, a un avis. Il l'exprime, il l'argumente, et il est écouté — jusqu'au moment où l'arbitrage se rend, où il n'est pas dans la pièce, ou bien il est présent mais sans le poids des trois autres.

C'est une différence de rang, pas de compétence ni de sérieux, et il faut y insister parce que la lecture facile serait de croire que ces gens font mal leur travail. Ils le font bien. Ils ne le font pas d'assez haut. Et un titulaire sans poids est, pour la suite de cette étude, plus intéressant qu'une absence de titulaire : il donne à l'organisation l'apparence d'une couverture, et cette apparence retarde d'autant le moment où quelqu'un s'aperçoit que la dimension cède.

Elle reste donc la seule des quatre à pouvoir absorber un écart **sans réunion, sans compte rendu, sans signature.**

Figure 3.2 — Qui répond de quoi

Quatre dimensions, trois instances. La quatrième absorbe l'écart sans réunion, sans compte rendu, sans signature.

DIMENSION	QUI EN RÉPOND	CE QUI EN GARDE TRACE
Périmètre	Comité de pilotage, product owner, et un client qui téléphone	Une décision d'arbitrage
Coût	Contrôle de gestion, un directeur qui signe	Une ligne budgétaire
Délai	Un jalon, un plan, un métier qui attend	Un engagement daté
Qualité interne structure, testabilité, réversibilité	— personne —	Aucun document ne s'ouvre

Personne ne décide de sacrifier la qualité : c'est la position de variable libre qui décide à la place de tout le monde.

Figure 3.2 — L'asymétrie ne porte pas sur l'importance accordée à chaque dimension, mais sur une question beaucoup plus prosaïque : qui en répond, et quel document s'ouvre quand elle bouge.

Nous tenons ici le point de bascule de cette section, et nous voudrions qu'il soit lu exactement pour ce qu'il est : personne ne décide de sacrifier la qualité ; c'est la position de variable libre qui décide à la place de tout le monde. L'argument est structurel, pas moral. Les directions dont nous parlerons dans cette étude arbitrent avec ce que leur système leur présente comme arbitrable, et elles le font sérieusement.

Deux objections se lèvent ici, et toutes deux méritent mieux qu'une réfutation.

« **Chez nous, nous fixons les quatre.** » Nous ne la contesterons pas — nous la déplacerons. Si les quatre dimensions sont réellement tenues et que la charge, elle, n'a pas diminué, alors l'écart est allé quelque part. Il est parti dans le turnover, dans la fatigue des équipes, dans un report silencieux vers l'exploitation qui absorbera les incidents, dans une sous-traitance qui portera la dette hors du bilan. La bonne question n'est jamais « fixez-vous les quatre ? », à laquelle chacun répond oui de bonne foi. Elle est : **où se fait l'ajustement, et où est passé l'écart ?**

« **Chez nous, la qualité n'est pas ajustable.** » Celle-là, nous la concédons franchement, et sans réserve mentale : dans l'aéronautique, le nucléaire, le dispositif médical, une norme et un régulateur transforment la qualité en seuil, pas en variable. La certification n'est pas négociable, le dossier de preuve non plus. Dans

ces industries, la variable d'ajustement est le délai ou le coût, et c'est parfaitement assumé. C'est le contre-exemple le plus solide qu'on puisse opposer à cette étude, et nous n'allons pas le traiter en note de bas de page : il a son chapitre, celui des limites. Notons seulement, en passant, ce qu'il démontre — quand une instance existe pour la qualité, elle cesse d'être la variable libre. C'est exactement notre thèse, vérifiée par son exception.

Le cordon

Les régimes de production qui ont réellement protégé la qualité ont tous dû résoudre ce problème. Ce qui frappe, quand on les regarde ensemble, c'est qu'ils l'ont résolu de la même façon sans jamais s'être lus.

Toyota, avec le jidoka et le cordon d'andon : n'importe quel opérateur peut arrêter la chaîne. Pas signaler, pas alerter, pas remplir un formulaire — **arrêter**. Le Manifeste agile et la définition of done : c'est l'équipe qui fixe le seuil, et c'est elle qui déclare que ce n'est pas fini. ADKAR, enfin, avec son Renforcement : une étape à part entière du modèle, qui ne se déclare pas depuis un tableau de bord.

Trois traditions, trois décennies, aucune filiation. Une seule et même trouvaille : **la décision de qualité se prend là où l'information existe, et elle est opposable.**

Écartons tout de suite la lecture sentimentale, qui affaiblirait l'argument : il ne s'agit pas de dire que le terrain serait plus vertueux que le sommet. La raison n'est pas morale. Elle est informationnelle, et c'est un cybernéticien qui l'a formulée le plus proprement.

Ross Ashby, en 1956 : un régulateur ne peut maintenir un système dans un état souhaité que s'il dispose d'au moins autant de variété — de nombre d'états distincts — que le système qu'il régule. C'est la loi de la variété requise. Sa conséquence est brutale et rarement tirée : **un déficit de répertoire ne se compense ni par l'effort, ni par l'autorité, ni par la bonne volonté.** Il ne se compense que par de la variété.

Ashby autorise une seconde stratégie, et il faut la traiter avant d'aller plus loin : réduire la variété à réguler plutôt qu'augmenter celle du régulateur. Elle est réelle — standardiser, réduire le parc applicatif, déléguer à un fournisseur — et ses limites sont strictes. Elle ne porte que sur la variété que l'organisation choisit, jamais sur celle qu'elle subit : failles découvertes, déplacements d'usage, obligations nouvelles. Et lorsqu'elle prend la forme d'une délégation, elle ne supprime pas la régulation : elle la déplace, en laissant intacte la part qui ne se délègue pas — et dont la qualité interne décide.

Appliquons. Celui qui décide au sommet ne voit pas le système : il voit ce que la remontée lui en transmet. Or un reporting est un encodage avec perte — et voici le point qu'il ne faut surtout pas manquer : **cette perte n'est pas un défaut, c'est la fonction**. Un reporting qui ne réduirait rien serait illisible, donc inutilisable ; on ne pilote pas une organisation en lisant son journal d'événements. Mais ce qui a été perdu à la compression ne se décompresse pas plus haut. Il n'est pas caché, il n'existe plus.

Vient alors la meilleure objection de toute l'étude, et elle vient toujours de la personne la plus compétente de la salle. « **Et si nous avions de meilleurs indicateurs ?** » Couverture de tests, analyse statique, mesure de dette technique, observabilité — la technologie a beaucoup progressé, pourquoi ne pas remonter la qualité comme on remonte le reste ?

La réponse tient en une distinction, et elle est décisive : **un indicateur transmet un état, jamais un arbitrage**. Votre tableau de bord vous dira qu'il y a de la dette sur ce composant. Il ne vous dira pas si la contracter ici, maintenant, avant cette échéance-là, sur ce module-ci, est raisonnable. Ce jugement-là exige de savoir ce que ce module fera dans six mois, qui le maintient aujourd'hui, ce qu'on y a rafistolé l'an dernier en attendant la prochaine version d'Angular ou de Kubernetes, et ce qu'on a promis à qui. C'est local, c'est largement tacite, et ça ne se met pas en colonne. Il existe une seconde réponse, plus redoutable encore — tout indicateur qui remonte devient un objectif et cesse de mesurer — mais nous ne la déploierons pas ici : elle a sa section plus loin.

Un mot, pour finir ce palier, sur le mot lui-même. On dit « déléguer » la décision de qualité, et nous continuerons de le dire parce que c'est le mot d'usage. Mais il ment. Déléguer suppose qu'on détenait quelque chose et qu'on consent à le confier. Ce que dit Ashby est autre chose : la décision n'était **jamais** correctement prenable en haut. Il n'y a pas de transfert de pouvoir, il y a une contrainte d'ingénierie.

Ici, le lecteur d'entreprise objectera qu'il connaît déjà cette idée, et il aura raison. Elle a un propriétaire, et c'est par lui qu'elle est entrée dans les organisations : Peter Drucker, *Management by Objectives — and Self-Control*, 1954. Les travailleurs du savoir — définis par leur asymétrie d'information, non par leur diplôme — sont les mieux placés pour décider de la façon dont leur travail se fait, et le seul contrôle qui fonctionne sur eux est l'auto-contrôle.

Et il vaut la peine de regarder ce que les cadres méthodologiques contemporains ont retenu de lui, car ils ont retenu autre chose. Le neuvième principe de SAFe s'intitule « Decentralize Decision-Making », mais l'épigraphe qui l'ouvre n'est pas

de Drucker — elle est de Jim Collins — l'auteur de *Built to Last* et *Good to Great*, deux enquêtes sur ce qui distingue les entreprises qui durent —, et elle dit ceci : « les entreprises les plus innovantes tendent à pousser les décisions aussi bas que possible dans l'organisation, donnant aux gens de tous niveaux l'occasion d'aller vite, d'exploiter leur créativité, d'appliquer leur intelligence et d'assumer des responsabilités ». Le texte qui suit justifie la décentralisation par la technologie disruptive, la compétition intense et la fugacité des opportunités.

C'est de la décentralisation **par la vitesse**. Ce n'est pas la même chose que la décentralisation par l'information, et l'écart n'est pas anodin : une décision qu'on décentralise pour aller plus vite se recentralise dès que la vitesse cesse d'être le problème. Une décision qu'on décentralise parce qu'elle n'est correctement prenable nulle part ailleurs ne se recentralise jamais — elle se perd, simplement.

Le rapprochement des deux hommes est instructif précisément par leur écart. **Drucker argumente en savoir, en motivation, en direction, et il aboutit à une délégation authentique et révocable : le sommet détenait légitimement les objectifs, et il choisit de confier. Ashby argumente en variété, et il aboutit à une impossibilité : il n'y avait rien à confier, parce que rien n'était détenu.** Drucker explique pourquoi la décentralisation fonctionne mieux. Ashby explique pourquoi la centralisation ne pouvait pas fonctionner. Nous avons besoin des deux : de Drucker pour que le lecteur entre dans le sujet par une porte qu'il connaît, d'Ashby parce que lui seul tient contre l'objection des indicateurs.

Et il faut dire ce qu'est devenu Drucker, parce que c'est un avertissement dont cette étude n'a pas fini d'avoir besoin. Le MBO tel qu'il s'est installé dans les entreprises, c'est une cascade d'objectifs chiffrés et un entretien annuel — c'est-à-dire très exactement le commandement qu'il voulait remplacer. La moitié qui mesure a survécu. La moitié qui rend la décision a disparu. Le nom complet, lui, est resté. Personne n'a décidé cette amputation : c'était simplement la seule moitié installable sans que le sommet cède quoi que ce soit. Gardez cette forme en tête — nous la retrouverons à la section suivante, appliquée non plus à un modèle de management, mais à un cordon.

Deux précautions avant de refermer ce palier, faute de quoi tout ceci deviendrait un slogan. D'abord, une équipe qui peut alerter mais pas refuser n'a rien reçu du tout : son alerte remonte compressée, et nous voilà revenus au point de départ. L'andon ne vaut que parce que la ligne s'arrête réellement. Ensuite, cette étude ne soutient nulle part que le terrain aurait toujours raison. Confier la décision à un périmètre qui n'a ni le contexte ni la compétence de juger ne supprime pas le

point de défaillance, il le déplace. La capacité de juger est une condition, pas un supplément — nous y reviendrons, c'est l'une des conditions d'une organisation robuste.

Ce que l'agilité a réellement renversé

Reprenons maintenant le triangle du début, avec tout cela en tête.

Gestion de projet classique : le périmètre est contractualisé, le délai est engagé, le coût se négocie à la marge. La qualité absorbe le reste — sans jamais s'écrire nulle part, puisqu'elle n'a aucun endroit où s'écrire.

Agilité : le renversement est terme à terme, et il est complet. Le coût est fixé par une équipe stable dont on ne fait pas varier la taille. Le délai est fixé par une timebox qui ne se prolonge pas. La qualité est déclarée non négociable et protégée par un dispositif opposable. Et le périmètre devient l'unique variable d'ajustement — **assumée, celle-là**. On livre moins, on le dit, et on le dit maintenant.

Figure 3.3 — L'inversion : ce que l'agilité renverse réellement

Même quadrilatère de contraintes. Ce qui change, c'est laquelle des quatre dimensions reste libre — et si quelqu'un l'a choisie.

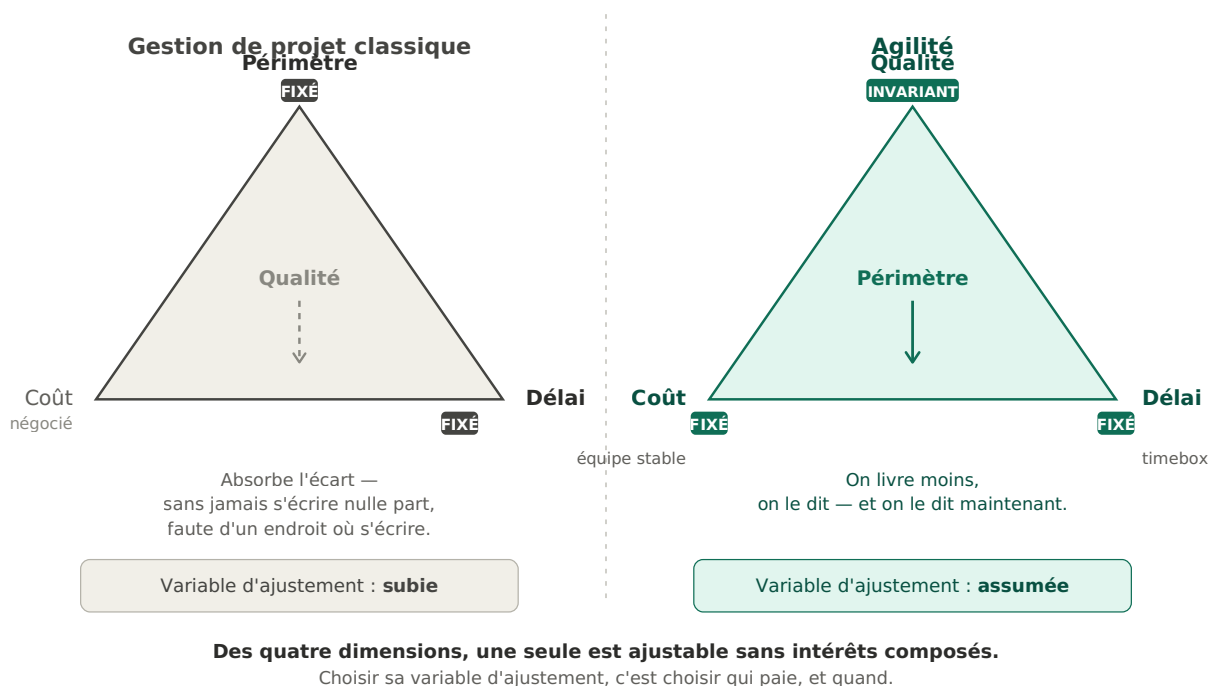


Figure 3.3 — Le renversement est terme à terme. Ce qui change n'est pas le nombre de contraintes, c'est laquelle des quatre dimensions reste libre — et si quelqu'un l'a choisie.

Ce renversement passe presque toujours pour une préférence méthodologique, une question d'école, un débat de praticiens. Ce n'en est pas une. C'est un arbitrage économique, et il repose sur une asymétrie que personne n'énonce jamais : **des quatre dimensions, une seule est ajustable sans intérêts composés.**

Réduire le périmètre coûte aujourd'hui, une fois, c'est visible, et quelqu'un sera mécontent — c'est désagréable, c'est traçable, et c'est fini. Réduire la qualité coûte demain, indéfiniment, sur chaque modification à venir, et personne ne saura jamais rattacher la facture à la décision qui l'a produite.

Choisir sa variable d'ajustement, ce n'est donc pas choisir une méthode. **C'est choisir qui paie, et quand.**

Nous voudrions clore cette section par une question, et nous la posons sans effet de manche, parce que la réponse nous intéresse réellement. Dans votre organisation, quelqu'un a-t-il le cordon ? Pas le droit d'alerter. Pas une ligne dans un registre de risques. Le pouvoir d'arrêter la ligne. Et si oui — que s'est-il réellement passé la dernière fois qu'il a tiré dessus ?

Posez la question autour de vous, elle est plus dérangeante qu'elle n'en a l'air. Car la plupart des organisations que nous décrivons dans cette étude **ont eu** ce cordon. Il a été installé entre 2015 et 2022, souvent à grands frais, parfois avec des cabinets, toujours avec des convictions. Puis il a disparu — et il ne s'est rien passé de spectaculaire ce jour-là.

C'est même exactement à cela qu'on le reconnaît.

§4. Le basculement silencieux

Où nous en sommes. La section précédente a posé la mécanique : quatre dimensions, trois instances, une variable libre — et la qualité interne occupe cette place parce qu'elle est la seule dont personne ne répond au rang où l'arbitrage se tranche. Nous avons vu aussi ce que trois traditions de management avaient inventé pour l'en sortir : un cordon, opposable, tenu par ceux qui disposent de l'information. Et nous avons terminé sur une remarque que cette section doit maintenant honorer — la plupart des organisations ont eu ce cordon, et il a disparu sans bruit.

Ce que nous allons voir ici. Comment il disparaît. Nous suivrons la séquence réelle, telle qu'elle s'est déroulée à partir de janvier 2023, puis nous chercherons la raison pour laquelle personne ne la voit se produire — c'est le cœur du diagnostic, et ce n'est ni la mauvaise foi ni l'incompétence. Nous donnerons enfin à ce phénomène le nom que les directions financières emploieraient si on le leur présentait dans leur langue.

En janvier 2023, Capital One supprime plus de mille cent postes de sa famille de métiers agile — Scrum Masters, coachs, delivery leads — en une seule décision. Le communiqué qui l'accompagne fera date, et il faut le lire de près, parce qu'il ne dit pas ce qu'on lui a fait dire. Il ne dit pas que l'agilité a échoué. Il dit qu'elle a réussi : elle est désormais « intégrée aux pratiques d'ingénierie ». C'est une déclaration de maturité, et elle est en grande partie vraie.

C'est bien ce qui rend la suite intéressante.

La section précédente a posé une mécanique : dans un système à quatre dimensions, quatre forces, dont trois sont tenues, la quatrième absorbe l'écart, car elle est la seule dont aucun titulaire ne siège là où l'arbitrage se rend. Dans les meilleurs cas, ce que j'ai vu parfois, l'architecte a seulement un avis consultatif, pas un droit de veto. Cette mécanique n'a jamais disparu pendant les transformations agiles. Elle y était tenue en échec par un dispositif construit exactement pour ça. Reste à raconter ce qui se passe quand on démonte le dispositif en croyant démonter un coût — et à comprendre pourquoi, dans les organisations où ce basculement s'est produit, personne n'a rien vu passer. Nous suivrons la séquence telle qu'elle s'est réellement déroulée, puis nous chercherons ce qui la rend invisible ; c'est cette invisibilité, et non le basculement lui-même, qui est le vrai sujet de cette section.

Ce qu'on démonte quand on démonte un coût

La déclaration de maturité vient d'abord, et elle est sincère. Beaucoup de ces organisations ont réellement acquis en dix ans des pratiques d'ingénierie qu'elles n'avaient pas — l'intégration continue est devenue banale, les tests automatisés ont cessé d'être exotiques, le déploiement fréquent ne fait plus peur. Ce qui a été intégré l'a vraiment été. Ce qui n'a pas été vu, c'est ce qui n'était pas intégrable.

Vient ensuite la dissolution, et elle est arithmétiquement imparable. Coachs, centres d'excellence, communautés de pratique : vus depuis un compte de résultat, ce sont des coûts de structure sans production directe, et leur suppression est l'arbitrage le plus simple qu'une direction puisse rendre. Royal London retire 90 % de ses coachs dans les mois qui suivent Capital One ; une grande banque britannique en supprime environ un millier début 2024 ; T-Mobile coupe ses équipes agiles dans le sillage de la fusion Sprint. Sur le marché britannique, les missions de coaching en contrat tombent de 127 à 44 entre mi-2023 et mi-2025 — moins soixante-cinq pour cent en deux ans. La même année, le Pulse of the Profession du PMI enregistre pour la première fois un recul du nombre d'organisations utilisant des méthodes agiles.

Un mot sur la provenance de ces exemples, car elle n'est pas neutre. Ils sont tous anglo-saxons, et ce n'est pas un choix. Une entreprise américaine ou britannique qui met fin à sa filière agile publie un communiqué ; en France, la même décision se prend sans annonce, ne figure dans aucun document public, et un plan de réorganisation ne détaille pas les métiers qu'il supprime. Le lecteur français ne doit donc pas lire cette absence comme le signe que le phénomène l'épargne. Les analyses de marché françaises décrivent depuis 2023 la même contraction des postes de facilitation, mais elles reposent sur l'observation de recruteurs et de praticiens, non sur des faits publiés — nous ne pouvons pas les employer au même rang.

Retenons plutôt ce que cette asymétrie dit d'elle-même : **la décision ne laisse pas de trace, y compris au niveau d'un marché entier.** C'est exactement ce que cette section soutient à l'échelle d'une organisation, et nous n'avions pas prévu de le retrouver à l'échelle d'un pays.

A ma modeste échelle, cette observation ma paru pertinente car c'est précisément ce que j'ai observé dans mes activités, et nombre de mes anciens consœurs et confrères ont abandonné les rôles agile et se sont reconvertis depuis 2023.

Vue depuis la mécanique de la section précédente, cette même dissolution a un tout autre sens. Le scrum master qui refusait de clore un sprint sur une définition of done non tenue — entendons-nous, la liste des conditions qu'un incrément doit

remplir pour être déclaré terminé — n'était pas un garant moral, et ce n'est pas son courage personnel qui est en cause ici. Il occupait une position : celle où l'information sur la qualité existe. Et il y disposait d'un droit reconnu, opposable, exercé devant témoins. En supprimant le rôle de SM, on ne supprime pas une dépense. On supprime le cordon d'andon dont la section précédente a fait l'invariant de tous les régimes robustes.

Puis vient le temps qu'aucun plan de réorganisation n'écrit, et c'est le seul qui compte vraiment : **la pression sur le périmètre, elle, ne bouge pas d'un pouce.** Les clients ont les mêmes besoins le lendemain de la dissolution que la veille. Les engagements commerciaux tiennent. Les feuilles de route restent chargées. Rien, dans la suppression d'une fonction d'accompagnement, ne retire une seule ligne de ce qu'il y a à produire.

La gouvernance revient alors au couple coût-délai. Budgets annuels, jalons, comités de pilotage — non par régression idéologique, non parce que quelqu'un aurait décidé de revenir en arrière, mais parce que ce sont les instruments qui restent debout quand les autres ont été démontés. Le périmètre est exigé, le coût et le délai sont verrouillés par la planification financière. Il reste une dimension libre. Une seule. Elle cède.

Aucun des quatre moments que nous venons de décrire ne contient de décision de sacrifier la qualité. Le résultat est pourtant exactement celui-là. Voilà ce qu'il s'agit maintenant d'expliquer.

Pourquoi personne ne le voit

Nous touchons ici au point le plus important de cette section, et nous voudrions y insister plus longuement qu'ailleurs, parce que c'est celui qu'on nous conteste le plus souvent en séance.

Le basculement est silencieux pour une raison précise, et cette raison n'est ni la mauvaise foi, ni l'incompétence, ni le cynisme. **Il n'existe aucune instance où la somme de ces décisions serait vue comme une décision.**

Écoutons les renoncements tels qu'ils se disent réellement, dans le langage des gens qui les prennent. Cette refonte, on la fera au prochain trimestre. Ces tests, on les écrira quand la pression sera retombée. Cette documentation, celui qui a écrit le module la porte dans sa tête, ça ira bien comme ça. Ce composant, on le duplique plutôt que de le factoriser, parce qu'on livre jeudi et qu'on ne va pas casser l'existant à trois jours de la mise en production.

Reconnaissez-vous ces phrases ? Nous les avons toutes prononcées. Chacune est raisonnable. Chacune est prise par quelqu'un de compétent, avec l'information dont il dispose, et le plus souvent avec raison à l'échelle où il décide. Chacune est réversible — en théorie. Et surtout, chacune passe sous le seuil qui déclencherait un compte rendu. Il n'existe pas de formulaire pour « nous n'avons pas refactorisé ce trimestre ». Il n'existe pas de ligne budgétaire pour la testabilité perdue. Il n'existe aucun comité qui reçoive, une fois par an, l'agrégat de ces mille renoncements et qui puisse dire : voilà ce que nous venons de choisir.

L'asymétrie mérite d'être dite crûment, parce qu'elle est le cœur du problème. Un dépassement de budget de trois pour cent remonte en comité de pilotage et il fait des remous. Un retard de deux semaines sur un jalon remonte en comité de pilotage et il commence à y avoir des vagues. Une fonctionnalité retirée du périmètre remonte en comité pilotage — et sera négociée avec le client, parfois âprement, avant l'avis de tempête. Une année entière sans le moindre investissement de robustesse ne remonte nulle part, non pas parce qu'on la cacherait, mais **parce qu'il n'existe nulle part où elle remonterait.**

J'ai accompagné récemment une équipe qui développait en agilité — sur le papier. Il n'y avait aucune automatisation des tests : ni tests unitaires, ni qualification automatisée, et l'intégration continue ne déclenchait aucun test de non-régression. Rien de tout cela n'existait.

Ce n'est pas que l'idée ne leur fût jamais venue. L'un des intervenants me l'a dit lui-même : il avait été recruté pour ça. Son parcours portait des chaînes de qualification automatisées, c'est ce qu'on était venu chercher chez lui, c'est pour cela qu'il était là. Le jour où nous avons eu cette conversation, il était en mission depuis six mois, peut-être un an. Il n'avait pas écrit une ligne d'automatisation. Il passait ses journées à qualifier à la main, il était l'un des rares à le faire, et comme rien n'était automatisé, il avait beaucoup à faire. Il n'avait donc pas le temps d'automatiser.

Arrêtons-nous sur ce cas, parce qu'il dit davantage que ce que nous venons d'écrire. Nous affirmions à l'instant qu'il n'existe pas de ligne budgétaire pour la testabilité perdue. Ici, il y en avait une. Quelqu'un avait identifié le besoin, ouvert un poste, lu des candidatures, retenu le bon profil et engagé la dépense. L'intention était formulée, la compétence était présente, le financement était acquis. Et au bout d'un an, l'organisation se trouvait exactement là où elle aurait été sans rien faire — à ceci près qu'elle avait payé.

Où cette information aurait-elle dû remonter ? Aucun jalon n'avait été manqué : les qualifications étaient faites. Aucun budget n'avait été dépassé. Personne n'avait décidé de renoncer à l'automatisation, et si vous aviez posé la question à cette direction, elle vous aurait répondu de bonne foi qu'un ingénieur y travaillait. Ce qui s'est produit là n'a simplement aucun nom dans les documents de cette organisation : **la marge qu'on avait réservée pour sortir du régime a été intégralement absorbée par le régime lui-même.**

C'est un défaut d'instrumentation, pas un défaut de caractère. Les directions dont nous parlons ici arbitrent avec ce que leur système leur présente comme arbitrage, et il serait malhonnête de leur reprocher de ne pas trancher une question qui ne leur est jamais posée.

La fuite mémoire

Nous avons déjà employé cette image, dans un article précédent, à propos des accompagnants eux-mêmes. Nous la reprenons ici pour le système de production tout entier, et cette répétition est délibérée : c'est le même défaut, à un autre étage.

Une fuite mémoire, en informatique, est un défaut d'une élégance perverse. Un programme réserve de la mémoire, s'en sert, et omet de la rendre. Une fois. Puis encore. Chaque omission est minuscule, aucune ne produit d'erreur, aucune n'apparaît dans les journaux, et le programme continue de tourner parfaitement — des heures, des mois, parfois des années. Il n'y a rien à corriger, puisqu'il n'y a rien qui échoue. Jusqu'au moment où la mémoire disponible s'épuise, et où le programme s'arrête net.

Ce qui nous intéresse dans cette image n'est pas le crash. C'est ce qui vient immédiatement après, et que tout développeur a vécu : **le crash sera attribué à sa dernière cause.** La requête qui tournait à ce moment-là. Le pic de charge du matin. La mise à jour de la veille — et le collègue qui l'a poussée passera un mauvais quart d'heure. Personne ne remontera aux dizaines de milliers d'allocations non libérées qui, individuellement, n'ont jamais rien cassé.

Traduisez maintenant dans votre organisation. Le composant que plus personne n'ose modifier. Le départ d'un développeur qui emporte la seule compréhension d'un module. Le délai de mise en production qui a doublé sans qu'on sache dire à partir de quand. L'incident dont le post-mortem conclut à une erreur humaine — alors qu'il mesure, en réalité, une architecture devenue impossible à tenir dans une seule tête. Chacun de ces événements sera traité comme un événement. Aucun ne sera lu comme le symptôme d'un régime.

Le nom que les directions financières comprendraient

Il nous faut, pour finir, quitter le vocabulaire de l'ingénierie. Ce qui se produit dans ce basculement porte un nom que les directions financières connaissent parfaitement, et qu'elles emploieraient immédiatement si on le leur présentait ainsi.

C'est un transfert de risque : du compte de résultat de l'exercice vers un engagement hors bilan — à ceci près qu'aucune annexe ne le mentionne.

La dissolution des dispositifs produit une économie certaine, datée, attribuable : elle apparaît dans l'exercice en cours, du bon côté de la ligne, et elle figurera dans l'évaluation de ceux qui l'ont décidée. La dégradation de la qualité interne produit un passif tout aussi réel, mais non daté, non attribuable, et — c'est le point décisif — **non comptabilisé**. Aucune norme n'oblige à le provisionner. Aucun auditeur ne le certifie. Aucun commissaire aux comptes n'a jamais refusé de signer parce qu'une entreprise avait cessé de refactoriser.

Cette dernière phrase est presque comique tellement elle est décorrélée de la réalité des entreprises.

Une organisation qui procède ainsi n'a donc pas fait un mauvais calcul. Elle a fait un calcul juste sur un périmètre trop petit, avec les seuls instruments dont elle disposait — et c'est précisément ce qui rend le phénomène si difficile à arrêter : il récompense ceux qui le déclenchent, sur l'horizon où ils sont évalués, et il facture ceux qui viendront après.

Alors une question, avant d'aller chiffrer ce passif dans la section suivante, et nous la posons sans ironie parce que nous n'en connaissons pas la réponse : dans votre organisation, qui recevrait cette information s'il elle existait ? Non pas qui devrait la recevoir — qui, aujourd'hui, siège quelque part où on pourrait la lui présenter ?

§5. Ce que la dette coûte vraiment

Où nous en sommes. Nous avons décrit un mécanisme (§3) et raconté comment il se réinstalle sans que personne le décide (§4). Nous avons dit que la dette technique était la fuite mémoire de l'organisation, et qu'elle finissait par se payer. Il reste à savoir combien — sans quoi tout ceci demeure une inquiétude d'ingénieur.

Ce que nous allons voir ici. Des chiffres, leurs sources, et leur méthode. Cette section est volontairement sèche : pas d'image, pas d'analogie, pas de récit. Nous établirons l'ordre de grandeur du coût de la non-qualité, puis nous montrerons que l'arbitrage qualité contre productivité est un marché de dupes — et qu'il est mesurable comme tel. Nous finirons par le seul langage qui permette de parler de qualité à une direction financière : celui du coût du délai.

L'ordre de grandeur

Commençons par le chiffre le plus large, qui a le défaut de tous les grands agrégats — il ne prouve rien sur votre organisation — et la vertu de fixer une échelle.

Le Consortium for Information & Software Quality établit, dans son rapport de décembre 2022, que le coût de la mauvaise qualité logicielle aux États-Unis s'élève à **2 410 milliards de dollars par an**. Le rapport couvre trois domaines : les pertes liées à la cybercriminalité exploitant des vulnérabilités existantes, les problèmes de chaîne d'approvisionnement logicielle — singulièrement autour des composants libres — et l'effet de la dette technique sur le développement.

Le même rapport chiffre par ailleurs la **dette technique accumulée** à environ 1 520 milliards de dollars. Il faut être précis sur le statut de ce second nombre, car il est régulièrement mal cité : il ne s'agit pas d'une part du premier. Les 2 410 milliards sont un coût annuel, constaté. Les 1 520 milliards sont un stock — le coût qu'il faudrait engager pour reprendre ce qui a été construit de façon sous-optimale. L'un se paie chaque année, l'autre attend.

C'est cette distinction, et non le montant, qui intéresse notre thèse. Elle décrit exactement le transfert dont nous parlions en §4 : un flux annuel qu'on optimise, et un stock qu'on laisse grossir sans qu'aucun exercice comptable ne l'enregistre.

Descendons maintenant d'un cran, vers un chiffre que les directions reconnaissent immédiatement parce qu'il leur est déjà arrivé. Splunk, avec Oxford Economics, a interrogé en 2026 deux mille dirigeants d'entreprises du Global 2000 dans vingt pays. Le coût agrégé de l'indisponibilité non planifiée atteint **600 mil-**

liards de dollars par an, en hausse de 50 % en deux ans. Rapporté à l'unité : **15 000 dollars la minute**. Une entreprise perd en moyenne 95 millions de dollars de chiffre d'affaires annuel du fait de ses pannes, et sa valeur actionnariale recule de 3,4 % après un incident unique.

Retenons surtout la tendance. Le montant a augmenté de moitié en deux ans, dans une période où les investissements en observabilité et en fiabilité n'ont jamais été aussi élevés. Nous ne prétendons pas expliquer cette hausse par notre seule thèse — ce serait précisément le genre de raccourci que nous aimerions éviter. Nous notons simplement que cette tendance est cohérente avec notre thèse, et qu'aucune des explications concurrentes que nous avons lues ne la couvre entièrement.

Le marché de dupes

Venons-en au point qui devrait, à lui seul, clore le débat — et qui ne le clôt jamais, parce qu'il contredit une intuition trop confortable.

L'arbitrage implicite que nous avons décrit tout au long de cette étude repose sur une croyance : en relâchant la qualité, on gagne de la vitesse. C'est ce qui rend le basculement acceptable dans l'esprit de ceux qui le laissent se produire. On sait qu'on abîme quelque chose, mais on croit acheter du délai avec.

Les travaux DORA mesurent cette croyance depuis **dix ans** — l'édition 2024 est la dixième —, sur des dizaines de milliers de répondants, et le résultat est constant : **le débit de livraison et la stabilité ne sont pas arbitrés l'un contre l'autre, ils sont corrélés**. Le rapport 2024 le formule sans détour : « le débit et la stabilité des changements tendent à évoluer ensemble ; nous avons vu des équipes atteindre un niveau élevé sur les deux, dans tous les secteurs d'activité » (p. 79, notre traduction). Les organisations qui livrent le plus souvent sont aussi celles dont les changements échouent le moins, et qui se rétablissent le plus vite. Ce ne sont pas deux populations qui auraient fait deux choix opposés : c'est la même.

Il faut prendre la mesure de ce que cela signifie. Sacrifier la qualité pour gagner en productivité **ne produit pas de productivité**. Ce n'est pas un arbitrage douloureux mais rationnel, où l'on paierait demain un gain obtenu aujourd'hui. C'est un marché de dupes : on paie demain, et on ne gagne pas aujourd'hui — ou si peu, et si brièvement, que la mesure ne le retient pas.

Nous devons ici une précaution, car cette corrélation ne dit pas son sens. Il est possible que les organisations qui livrent souvent soient stables parce qu'elles livrent souvent — les petits lots réduisent le risque de chaque changement. Il est

possible que ce soit l'inverse. Il est possible, et c'est notre lecture, qu'un troisième facteur produise les deux : un régime où la qualité est tenue. DORA ne tranche pas, nous non plus. Mais dans les trois cas, l'idée qu'on gagnerait du débit en relâchant la stabilité est démentie par la donnée statistique.

Parler de qualité sans dire le mot

Reste un problème pratique, et il est le vrai obstacle de terrain : les chiffres qui précèdent ne se transportent pas dans une réunion de direction. Un directeur financier n'a que faire d'un agrégat national, et l'expression « dette technique » a été trop employée par des gens qui ne pouvaient pas la chiffrer.

Le vocabulaire qui fonctionne est celui de Donald Reinertsen, consultant en développement de produits et auteur en 2009 de *The Principles of Product Development Flow*, et il consiste à ne jamais prononcer le mot qualité. Reinertsen raisonne en **coût du délai** : combien coûte, en euros par semaine, le fait qu'une capacité arrive plus tard qu'elle n'aurait pu. Cette question-là se répond, elle se chiffre, et elle est déjà dans le langage de la finance.

Son intérêt pour notre thèse est le suivant. Dès qu'on sait chiffrer le coût du délai, on peut chiffrer ce que coûte l'allongement progressif des cycles de livraison — cet allongement dont personne ne sait dire quand il a commencé, et que nous avons décrit en §4 comme l'un des symptômes qui sera attribué à sa dernière cause. La dette technique cesse alors d'être une préoccupation d'ingénieur pour devenir ce qu'elle est réellement : **une augmentation progressive du coût du délai sur toutes les capacités futures.**

Formulée ainsi, elle entre dans une discussion budgétaire. Formulée en « il faudrait refactoriser », elle n'y entre pas — et n'y est jamais entrée.

C'est le moment de rappeler ce que disait la section 3 : la qualité interne n'a pas le poids suffisant parce qu'elle n'a pas de langage dans lequel s'exprimer au bon niveau. Reinertsen ne crée pas l'instance manquante. Il fournit la langue dans laquelle elle pourrait, un jour, délibérer.

§6. Performance et robustesse

Où nous en sommes. Les sections 3 à 5 ont établi un fait et l'ont chiffré : une dimension dont personne ne répond au rang de l'arbitrage finit par céder, et cela coûte. Mais tout ce qui précède est formulé dans le vocabulaire de l'agilité et de la production logicielle — ce qui expose l'étude à une objection facile, et qu'il vaut mieux traiter que contourner : tout ceci n'est qu'une querelle de méthodes entre praticiens.

Ce que nous allons voir ici. Un cadre qui ne vient pas de ce monde-là. L'arbitrage décrit jusqu'ici est un cas particulier d'un dilemme beaucoup plus général, documenté en biologie, en stratégie d'entreprise et en ingénierie de la sécurité, et qui porte un nom : **performance contre robustesse**. Nous verrons ensuite que la période présente retourne l'intuition commune sur ce qui, des deux, est le luxe.

Palier 1 — Le dilemme, hors du monde logiciel

Olivier Hamant est biologiste — directeur de recherche à l'INRAE, au laboratoire de Reproduction et Développement des Plantes de l'École normale supérieure de Lyon, où il étudie comment les plantes se servent des forces mécaniques pour contrôler leur croissance. Il n'est donc pas un auteur de management qui aurait emprunté une image au vivant : c'est un chercheur du vivant qui constate que le vivant ne fonctionne pas comme nos organisations, et qui s'en étonne. Cette provenance importe pour la suite, et nous dirons plus bas ce qu'elle interdit d'en tirer.

Il pose sa définition dès l'introduction de L'Entreprise Robuste, et elle tient en une phrase que nous pouvons citer telle quelle : « robustesse : maintenir le système stable (à court terme) et viable (à long terme) malgré les fluctuations. » En face, la **performance** est « la somme de l'efficacité et de l'efficience, c'est-à-dire atteindre son objectif avec le moins de moyens possibles ». Deux stratégies, deux régimes d'environnement : la performance gagne dans un monde stable, où l'optimisation ne coûte rien parce que rien ne change ; la robustesse gagne dans un monde fluctuant, où la même optimisation devient le chemin le plus court vers la rupture.

Ce n'est pas une préférence esthétique pour la lenteur ou pour la marge — c'est une lecture du vivant lui-même, chiffres à l'appui. Le rendement photosynthétique est inférieur à 1 % ; 90 % de l'énergie se perd à chaque échelon de la chaîne alimentaire ; les réseaux moléculaires sont saturés de redondances, d'aléas et d'incohérences délibérées (des boucles d'accélération systématiquement doublées de boucles de freinage). « Le vivant se construit donc sur des fuites permanentes », écrit Hamant. Ce qui ressemblerait à un défaut de conception est ce qui permet au vivant de durer : il construit sa robustesse contre la performance, pas malgré elle.

Un mot, avant d'aller plus loin, sur « résilience ». Le terme est trop chargé pour servir de synonyme commode à robustesse, et Hamant ne le traite pas comme une nuance technique — il le traite comme une récupération idéologique. Citant Thierry Ribault, il rappelle que le vocabulaire de la résilience, appliqué aux sociétés, tend à « fataliser le désastre » en valorisant ceux qui se relèvent individuellement d'un choc, pendant qu'on oublie ceux qui ne se relèvent pas — et le tissu collectif qui aurait dû amortir ce choc pour tous. La robustesse, chez Hamant, n'est pas une vertu de rétablissement après coup ; c'est une propriété du système avant la crise.

Mais il faut aller un cran plus loin que la mise en garde de Hamant, jusqu'à la source écologique dont dérive tout ce vocabulaire. Holling, en 1973, sépare rigoureusement deux propriétés que la définition de Hamant fusionne dans un seul mot : la **stabilité** — la vitesse et la précision avec lesquelles un système revient à l'équilibre après une perturbation, peu de fluctuation étant signe de forte stabilité — et la **résilience** — la capacité à absorber le changement sans disparaître, mesurée non par la fluctuation mais par la persistance. Or ces deux propriétés peuvent **s'opposer**. Holling l'illustre par les pêcheries des Grands Lacs : climatiquement tamponnées, homogènes, peu de fluctuation — donc hautement stables — mais si peu résilientes qu'une pression de pêche pourtant relâchée à temps n'a pas empêché leur effondrement. Sa formule tient en une phrase : « poursuivre un rendement soutenu maximal peut produire un système plus stable, mais à résilience réduite. » Optimiser pour la stabilité — ce que Hamant et cette étude appellent déjà performance — ne se contente donc pas de coûter de la robustesse : **il l'use activement, et sans le signaler**, puisque Holling note que le déclin des stocks du lac Michigan n'a montré « aucun signe visible de problème » pendant toute la période où la résilience se défaisait. C'est, quarante ans avant l'ingénierie de la résilience présentée au palier 4, le même argument que Hollnagel : l'absence d'incident n'est pas la preuve d'un système sûr. Ce document retient donc « robustesse » comme le fait Hamant — le mot qui couvre les deux propriétés — mais avec cette réserve : le point 3, plus bas, n'est

pas seulement que la performance est risquée en monde fluctuant ; c'est que la recherche consomme par construction la capacité d'absorber le choc rare, avant même qu'aucun signal ne le montre.

Reste la question de l'analogie, et Hamant lui consacre un chapitre entier avant de se prononcer — un signal qu'il faut prendre au sérieux plutôt que sauter. Sa conclusion est nette, et vaut d'être citée plutôt que résumée : « nous pourrions donc nous méfier de toute analogie béate qui utiliserait le vivant pour le "plaquer" sur les modèles d'entreprise [...]. Cet essai propose plutôt d'utiliser les règles du vivant dans une démarche d'ouverture d'esprit et de stimulation intellectuelle. » Pas de transposition, donc : une analogie assumée et bornée dans le même geste — l'entreprise n'a ni évolution darwinienne, ni l'absence de décision du vivant, et Hamant le dit lui-même sans détour. Nous reprenons ici la même prudence : ce document s'autorise l'éclairage du vivant, pas son mimétisme.

Sur le coût, en revanche, Hamant reste qualitatif. Il parle d'un coût « en termes de performance » pour les fonctions redondantes maintenues au cas où, et renvoie l'arbitrage à un dosage — au décideur de « jauger la dose de lenteur, de redondance ou d'aléas à préserver ». Aucun chiffrage. Ce sera au palier suivant, et à notre propre matériau, de donner à ce coût un ordre de grandeur que le seul cadre biologique ne fournit pas.

Ce dilemme a aussi été formulé dans le monde du conseil, et par quelqu'un qu'on ne peut pas soupçonner d'hostilité à l'efficience.

Palier 2 — Le renfort, et la sortie de l'entre-soi

Trois auteurs portent une idée chacun et nous ne les avons pas invoqué pour faire nombre : chacun ferme une porte de sortie différente au lecteur qui voudrait ranger ce document du côté des "préférences méthodologiques".

Roger Martin, stratège et ancien doyen de la Rotman School of Management de Toronto — un homme du conseil, et c'est ce qui rend son constat gênant pour son propre monde —, publie en 2019 dans la Harvard Business Review un article au titre sans ambiguïté : « The High Price of Efficiency ». L'appui compte moins pour son contenu que pour sa provenance : il vient exactement du monde qui a promu l'efficience comme finalité, et il est publié par la revue qui l'a diffusée. Son argument est que la poursuite de l'efficience, au-delà d'un certain point, cesse de produire de la valeur et commence à

produire du **risque systémique** — une organisation optimisée n'est pas seulement plus performante, elle est plus étroitement couplée, et un incident local s'y propage au lieu de s'y amortir.

Nassim Nicholas Taleb, ancien opérateur de marché devenu essayiste du risque, apporte la deuxième idée, et de poids. Nous ne mobilisons ni son vocabulaire complet ni sa posture : une seule proposition, qui suffit et qui est la bonne. **L'optimisation n'est pas un état neutre — elle est elle-même un facteur de fragilisation.** Retirer les marges d'un système ne le laisse pas identique en plus économe : cela change sa réponse aux perturbations. C'est la formulation la plus concise du coût caché que cette étude documente depuis cinq sections.

Tom DeMarco, enfin, fait le pont le plus direct avec les directions informatiques. Il vient du génie logiciel, où il avait déjà signé avec Timothy Lister Peopleware, le livre qui posait que les échecs de projet sont sociologiques avant d'être techniques. Dans Slack (2001), il pose que le mou n'est pas du gaspillage résiduel qu'une organisation efficace aurait éliminé : c'est **la capacité d'adaptation elle-même**. Sa conséquence est d'une brutalité utile en comité de direction : une organisation dont toutes les ressources sont affectées à 100 % n'est pas au sommet de son efficacité, **elle est dans l'incapacité de changer**. Elle peut encore exécuter ; elle ne peut plus bifurquer.

Ces trois appuis convergent, et ils viennent de trois mondes qui ne se lisent pas mutuellement — le conseil stratégique, la finance des événements rares, l'ingénierie logicielle. C'est cette convergence qui compte : elle indique que nous ne décrivons pas une particularité de notre secteur.

Reste à savoir dans quel environnement nous nous trouvons réellement.

Palier 3 — Le retournement

Tout ce qui précède tient en deux propositions, et comme vous allez le constater, leur mise en contact produit l'argument central de cette étude.

La première, déjà posée au palier 1 : la robustesse est la stratégie de survie d'un environnement fluctuant, la performance celle d'un environnement stable — Hamant le montre sur le vivant, ce n'est pas une préférence mais une constatation sur ce qui persiste.

La seconde : la période 2024-2026 est l'environnement le plus instable qu'aient connu les directions informatiques depuis des décennies. Rupture technologique majeure et non stabilisée, contrainte économique durable, recomposition brutale des équipes, incertitude réglementaire sur l'IA, chaînes d'approvisionnement logicielles sous tension.

Et c'est exactement le moment qui a été choisi pour liquider les marges.

Il faut mesurer ce que cette rencontre implique, parce qu'elle renverse une intuition partagée par à peu près tout le monde, y compris par ceux qui prennent les décisions de bonne foi. L'intuition dit : quand la conjoncture se durcit, on resserre ; l'investissement de qualité, la redondance, le temps de revue sont ce qu'on peut se permettre quand ça va bien.

Or c'est l'inverse qui est vrai : **La performance est le luxe, pas la robustesse.**

Une organisation peut se permettre d'être optimisée et étroitement couplée tant que son environnement ne bouge pas — les surprises y sont rares, et le rendement supérieur les paie largement. Dès que l'environnement devient imprévisible, cette même optimisation devient le comportement le plus risqué disponible : elle a supprimé précisément ce qui sert à absorber ce qu'on n'avait pas prévu. Resserrer en période instable, ce n'est pas de la prudence — c'est augmenter simultanément la probabilité du choc et son coût.

Ceux qui viennent de l'exploitation reconnaîtront la mécanique sous un autre nom. Un système que l'on charge à 95 % de sa capacité n'est pas « efficace à 95 % » : il est dans la zone où le temps d'attente explose de façon non linéaire, et où la moindre variation de la demande produit une file qui ne se résorbe plus. Personne, en informatique de production, ne dimensionne délibérément à 95 %. **C'est pourtant le dimensionnement que l'on impose depuis trois ans à des organisations entières**, et sans que quiconque ait eu à décider que c'était la cible.

Cette section ne défend donc aucun état antérieur, et c'est important pour lire correctement tout le document : il ne s'agit pas de dire que les transformations agiles étaient bonnes et qu'on a eu tort de les arrêter. Il s'agit d'un constat de calendrier — **la période présente demande plus de robustesse qu'avant, pas moins**, et elle en a reçu moins.

Palier 4 — Fermeture, et ce que la section suivante devra porter

Deux corpus, brièvement, parce qu'ils fournissent les conditions dont la section 9 fera son matériau.

L'ingénierie de la résilience — Hollnagel et Woods — pose que la sécurité n'est pas l'absence d'incident, mais une capacité d'adaptation. La conséquence pour notre thèse est directe, et elle est inconfortable : une organisation qui n'a plus d'incidents n'est pas nécessairement devenue plus sûre. Elle peut simplement avoir cessé de regarder. C'est le même mécanisme que celui de la section 3, énoncé par une autre discipline : ce qui n'est plus observé cesse d'exister pour le système de pilotage, et cesser d'exister au niveau pilotage n'est pas la même chose que cesser de se produire.

Weick et Sutcliffe, à partir de l'étude des organisations à haute fiabilité, en retiennent deux principes — nous ne déroulons pas les cinq, seuls ces deux préparent la suite. La **préoccupation pour les défaillances** : traiter le petit signal comme une information et non comme du bruit, y compris quand il ne s'est rien passé. Et la **déférence à l'expertise** : en situation critique, la décision migre vers celui qui sait, indépendamment de son rang.

Ce second principe mérite qu'on s'arrête, parce qu'il ferme la boucle de ce document. **Weick observe en organisation ce qu'Ashby démontre en théorie** (§3). Là où Ashby énonce qu'un système de contrôle doit posséder au moins autant de variété que ce qu'il régule — donc que la décision doit se prendre là où l'information se trouve —, Weick constate que les organisations qui échouent le moins souvent sont précisément celles qui ont organisé cette migration. Le théorème et l'observation de terrain disent la même chose à soixante ans d'écart, et par des chemins qui ne se sont jamais croisés.

Ce cadre existe depuis des décennies. Il est enseigné, documenté, publié dans les revues que lisent les dirigeants. Reste à comprendre pourquoi il ne s'installe jamais durablement — et c'est l'objet de la section suivante.

§7. Le motif du dévoiement

Où nous en sommes. La section précédente a posé le cadre conceptuel : performance et robustesse ne sont pas deux degrés d'une même vertu, mais deux régimes distincts, et un système optimisé pour l'une perd nécessairement de l'autre. Nous disposons désormais du diagnostic (§3 à §5) et du vocabulaire pour le penser (§6).

Ce que nous allons voir ici. Que ce qui est arrivé à l'agilité ne lui est pas arrivé à elle seule. Nous examinerons trois méthodes de management qui n'ont aucun lien de filiation — le lean, l'agilité, le change management — et nous montrerons qu'elles ont subi exactement la même trajectoire, dans le même ordre, pour la même raison. Nous nommerons ensuite le mécanisme, qui porte un nom en économie depuis cinquante ans. Cette section est celle qui fait passer l'étude d'une critique sectorielle à une loi générale ; elle est aussi celle où nous prenons le plus grand soin de n'accuser personne.

Une histoire que vous avez déjà entendue

Commençons par Toyota, puisque tout le monde y commence.

Le jidoka est un dispositif d'une radicalité qu'on a peine à imaginer aujourd'hui : n'importe quel opérateur, à n'importe quel poste, peut arrêter la chaîne de production entière. Pas signaler un défaut. Pas remplir une fiche. Arrêter ! Une usine à plusieurs milliers de personnes s'immobilise parce qu'un opérateur a jugé que quelque chose n'allait pas — et le système est conçu pour que ce jugement soit respecté, y compris quand il se révèle erroné, parce qu'une organisation qui punit les arrêts injustifiés n'obtient plus d'arrêts du tout.

Vous connaissez la suite, et vous l'avez peut-être vécue. Le lean traverse l'Atlantique puis la Manche, il est adopté massivement, il est enseigné, certifié, décliné en programmes d'entreprise. Et ce qui en arrive dans la plupart des organisations occidentales, trente ans plus tard, ce n'est pas le cordon. C'est un tableau de bord. Des indicateurs de gaspillage, des taux de rendement synthétique, des chantiers d'amélioration continue comptabilisés au trimestre. Le vocabulaire est intact — on dit kaizen, on dit muda, on dessine des VSM. Le dispositif qui donnait à l'opérateur le pouvoir d'arrêter la ligne, lui, n'a pas fait le voyage.

Je dois ici un témoignage plus ancien, et plus personnel, parce qu'il a décidé de la suite de ma vie professionnelle.

Il y a une vingtaine d'années, j'étais en prestation dans une équipe, et le lean y a été déployé — massivement, sur toutes les divisions, la nôtre comprise. La pratique qu'on nous a imposée tenait en peu de chose : chaque matin, l'équipe se retrouvait debout devant un tableau, et chacun répondait à trois questions. Qu'ai-je fait hier ? Que vais-je faire aujourd'hui ? **Ai-je besoin d'aide ?**

Arrêtons-nous sur la troisième, car c'est un cordon. C'en est même un remarquablement bien placé : il est tendu tous les jours, à voix haute, au moment où l'information est fraîche, et il ne demande aucune démarche. Il suffit de répondre oui.

Une personne de l'équipe parlait peu. Elle disait ce qu'elle avait à faire, et c'était tout. Un matin, après le rituel, je l'ai trouvée à la machine à café ; elle n'allait pas bien. Nous avons parlé. Elle m'a dit qu'elle ne s'en sortait pas avec une macro Excel, et qu'elle ne pouvait pas le dire devant les autres — pas cela, pas devant tout le monde, précisément parce que le sujet était trivial. Le lendemain, elle était en arrêt maladie pour deux semaines.

Je ne rapporte pas cela pour désigner un coupable, et il n'y en a pas. Personne dans cette équipe n'aurait voulu ce qui s'est passé, et le rituel avait été installé par des gens convaincus de bien faire. Je le rapporte parce que le dispositif a fonctionné exactement comme prévu, et qu'il a produit l'inverse de son objet. La question a été posée. Elle a été posée chaque matin, à quelqu'un qui avait besoin d'aide. Ce qui manquait n'était pas la question : c'était ce qui permet d'y répondre. Et cela ne s'installe pas, cela ne se déploie pas sur toutes les divisions, cela ne figure dans aucun plan.

Ce que Toyota avait compris et que le voyage a laissé à quai, ce n'est donc pas le tableau. C'est qu'une organisation qui punit les arrêts injustifiés n'obtient plus d'arrêts du tout — et qu'une organisation où l'on perd la face en demandant de l'aide n'obtient plus de demandes d'aide.

C'est ce matin-là que j'ai décidé de devenir coach.

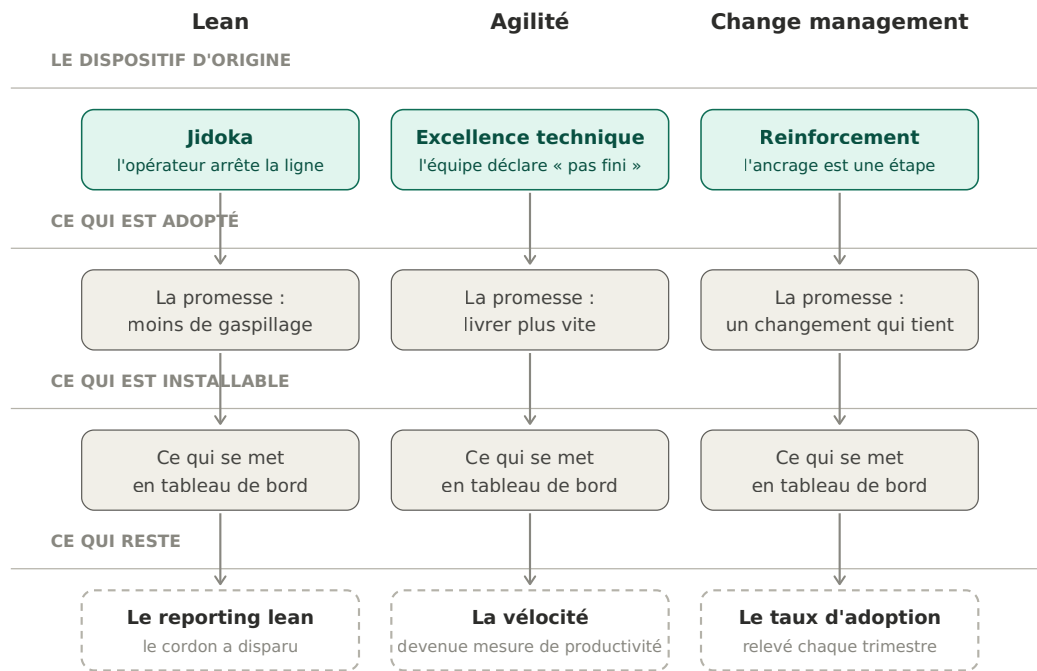
Prenons maintenant l'agilité, dont cette étude a parlé jusqu'ici. Au commencement, une exigence : l'excellence technique, la définition of done, le droit de l'équipe de déclarer que ce n'est pas fini. Trente ans plus tard, ce qui subsiste dans le plus grand nombre d'organisations tient en un mot, et c'est la vélocité — devenue une mesure de productivité d'équipe, comparée d'un sprint à l'autre et parfois d'une équipe à l'autre, alors qu'elle avait été conçue comme un outil de prévision interne et rien d'autre.

Et prenons enfin le change management, qui est la génération la plus récente et sur laquelle nous voulons être particulièrement précis. Le modèle ADKAR compte cinq étapes, et la dernière — le Renforcement, l'ancrage durable — n'est ni un supplément ni une bonne intention : c'est une étape à part entière, sans laquelle le modèle ne prétend pas fonctionner. Elle exige du temps après le déploiement, de la présence, une observation de ce qui se passe réellement une fois que les consultants sont partis. Ce qui remonte aujourd'hui dans les comités, dans l'écrasante majorité des programmes de transformation, ce sont des taux d'adoption trimestriels.

Trois méthodes. Trois décennies. Aucune n'a inspiré les autres. Et pourtant, si vous les alignez, la trajectoire est la même au détail près.

Figure 7.1 — Le motif du dévoiement

Trois générations, aucune filiation entre elles, et la même trajectoire en quatre temps.



Chaque fois, la moitié qui mesure survit ; la moitié qui rend la décision disparaît. Le nom, lui, reste.

Figure 7.1 — Trois générations sans filiation, et la même trajectoire en quatre temps. Ce n'est pas une analogie : c'est le même mécanisme qui opère trois fois.

Ce qui se conserve, et ce qui s'évapore

Regardons le motif de plus près, parce que c'est dans son détail qu'il devient utile.

Chacune de ces méthodes a été conçue **autour** d'un dispositif de qualité — pas avec, pas en plus : autour. Le jidoka n'est pas une pratique du lean parmi d'autres, c'en est le pilier. L'excellence technique n'est pas une option de l'agilité, c'est ce qui rend l'itération soutenable. Le Renforcement n'est pas la cinquième étape d'ADKAR au sens où elle viendrait après les choses sérieuses, c'est celle qui décide si les quatre premières auront servi à quelque chose.

Chacune a ensuite été adoptée pour ses **promesses** — moins de gaspillage, livrer plus vite, un changement qui tient — et ces promesses sont réelles. Aucune de ces adoptions n'a été un malentendu.

Chacune a enfin été **industrialisée** : formalisée, certifiée, outillée, vendue à grande échelle. C'est ici que se produit le tri, et il faut décrire ce tri exactement, parce que c'est le cœur de cette section.

Quand une méthode s'industrialise, elle ne se transmet pas en entier. Elle se transmet par ce qui est **transmissible** — c'est-à-dire par ce qui se met en tableau de bord, se compare entre entités, se présente en comité, et s'inscrit dans un système d'information. Un indicateur voyage admirablement. Un droit de veto, non. Un taux de rendement se consolide sur quarante sites ; le pouvoir d'arrêter la ligne ne se consolide pas, parce qu'il ne se mesure pas — il s'exerce, ou il n'existe pas.

Le résultat est toujours le même, et nous l'avons déjà rencontré en §3 à propos de Drucker : **la moitié qui mesure survit, la moitié qui rend la décision disparaît, et le nom complet reste**. Personne n'a décidé cette amputation. C'était simplement la seule moitié installable sans que le sommet cède quoi que ce soit — et une organisation qui adopte une méthode choisit toujours, sans y penser, la version qui ne lui coûte pas de pouvoir.

Le nom du mécanisme

Ce mécanisme a un nom, et il vient de l'économie. Charles Goodhart, en 1975, observant les autorités monétaires britanniques, constate qu'une régularité statistique s'effondre dès qu'on entreprend de s'en servir comme cible de politique. La formule ramassée que tout le monde connaît — lorsqu'une mesure devient un objectif, elle cesse d'être une bonne mesure — n'est d'ailleurs pas de lui : c'est la reformulation qu'en a donnée l'anthropologue Marilyn Strathern en 1997, en observant l'audit du système universitaire britannique. Nous la retenons parce qu'elle est plus juste que l'originale pour notre objet, mais il faut rendre à chacun ce qui lui revient. La loi a circulé jusqu'à devenir un adage de comptoir chez les

praticiens de la donnée, et elle mérite mieux que ce sort — car ce qu'elle décrit n'est pas une bêtise administrative, c'est une propriété structurelle de tout système où l'on est évalué.

V. F. Ridgway l'avait d'ailleurs formulée près de vingt ans avant Goodhart, dès 1956, et dans des termes bien plus proches de notre sujet puisqu'il parlait d'organisations et non de politique monétaire : les mesures de performance produisent des conséquences dysfonctionnelles — et elles les produisent d'autant plus sûrement qu'elles sont mieux conçues.

Appliquons-la ici. Le jidoka, la définition of done, le Reinforcement étaient des **dispositifs de décision**. Le taux de rendement, la vitesse, le taux d'adoption sont des **mesures**. Le passage de l'un à l'autre n'est pas une dégradation progressive, c'est une substitution : l'indicateur qui accompagnait le dispositif lui survit, puis prend sa place, puis finit par le désigner. On dit « faire du lean » pour dire « suivre des indicateurs lean ». On dit « être agile » pour dire « tenir sa vitesse ». Et à la fin, ce qui est piloté est l'indicateur, pendant que la chose qu'il mesurait a cessé d'exister sans que rien ne le signale.

Nous avons annoncé cette idée en §3, en refusant de la déployer. La voici à sa place. Et remarquez ce qu'elle nous apprend en retour sur l'objection des indicateurs, celle du praticien compétent qui demandait s'il ne suffirait pas de mieux mesurer la qualité. Nous y avons répondu qu'un indicateur transmet un état et jamais un arbitrage. Il faut maintenant y ajouter ceci, qui est plus dérangeant : **un indicateur de qualité qui remonterait effectivement en comité de direction cesserait, par ce mouvement même, de mesurer la qualité**. Il deviendrait une cible, il serait tenu, et il serait tenu de la façon la moins coûteuse — ce que chaque équipe de développement du monde sait faire avec une couverture de tests.

Laissez-moi décrire ce que je fais quand j'arrive dans une équipe, parce que ce geste est devenu, avec les années, mon meilleur instrument de diagnostic.

Je demande à voir les artefacts, et je m'arrête sur la vitesse. Je ne demande pas combien elle vaut : cela ne m'apprendrait rien. Je demande comment elle est définie. Les story points sont-ils utilisés ? Existe-t-il une référence ? Combien vaut le 1 ? Quelle est la user story qui sert de maître étalon, celle à laquelle on compare toutes les autres ?

Il est très rare que je trouve ces réponses. La mesure est là, elle est produite à chaque itération, elle est reportée quelque part — et l'étalon qui lui donnerait un sens a disparu, si tant est qu'il ait jamais existé. Mesurons bien ce que cela veut dire, car ce n'est pas une question de rigueur : une vitesse sans étalon n'est pas une mesure imprécise, c'est un nombre.

Ce que je rencontre en revanche, souvent, c'est un responsable qui protège. Quelqu'un qui s'arrange pour que la vitesse de son équipe ne sorte pas de son équipe. Ses raisons sont excellentes : il est arrivé trop souvent que des équipes soient comparées sur cette base. La vitesse est pourtant une mesure strictement locale — jamais deux équipes ne font la même chose, et un point chez l'une n'a aucun rapport avec un point chez l'autre. Mais vue de loin, dans un tableau, une colonne se compare. De bonne foi, quelqu'un constate que la productivité — traduction : la vitesse — paraît meilleure d'un côté que de l'autre. Et cela redescend en injonction : allez voir comment fait l'autre équipe, faites mieux, il faudrait peut-être recruter autrement. L'ambiance de l'équipe ne s'en remet pas toujours.

Voilà pourquoi les Scrum Masters les plus expérimentés que j'ai rencontrés refusent de diffuser leur vitesse. Certains vont plus loin : ils communiquent une estimation en jours-homme dont ils savent qu'elle est fautive, et ils cessent purement et simplement de regarder la vitesse. Pas de burndown affiché. Aucune amélioration continue fondée dessus.

Arrêtons-nous sur ce qu'ils font, parce que c'est un cran au-delà de Goodhart. Ils ont compris que leur mesure serait capturée, et leur seule défense est d'y renoncer. Ils ne sauvent pas la vitesse de la comparaison : ils la sacrifient pour sauver leur équipe. Et ils perdent au passage précisément ce que l'outil devait leur apporter — la capacité de voir leur propre régularité et de travailler dessus. Le dévoiement ne se contente donc pas de vider la méthode de sa substance chez ceux qui s'y soumettent. **Il coûte aussi son outil à celui qui lui résiste.**

Rendre les méthodes à elles-mêmes

Il faut dire ici, très clairement, ce que cette étude ne soutient pas.

Elle ne soutient pas que le lean serait une méthode taylorienne déguisée. Elle ne soutient pas que l'agilité aurait échoué. Et elle ne soutient surtout pas que le change management serait un habillage de la contrainte — nous connaissons trop bien ce corpus pour le caricaturer, et le modèle ADKAR contient précisé-

ment, dans son étape finale, ce dont l'absence est ici diagnostiquée. Ces trois méthodes ont été **dévoyées, non fautives**. Le dévoiement est un sort qu'elles ont subi, pas un vice qu'elles portaient.

C'est même l'inverse d'une critique. Ces trois corpus avaient tous compris ce que nous essayons de redire ici — que la qualité ne se maintient que si quelqu'un, quelque part, a le pouvoir de dire non. Ils l'avaient compris avant nous, chacun dans son vocabulaire, et chacun l'avait inscrit dans son dispositif central. Ce que l'étude propose n'est donc pas de les remplacer par une méthode de plus. C'est de les rendre à eux-mêmes.

Et ce n'est pas une position de confort, car elle a une conséquence immédiate et inconfortable : **le mécanisme est en train d'opérer une quatrième fois, sous nos yeux**. L'accompagnement de l'adoption de l'IA générative est aujourd'hui à l'endroit exact où se trouvaient le lean en 1995 et l'agilité en 2012. Les dispositifs sérieux existent déjà — revue humaine du code produit, critères d'acceptation d'un usage, formation au jugement plutôt qu'à l'outil. Et l'indicateur qui les accompagne existe aussi, il est déjà dans les comités, et vous l'avez sans doute vu passer : le taux d'adoption des licences.

Lequel des deux, à votre avis, sera encore là dans cinq ans ?

Nous poserons donc la question qui donne son titre à cette section, et nous la posons pour de bon, sans savoir y répondre : combien de fois faudra-t-il ne pas tirer la leçon ? Non pas combien de fois faudra-t-il l'apprendre — elle a été apprise trois fois, écrite trois fois, enseignée trois fois. Combien de fois faudra-t-il la perdre au même endroit, de la même façon, en gardant le nom ?

§8. L'IA générative, amplificateur de régime

Où nous en sommes. Nous avons établi le diagnostic (une variable d'ajustement sans titulaire, un basculement silencieux, une dette qui se chiffre), le cadre qui permet de le penser (performance contre robustesse), et le motif qui explique pourquoi la leçon se perd à chaque génération. Tout cela est rétrospectif : nous avons parlé de ce qui s'est produit entre 2015 et 2025.

Ce que nous allons voir ici. Ce que la période qui s'ouvre va faire de ce diagnostic. Nous examinerons ce que les données disponibles disent réellement de l'effet de l'IA générative sur la production logicielle — en distinguant soigneusement ce qui est mesuré de ce qui est déduit, car cette section est celle où il serait le plus facile, et le plus coûteux, de surinterpréter. Nous en tirerons une seule conclusion, mais elle change la portée de toute l'étude : l'IA générative ne choisit pas de régime, elle amplifie celui qu'elle trouve.

Le paradoxe que personne n'avait prévu dans ce sens

Il faut commencer par une donnée, et elle est inconfortable pour tout le monde — pour les enthousiastes comme pour les sceptiques.

L'adoption de l'IA générative produit **le plus fort gain d'efficacité individuelle jamais mesuré** dans notre industrie. Ce n'est pas une concession de politesse avant la critique, c'est le résultat, et il est solide. Les développeurs qui utilisent ces outils produisent davantage, plus vite, et le rapportent eux-mêmes.

Et le rapport DORA 2024 observe, dans le même mouvement, une **baisse de la stabilité de livraison** au niveau de l'équipe. Le chiffre mérite d'être lu lentement : pour chaque hausse de 25 % de l'adoption de l'IA, DORA estime une baisse de **7,2 % de la stabilité de livraison** et de **1,5 % du débit de livraison**. Non seulement le débit collectif ne monte pas — il baisse légèrement — mais **la stabilité baisse cinq fois plus vite que lui**.

C'est le résultat le plus important de cette section, et il est contre-intuitif au point que les auteurs le signalent eux-mêmes comme contraire à leurs attentes.

Lisez ces deux phrases ensemble. Le gain est individuel et il est réel. La perte est collective et elle est réelle aussi. Ce n'est pas une contradiction dans les données : c'est le signe que quelque chose se passe **entre** l'individu et l'équipe, dans cet espace où le travail produit par l'un devient le système que tous doivent maintenir.

Ce quelque chose porte un nom que cette étude a déjà employé : c'est la capacité d'absorption en aval. Revue, tests, observabilité, temps de compréhension. Le gain individuel ne se convertit en gain collectif que s'il traverse cet étage — et si l'étage n'a pas été dimensionné pour le débit nouveau, il ne le convertit pas. Il l'accumule.

Ce que montrent les mesures de code

Deux jeux de données viennent éclairer ce que devient concrètement ce code accumulé, et il faut les prendre pour ce qu'ils sont : de la mesure massive et automatisée, sur du réel.

GitClear a analysé 211 millions de lignes de code modifiées entre 2020 et 2024. Quatre indicateurs y bougent dans le même sens et sur la même période. La fréquence des **blocs de cinq lignes dupliquées ou plus** est multipliée par huit sur la seule année 2024. Le **copié-collé** passe de 8,3 % à 12,3 % des changements, et dépasse pour la première fois les lignes simplement déplacées — renversement significatif, car déplacer du code est le geste de quelqu'un qui organise, le copier celui de quelqu'un qui ajoute. Les lignes **déplacées**, que le rapport tient pour un indice d'activité de refactoring sans prétendre la mesurer directement, s'effondrent de 24,1 % des changements en 2020 à 9,5 % en 2024. Et le **churn** — la proportion de code réécrit peu après avoir été écrit — passe de 3,1 % à 5,7 %.

De son côté, OX Security a examiné plus de 300 dépôts open source et y a cherché dix anti-patterns d'ingénierie. Le résultat n'est pas qu'un défaut unique serait omniprésent, mais que **les plus fréquents le sont presque toujours** : les commentaires superflus dans 90 à 100 % du code généré, l'évitement du refactoring, la sur-spécification et l'application littérale des règles dans 80 à 90 %. À l'autre bout de l'échelle, les défauts les plus spécifiques — faux bugs, couverture de test factice — ne dépassent pas 20 à 50 %. La distribution compte autant que les chiffres : **ce qui est systématique, ce sont les défauts de jugement, pas les défauts de fonctionnement.**

Le rapport résume cet écart par une image que nous reprenons parce qu'elle est marquante : l'IA se comporte comme une armée de juniors — rapide, compétente, produisant du code qui marche, et dépourvue du jugement architectural et de la conscience de sécurité qui distinguent un développeur expérimenté.

Arrêtons-nous sur cette formule, parce qu'elle dit exactement la bonne chose. Il ne s'agit pas de code mauvais. Il fonctionne — c'est même précisément ce qui le rend difficile à refuser en revue. Ce qui lui manque est d'un autre ordre : la déci-

sion de savoir si ce composant devait exister, s'il fallait le dupliquer ou le factoriser, ce qu'il coûtera à maintenir dans dix-huit mois. C'est-à-dire, très exactement, l'arbitrage dont nous disions en §3 qu'aucun indicateur ne le transmet.

Ce que ces chiffres ne prouvent pas

Nous devons maintenant poser nous-mêmes les limites de ce que nous venons de citer, et nous préférons le faire avant qu'on nous les oppose — c'est une règle que cette étude s'applique partout :

Ces données sont corrélationnelles, pas causales. GitClear le reconnaît explicitement, et il faut lui en savoir gré. Le point mérite d'être dit dans toute sa force, parce qu'il est plus radical qu'il n'y paraît : **le rapport ne compare à aucun moment du code écrit avec ces outils à du code écrit sans eux.** Il mesure l'évolution de tout son corpus dans le temps et la met en regard de la courbe d'adoption. Il n'existe donc aucun groupe de contrôle : personne n'a fait travailler deux populations comparables sur les mêmes projets, l'une avec ces outils et l'autre sans, pendant quatre ans. Nous ne pouvons pas écrire que le code généré par l'IA duplique huit fois plus ; nous écrivons que la duplication a été multipliée par huit pendant l'année où ces outils se sont généralisés. La période 2020-2024 a par ailleurs connu bien d'autres bouleversements — le travail à distance généralisé, une vague de départs puis une vague de licenciements dans la tech, une pression économique inédite sur les délais. Chacun de ces facteurs pourrait à lui seul dégrader les mêmes indicateurs.

Nous ne présentons donc pas ces chiffres comme une preuve. Nous les présentons comme un **faisceau convergent** avec les observations DORA, qui portent sur d'autres populations, avec d'autres méthodes et d'autres indicateurs. Quand deux appareils de mesure indépendants pointent dans la même direction, cela ne démontre pas la cause, mais cela rend l'hypothèse inverse difficile à soutenir.

Un lecteur exigeant nous dira qu'il aurait fallu davantage. Il aura raison, et nous l'écrirons encore au chapitre des limites : sur ce point précis, l'étude expose une conviction adossée à des faits convergents, pas une démonstration.

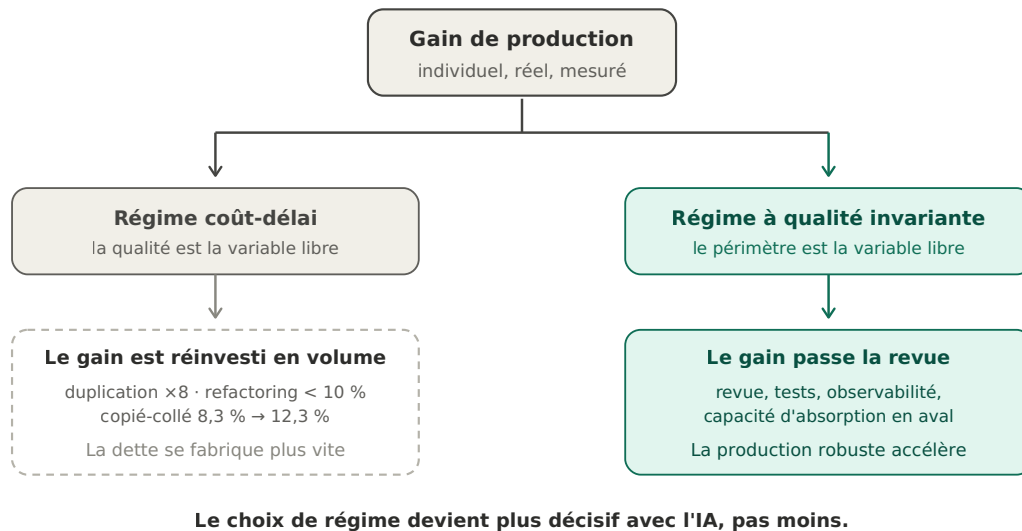
L'IA n'a pas de direction

Venons-en à la seule conclusion que nous tirons de tout ceci, et elle tient en une phrase : **l'IA générative est un amplificateur de régime.**

Elle n'a pas de direction propre. Elle ne rend pas une organisation meilleure ou pire ; elle augmente le débit de ce que cette organisation faisait déjà.

Figure 8.1 — L'IA générative n'est pas une direction, c'est un gain

Le même gain de productivité individuelle, injecté dans deux régimes, produit deux résultats opposés.



Chiffres : GitClear, 211 M de lignes, 2020-2024 — corrélationnels, sans groupe contrôle. Voir la précaution méthodologique dans le texte.

Figure 8.1 — Le même gain individuel, injecté dans deux régimes, produit deux résultats opposés. Ce n'est pas l'outil qui décide, c'est ce qu'il rencontre en aval.

Dans une organisation où la qualité est restée l'invariant — où la revue est réelle, où les tests conditionnent la livraison, où quelqu'un peut encore dire que ce n'est pas fini — le gain individuel traverse l'étage d'absorption et devient un gain collectif. La production robuste accélère. C'est le cas favorable, il existe, et nous en connaissons.

Nous connaissons de tels contextes, et ils ont un trait commun qui mérite d'être dit, parce qu'il est contre-intuitif : **l'outil n'y est pas entré par le développement**. Il est arrivé par le haut de la chaîne — l'écriture d'un dossier d'initiative, le découpage d'une fonctionnalité en histoires, la formulation des critères d'acceptation. C'est-à-dire aux endroits où l'on clarifie une intention, et non à celui où l'on produit des lignes.

Le second trait tient à ce que ces organisations n'avaient pas démonté. Leurs temps de respiration étaient encore là — démonstrations, rétrospectives, itérations sans engagement de livraison où l'on peut explorer, prototyper, se tromper. Le gain individuel y a rencontré un aval capable de l'absorber autrement qu'en volume, parce qu'il subsistait des endroits où mettre autre chose que de la production.

Nous nous garderons d'en tirer une loi : ce sont des observations, pas une étude. Mais elles vont dans le même sens que tout ce qui précède — ce n'est pas l'outil qui décide, c'est ce qu'il rencontre en aval.

Dans une organisation qui a basculé vers le régime coût-délai — celle de la section 4, celle qui a démonté son cordon sans que rien ne se passe — le même gain rencontre un aval qui n'a pas été dimensionné, aucune instance qui puisse le refuser, et une pression de périmètre inchangée. Il se convertit alors en volume. Ce qui s'industrialise, dans ce cas, ce n'est pas la production : c'est la fabrication de dette.

Il faut mesurer ce que cela change. Beaucoup de dirigeants attendent aujourd'hui de l'IA qu'elle règle un problème de productivité dont ils sentent qu'il s'est aggravé — sans toujours savoir dire pourquoi, et cette étude propose une réponse à ce pourquoi. Or si notre lecture est juste, l'IA ne réglera rien de tel dans une organisation qui a perdu son invariant. Elle y accélérera précisément ce qui l'a dégradée.

Et c'est ici que nous voudrions retourner l'objection la plus fréquente qu'on nous fait sur cette étude — celle qui consiste à dire que ces débats de régime appartiennent au passé, que l'IA rebat les cartes et rend caduques les querelles de méthode des années 2010.

C'est l'inverse : si l'IA amplifie le régime qu'elle trouve, alors **le choix de régime devient plus décisif avec elle, pas moins**. Il produisait autrefois un écart lent, sur plusieurs années, que la plupart des organisations pouvaient absorber sans le voir. Il produit maintenant cet écart au débit de la machine. Deux entreprises comparables, l'une ayant gardé son invariant et l'autre l'ayant perdu, adopteront les mêmes outils la même année — et divergeront beaucoup plus vite qu'elles n'auraient divergé sans ces outils.

Disons-le sans détour, car la lecture inverse serait facile et nous ne la soutenons pas : **rien de ce qui précède ne dit que ces outils seraient incompatibles avec un régime lean-agile**. C'est même le contraire que nous observons. Une organisation qui a gardé son invariant est celle qui en tire le plus, et qui en tire autre chose que du volume — parce qu'elle dispose encore des endroits où un gain peut aller se loger. Ce que nous décrivons n'est pas un danger de l'outil, c'est une propriété de l'aval. Le reproche adressé ici ne s'adresse pas à l'IA générative, qui n'a pas d'intention ; il s'adresse à l'idée qu'on pourrait s'en remettre à elle pour réparer ce qu'on a démonté.

Alors la question que nous laissons ouverte n'est pas de savoir si vous devez adopter ces outils : vous les avez déjà, et vos équipes s'en servent, avec ou sans vous. Elle est de savoir ce qu'ils vont trouver en aval. Chez vous, aujourd'hui,

qu'est-ce qui se passe quand un développeur soumet en revue trois cents lignes de code qui fonctionnent parfaitement et que personne n'a le temps de lire ? Parce qu'il soumet à nouveau trois cent lignes de code une heure plus tard.

C'est cette question, et non la précédente, qui décidera de ce que l'IA vous aura coûté.

§9. À quoi ressemble une organisation robuste

Où nous en sommes. Le diagnostic est complet. Nous savons quelle dimension cède (§3), comment elle se remet à céder sans que personne le décide (§4), ce que cela coûte (§5), dans quel cadre le penser (§6), pourquoi la leçon se perd à chaque génération (§7), et ce que l'IA générative va en faire (§8).

Ce que nous allons voir ici. Ce qui distingue les organisations où cela ne s'est pas produit. Nous en proposons cinq conditions, et nous prévenons tout de suite : aucune n'est une méthode, aucune ne se déploie, aucune ne s'achète. Ce sont des propriétés observables, formulées de sorte que vous puissiez vérifier vous-même, aujourd'hui, si votre organisation les remplit. Elles ne forment pas un modèle : rien n'indique qu'il faille les cinq, et rien ne dit qu'elles suffisent.

Ce chapitre ne vous donnera pas de méthode, et c'est délibéré

Vous attendez peut-être, à ce stade, que nous sortions un cadre. Des étapes, des rôles, une échelle de maturité, une roue à cinq quartiers. C'est ce que ferait un cabinet, et c'est ce que le marché récompense.

Nous n'allons pas le faire, et la raison tient en une phrase : nous venons de passer une section entière à montrer ce qui arrive aux bonnes idées lorsqu'elles deviennent des référentiels (§7). Écrire ici « le modèle en cinq paliers » reviendrait à fabriquer sous vos yeux la sixième occurrence du motif que nous dénonçons. Vous seriez fondé à nous le reprocher.

Nous proposons donc des **conditions**, et le mot est à prendre au sens où l'entend un ingénieur : une propriété qu'un système possède ou ne possède pas, qu'on constate en regardant, et dont l'absence explique un comportement. Une pratique, on l'installe. Une condition, on l'a ou on ne l'a pas.

Ce choix se paie, et autant le dire tout de suite : ce chapitre est moins actionnable qu'un référentiel, et personne n'en fera de slide. Il est aussi celui qui ne se dévoiera pas, puisqu'on ne dévoie que ce qui s'installe.

Condition 1 — Quelqu'un répond de la qualité interne

Ou quelqu'un en répond sans peser suffisamment, ce qui revient au même — et c'est le cas général. Cette condition découle en droite ligne de la section 3 : la qualité interne cède parce qu'elle est la seule dimension dont le titulaire ne siège pas au rang où l'arbitrage se tranche.

Les quatre conditions suivantes s'appuient sur celle-ci.

Ceux qui viennent du développement reconnaîtront la mécanique. En informatique, un objet vers lequel plus aucune référence ne pointe n'est pas conservé au cas où : il est ramassé, silencieusement, par le ramasse-miettes. Le programme continue de tourner, rien ne signale la disparition, et personne ne s'en aperçoit — jusqu'au jour où l'on cherche l'objet. La qualité interne, dans une organisation dont aucune instance ne la porte, subit exactement ce sort. Elle n'est pas arbitrée contre autre chose. Elle n'est pas non plus supprimée par décision. **Elle est ramassée.**

Une organisation robuste a corrigé cela, et il faut être précis sur ce que « corrigé » veut dire ici, parce que le contresens est facile. Ce n'est pas qu'elle mesure la qualité : nous avons vu en §7 qu'un indicateur qui remonte à un comité cesse assez vite de mesurer ce qu'il mesurait. C'est qu'il existe **un endroit où la dette est arbitrée, au même niveau et avec la même solennité que le budget.** Pas un tableau de bord. Un endroit, avec des gens dedans, et un ordre du jour.

Le test ne pardonne pas : dans votre organisation, quel document s'ouvre quand la qualité interne se dégrade ? Si la réponse est « aucun », la condition n'est pas remplie, quel que soit le discours. Si la réponse est « un tableau de bord », elle ne l'est pas davantage : un tableau de bord informe, il n'arbitre pas.

La forme, elle, varie beaucoup. Comité d'architecture doté d'un droit de veto, ligne budgétaire de remédiation votée annuellement et non redéployable, jalon de reprise technique inscrit dans les mêmes documents que les jalons fonctionnels. Ce qui compte est qu'un arbitrage y ait lieu, et qu'il laisse une trace.

Condition 2 — Quelqu'un peut arrêter la chaîne de production, et l'arrêt tient

C'est l'andon de la section 3, transposé hors de l'usine.

Quelqu'un, quelque part, peut arrêter. Pas alerter : arrêter. Et le mot qui porte n'est pas « peut », il est **opposable** : l'arrêt tient contre la volonté de ceux qui voudraient continuer, y compris quand ils ont de bonnes raisons, y compris quand la date approche.

Nous avons posé cette nuance en §3 et nous la répétons, parce que c'est ici que la plupart des organisations se croient conformes à tort : **une équipe qui peut alerter mais pas refuser n'a rien reçu.** Son alerte remonte compressée, elle est arbitrée par quelqu'un qui ne dispose pas de l'information, et nous voilà revenus au point de départ.

Le test, ici encore, est comportemental et non déclaratif. Non pas « existe-t-il un droit d'arrêt ? », qui figure dans beaucoup de chartes, mais : **quand l'a-t-on exercé pour la dernière fois, et qu'est-il arrivé à la personne qui l'a exercé ?** Une organisation où le cordon existe et n'a jamais été tiré est une organisation où il n'existe pas.

Le monde de l'agilité à l'échelle offre ici un cas d'école, et il est instructif justement parce que le dispositif y est explicitement prévu. À la fin d'un PI planning — ces deux journées où plusieurs équipes construisent ensemble le plan du trimestre — se tient un vote de confiance. Chacun lève la main, de un à cinq doigts, sur une question simple : croyons-nous à ce plan ? Le mécanisme est exactement un cordon. Il est institué, il est public, il arrive au bon moment, et il ne coûte rien à actionner.

J'ai accompagné beaucoup de ces séances, auprès d'équipes comme de core teams. Ce que j'y ai vu souvent n'est pas un vote négatif que l'on ignorerait : c'est un vote favorable rendu par des gens qui, une heure plus tôt, devant leur mur, venaient de dire que la feuille de route affichée ne serait pas tenue. L'information existait. Elle était précise, elle était détenue par ceux qui allaient faire le travail, et elle n'a pas été portée au moment où le dispositif la demandait — parce que personne ne s'est senti autorisé à la dire là.

Le cordon existait donc, et il a remonté un signal vert. Vous noterez que rien, dans cette séquence, ne ressemble à un dysfonctionnement : la séance s'est bien passée, le plan est sorti, le vote a été favorable. C'est précisément ce qui rend la condition 2 difficile à vérifier de l'extérieur, et pourquoi le test qui la vérifie est comportemental.

Je dois ici une précision qui n'est pas flatteuse pour notre métier. En une vingtaine d'années de pratique, je n'ai vu cette condition satisfaite qu'une seule fois. C'était chez un assureur, il y a une douzaine d'années, et elle ne venait d'aucune charte : des architectes et des responsables techniques avaient donné ce droit aux équipiers, et l'ont tenu le jour où il a été exercé. Une fois. Je n'en tire aucune statistique et je me garde de conclure sur ce que je n'ai pas vu ; je signale seulement que la condition la plus facile à inscrire dans un référentiel est celle que j'ai le moins rencontrée sur le terrain.

C'est le Safety-II de Hollnagel et Woods (§6) : la sécurité n'est pas l'absence d'incident, c'est une capacité d'adaptation. Une organisation qui n'a plus rien à signaler a pu cesser de regarder plutôt que devenir plus sûre, et un cordon jamais tiré ne prouve donc pas que le système fonctionne bien. Weick et Sutcliffe

nomment le même réflexe autrement : la **préoccupation pour les défaillances** traite le petit signal comme une information plutôt que comme du bruit, et la **déférence à l'expertise** dit qui doit pouvoir tirer le cordon — celui qui dispose de l'information, indépendamment de son rang. Une organisation où seule la hiérarchie peut l'actionner n'a pas satisfait la condition 2 ; elle a déplacé le goulot d'étranglement.

Condition 3 — Les projets clos ont ajusté leur périmètre, et leurs documents le disent

Cette condition sépare le discours de la pratique, et elle se vérifie brutalement : elle s'établit en ouvrant des documents, pas en interrogeant des intentions.

Une organisation robuste a fait le choix décrit en §3 : coût et délai tenus, qualité invariante, périmètre ajustable. Ce choix ne vaut que s'il est **visible dans les documents de pilotage** — des livraisons passées qui montrent des périmètres réduits, tracés, communiqués au moment de la décision, sans sanction pour ceux qui les ont portées.

Si vos documents des trois dernières années montrent au contraire un périmètre systématiquement tenu, l'une des trois autres dimensions a absorbé l'écart. C'est arithmétique, et c'est la question de §3 : où est passé l'écart ?

Il existe une manière plus discrète d'échouer, et elle mérite d'être décrite parce qu'elle ressemble à une réussite. Dans les revues de portefeuille, il arrive qu'un produit s'arrête parce que son budget s'achève, alors qu'il se trouve dans un état que les équipes appellent entre elles un « MVP++ » : le produit minimum viable, augmenté de ce qui a pu être ajouté avant la fin, mais pas achevé au sens fonctionnel. Le périmètre a donc bien été la variable qui a cédé, et la première moitié de la condition est satisfaite. Mais le document de clôture, lui, ne dit pas que le produit a été **arrêté** ; il dit qu'il est **fini**. Cette substitution d'un mot à l'autre suffit à annuler tout le bénéfice, car elle efface ce que la condition cherchait à rendre lisible : qu'un arbitrage a eu lieu, à quelle date, et au profit de quoi. Une organisation peut ainsi ajuster son périmètre en permanence sans en garder aucune trace exploitable.

Condition 4 — Les temps de respiration tiennent même sous tension

Une marge qui ne survit pas au premier trimestre difficile n'est pas une marge : c'est un reliquat que personne n'avait encore réclamé. Cette condition-ci ressemble à du gras, et c'est pourquoi elle part la première dès que la pression monte.

Tom DeMarco l'a formulée mieux que quiconque, dans Slack en 2001 : le mou n'est pas du gaspillage résiduel qu'une organisation efficace aurait éliminé, c'est **la capacité d'adaptation elle-même**. Une organisation dont toutes les ressources sont affectées à cent pour cent ne peut, par construction, absorber aucun imprévu : ni un incident, ni une opportunité, ni une reprise technique. Elle n'est pas efficace, elle est rigide, et la différence ne se voit que le jour où quelque chose arrive.

La condition n'est donc pas « garder du mou », qui ne survivrait pas au premier exercice budgétaire contraint. Elle est plus exigeante : **le mou est budgété, nommé, et défendu comme une ligne**. Un pourcentage de capacité réservé à la remédiation. Un temps d'ingénierie affecté à aucun projet. Ce qui n'a pas de ligne budgétaire n'existe pas dans une discussion budgétaire, et c'est le mécanisme de la section 3 appliqué cette fois à la marge plutôt qu'à la qualité.

Le monde de l'agilité à l'échelle a d'ailleurs produit cette ligne et lui a donné un nom : le sprint dit « I&P », pour innovation and planning, une itération entière sur cinq, sans engagement de livraison, réservée à ce qui ne trouve jamais sa place ailleurs. Sur le papier, la condition est donc satisfaite par construction, et c'est ce qui rend son sort instructif — comme pour le vote de confiance de la condition 2, qui vient du même référentiel. Notons-le, car la conclusion inverse serait facile et fausse : **ces deux dispositifs sont bien conçus**. Le vote de confiance demande l'information à ceux qui la détiennent, au bon moment ; la marge d'innovation la budgète et la nomme. Ce n'est donc pas le dispositif qui manque à ces organisations, c'est ce qui le tient.

Le référentiel le sait d'ailleurs parfaitement, et c'est peut-être le détail le plus parlant de toute cette section. Son plan de déploiement ne commence pas par les équipes : il commence par la formation des décideurs. L'ordre n'est pas administratif, il est logique — une équipe formée à voter sa confiance dans un environnement qui ne supporte pas d'entendre un doute apprendra surtout à voter favorablement. Ce qui se saute, dans les déploiements que j'ai accompagnés, n'est presque jamais la formation des équipes. C'est celle d'au-dessus. J'ai vu ce sprint, plus souvent qu'à son tour, se transformer en sprint de rattrapage. Le mou était budgété, il était nommé, il était visible de tous, et il a été repris sans qu'aucune décision ne soit prise nulle part, simplement parce que le trimestre avait du retard. Budgéter une marge ne suffit donc pas à la protéger : il faut encore que quelqu'un ait le mandat de refuser qu'on la consomme. Ce qui nous ramène à la condition 2, et explique pourquoi ces cinq conditions ne tiennent pas séparément.

DeMarco écrivait en 2001, depuis l'ingénierie logicielle. Hollnagel et Woods (§6) arrivent au même constat depuis l'ingénierie de la sécurité : la résilience d'un système est sa capacité d'adaptation, et elle se **réserve** avant, elle ne se déduit pas après coup. Le mou est donc la condition de possibilité du budget de remédiation, pas sa variable d'ajustement. Point vérifié le 03/09/2026 : Hollnagel cite nommément Holling (1973) comme l'origine du mot « résilience » — filiation directe, pas convergence indépendante. Pour DeMarco, l'affaire reste ouverte : Slack n'a pas de bibliographie, ce qui prouve seulement qu'il ne cite personne dans ce livre.

Holling donne à cette condition son cas empirique, chiffré sur soixante ans de données de terrain plutôt que déduit d'un principe d'ingénierie. Les pêcheries des Grands Lacs ont été gérées pour l'exact inverse de la condition 4 : un rendement soutenu maximal, une extraction poussée jusqu'à la limite du stock, sans marge maintenue en retrait. La résilience s'est vidée **sans aucun signe visible**, jusqu'à l'effondrement du lac trout et du corégone — alors même que la pression de pêche avait déjà été relâchée. Même mécanisme qu'en condition 2 : l'absence de problème visible ne prouve pas que la capacité d'absorber existe encore.

Condition 5 — Ceux qui disent non ont un endroit où reprendre leurs arbitrages

Les quatre premières conditions portent sur le dispositif. Celle-ci porte sur les personnes, et elle est de nature différente — mais elle s'énonce de la même façon : cet endroit existe, ou il n'existe pas.

Toutes les précédentes reposent, au bout du compte, sur des individus qui disent non. Le responsable d'architecture qui refuse une mise en production. L'équipe qui déclare que ce n'est pas fini. Le directeur qui défend une ligne de remédiation dans un arbitrage budgétaire où elle ne rapporte rien.

Ces refus ont un coût qu'on compte rarement. Ils s'exercent en isolement, souvent contre une majorité pressée, parfois contre sa propre hiérarchie, et toujours sans preuve immédiate qu'on avait raison : la catastrophe évitée ne se voit pas. Une personne qui tient une telle position pendant trois ans sans espace où l'examiner finit par céder, ou par partir. Les deux passent pour des trajectoires individuelles.

Une organisation robuste a prévu un espace où ces arbitrages se reprennent — hors de la ligne hiérarchique, hors de l'évaluation, avec quelqu'un dont ce n'est pas l'intérêt que la décision aille dans un sens plutôt que dans l'autre. Ce n'est

pas du soutien psychologique, et ce n'est pas de la formation. C'est un lieu où l'on peut se demander à voix haute si l'on a eu raison, sans que la question compte contre soi.

Son absence ne produit pas d'incident. Elle produit une érosion : les gens qui savaient dire non s'en vont, ou apprennent à ne plus le dire. Et comme les quatre autres conditions dépendent d'eux, elles tombent ensuite, sans que rien n'ait été décidé.

Ce que ces cinq conditions ne disent pas

Une dernière précaution, qui est aussi une dette contractée en §3.

Nous avons écrit que la décision de qualité devait se prendre là où l'information existe. Cela suppose que celui qui la prend ait la **capacité** de juger : contexte, expérience, compréhension des conséquences. Confier cette décision à qui ne les a pas ne supprime pas le point de défaillance, il le déplace, et parfois l'aggrave.

Aucune de nos cinq conditions ne garantit cette capacité. Elles créent les conditions d'un jugement ; elles ne le produisent pas. C'est, à notre sens, la limite principale de ce chapitre, et nous préférons l'écrire ici plutôt que de laisser croire à un dispositif complet.

Ce qui nous amène à la question que nous vous laissons, et que nous poserions à un dirigeant qui n'aurait le temps que d'une : **des cinq conditions ci-dessus, combien pourriez-vous démontrer par un document plutôt que par une intention ?**

§10. Les cinq questions sur votre variable d'ajustement

Où nous en sommes. La section précédente a décrit cinq conditions de robustesse. Elles sont observables, mais elles demandent du recul — et le recul est exactement ce dont on manque quand on est dedans.

Ce que nous allons voir ici. Cinq questions, à poser dans votre organisation. Elles tiennent sur une page, elles se posent en réunion, et elles ne demandent aucune préparation. Nous ne proposons ni score ni niveau : chaque question est suivie de ce que sa réponse révèle. C'est le seul chapitre de cette étude qui ne cherche pas à vous convaincre de quelque chose — il cherche à vous faire regarder, et c'est une démarche de coaching.

Un mot sur ce que ces questions ne sont pas : elles ne recoupent pas une à une les cinq conditions du chapitre précédent. Deux d'entre elles portent sur la même condition, et la dernière condition — celle qui concerne les personnes qui tiennent ces arbitrages — n'a volontairement aucune question ici. Elle ne se vérifie pas en réunion.

1. Qui, nommément, répond de la qualité interne ?

Une personne, pas une fonction. Un nom, pas une équipe.

Si vous n'avez pas de nom — c'est le cas le plus fréquent, et c'est le diagnostic complet de cette étude en une réponse. La qualité est votre variable libre, et elle absorbe déjà tout ce que les trois autres dimensions ne veulent pas prendre.

Si vous avez un nom, posez la question suivante : cette personne peut-elle bloquer une livraison ? Si elle ne peut qu'alerter, vous avez un porte-parole, pas un titulaire.

2. À quel niveau la dette technique est-elle arbitrée ?

Cherchez l'instance, pas le rapport.

Si elle est arbitrée dans les équipes — elle est traitée là où l'information existe, ce qui est la bonne place, mais elle perd contre toute pression venue d'au-dessus. Vous avez le jugement sans le pouvoir.

Si elle est arbitrée en comité de direction — vérifiez alors ce qui y remonte. S'il s'agit d'un indicateur, vous avez le pouvoir sans le jugement, et l'indicateur deviendra une cible.

Si elle n'est arbitrée nulle part — elle est arbitrée quand même, chaque jour, par ceux qui manquent de temps.

3. Quand quelqu'un a-t-il dit « ce n'est pas prêt » pour la dernière fois, et que s'est-il passé ?

Une date, et une conséquence.

Si personne ne s'en souvient — le cordon n'existe pas, quoi qu'en disent vos chartes. Un dispositif d'arrêt qu'on n'a jamais tiré n'a jamais été éprouvé, et rien ne dit qu'il tiendrait.

Si la livraison a eu lieu quand même — vous savez maintenant ce que vaut le dispositif, et vos équipes le savent depuis plus longtemps que vous.

Si la livraison a été décalée — regardez ce qui est arrivé ensuite à la personne qui l'a demandé. C'est là, et nulle part ailleurs, que se lit la réponse à cette question.

4. Sur les douze derniers mois, quels périmètres avez-vous réduits ?

Des exemples, avec leurs dates.

Si vous n'en trouvez aucun — et que la charge n'a pas diminué, alors l'écart est allé ailleurs. Il n'a pas disparu : il est dans la qualité, dans la fatigue de vos équipes, ou dans un report vers l'exploitation qui paiera plus tard.

Si vous en trouvez — vérifiez qu'ils ont été annoncés au moment de la décision, et non constatés après coup. Un périmètre qui se réduit en silence n'est pas une variable assumée, c'est un retard qu'on n'a pas nommé.

5. Combien de capacité d'ingénierie n'est affectée à aucun projet ?

Un pourcentage, tiré de vos propres plans de charge.

Si la réponse est zéro — votre organisation ne peut absorber aucun imprévu, et le prochain incident sera payé en dette ou en heures supplémentaires, faute d'un troisième endroit où le mettre.

Si la réponse est un chiffre non nul mais non budgété — il disparaîtra au premier arbitrage contraint, parce que ce qui n'a pas de ligne n'existe pas dans une discussion budgétaire.

Si la réponse est un chiffre budgété et défendu — alors vous avez quelque chose que peu d'organisations conservent sous pression. Regardez qui l'a défendu la dernière fois, et assurez-vous que cette personne n'est pas seule.

Comment lire vos réponses

Il n'y a pas de total, et c'est délibéré : un score transformerait ces questions en indicateur, et nous avons consacré une section entière à ce qui arrive aux indicateurs.

Ce que nous suggérons plutôt, et qui prend une heure : posez ces cinq questions à trois personnes qui occupent des positions différentes — un membre de direction, un responsable d'équipe, quelqu'un qui écrit du code. **L'information n'est pas dans les réponses. Elle est dans l'écart entre elles.** Là où les trois disent la même chose, vous avez un fait. Là où elles divergent, vous avez trouvé exactement l'endroit où l'information se perd en remontant — et cette étude vous aura servi à quelque chose.

§11. Ce que nous n'avons pas prouvé

Où nous en sommes. L'argumentation est close. Nous avons proposé un diagnostic, un cadre, un motif historique, une lecture de la période qui s'ouvre, et cinq conditions.

Ce que nous allons voir ici. Ce que tout cela ne démontre pas. Un document qui ne dit pas où il s'arrête ne permet pas de savoir jusqu'où le suivre.

Les données sur l'IA sont corrélationnelles. Nous l'avons écrit en §8 et nous le répétons ici : ni GitClear ni Ox Security ne disposent d'un groupe de contrôle, et la période 2020-2024 a connu d'autres bouleversements que l'arrivée des assistants de code. Nous présentons un faisceau convergent avec les observations DORA. Un faisceau n'est pas une démonstration causale, et nous n'en tirons aucune.

La fenêtre 2026-2028 est un cadre de lecture, pas une prédiction datée. Rien dans nos matériaux ne permet de dire quand les effets décrits deviendront visibles, ni s'ils le deviendront de façon spectaculaire ou par une lente érosion. Nous soutenons qu'un régime produit une trajectoire ; nous ne prétendons pas savoir à quelle vitesse il la parcourt. Quiconque lira ce document en 2029 pour vérifier une échéance se trompera de lecture.

Les observations de terrain sont des témoignages directs, et n'ont rien d'un protocole. Ce document en rapporte six, aux chapitres 2, 4, 7, 8 et 9. Aucune n'est un matériau recomposé à partir d'un corpus : ce sont des scènes vécues par un seul observateur — moi-même — et racontées de mémoire. Elles n'ont ni échantillonnage, ni terme de comparaison, ni rien qui ressemble à un contrôle ; un autre praticien placé aux mêmes endroits n'en aurait pas nécessairement retenu les mêmes moments. Nous les présentons pour ce qu'elles font, et pour rien d'autre : elles donnent à **reconnaître** un mécanisme, elles ne l'établissent pas. Ce qui l'établit, ce sont les données publiques des sections 5 et 8.

Aucun entretien n'a été conduit. Ce n'est pas un manque à combler plus tard : des entretiens auraient apporté de la matière, et rendu l'étude invérifiable — personne d'autre que nous n'y aurait eu accès. Tout ce qui est affirmé ici est contrôlable par un lecteur qui n'a que les mêmes sources. La contrepartie est que ce document ne contient aucun cas d'organisation identifiable.

Certains chiffres viennent d'un communiqué et non du rapport complet.

Le dossier de méthode, publié à côté de cette étude, dit lesquels, et ce qui n'y est donc pas vérifiable — au premier rang de quoi la méthode de production de ces chiffres.

Des convergences repérées, et laissées de côté. Plusieurs corpus recoupent cette étude sans y être mobilisés : l'écologie de l'action d'Edgar Morin, le cadre Cynefin de Dave Snowden, la systémique et la notion d'homéostasie. Chacun éclairerait le mécanisme décrit ici. Aucun ne le contraint davantage que la loi de la variété requise, déjà retenue — et nous nous sommes tenus à la règle énoncée dans ce même dossier : une référence n'entre que si elle interdit quelque chose. Elles restent ouvertes pour une version ultérieure.

Nos cinq conditions ne garantissent pas la capacité de juger. C'est la limite la plus sérieuse de la quatrième partie, et nous l'avons déjà posée en §9. Une organisation peut remplir les cinq conditions et confier ses arbitrages de qualité à des personnes qui n'ont pas le contexte pour les rendre. Le dispositif serait alors formellement conforme et substantiellement vide. Nous ne savons pas énoncer une condition qui garantirait cette capacité, et nous n'en proposons donc aucune.

Nous n'avons pas prouvé que les organisations décrites ici se porteraient mieux en changeant de régime. Nous avons montré qu'un régime transfère un coût dans le temps et qu'il le rend invisible. Nous n'avons pas montré que l'autre régime soit préférable dans toutes les circonstances, ni pour toutes les organisations, ni sur tous les horizons. Une entreprise dont l'horizon réel est de dix-huit mois — parce qu'elle sera vendue, parce que son marché disparaît, parce que son produit est jetable par nature — fait peut-être un calcul juste en sacrifiant sa robustesse.

Ce que nous soutenons est plus étroit, et nous nous y tenons : **ce calcul devrait être fait sciemment.** Aujourd'hui, dans la plupart des organisations que décrit cette étude, il n'est pas fait du tout.

§12. Ouverture — qui valide ce qu'on ne voit plus

Où nous en sommes. Nous avons dit ce que nous soutenons, et nous venons de dire ce que nous ne prouvons pas. Le document pourrait s'arrêter là.

Ce que nous allons voir ici. Une question que cette étude a produite sans la traiter, et qui nous paraît plus importante que ses conclusions. Nous l'énonçons et nous nous arrêtons : elle appelle un autre travail que celui-ci.

Il y a une phrase, dans tout ce qui précède, dont nous n'avons pas fini de mesurer la portée.

Ce qui ne se voit pas dans un reporting se sacrifie sous pression de performance.

Nous l'avons appliquée à la qualité interne du logiciel, parce que c'est là qu'elle se démontre le mieux et qu'elle s'y chiffre. Mais rien, dans le raisonnement, ne la limite au code. Elle vaut pour tout ce qu'une organisation produit et qui échappe à ses instruments de remontée — et c'est ici que la période qui s'ouvre nous préoccupe.

L'IA générative, avons-nous dit en §8, n'a pas de direction propre : elle amplifie le régime qu'elle trouve. L'agentique, qui vient après, fait quelque chose de différent et de plus inquiétant. Elle ne change pas la loi que nous venons d'énoncer : **elle en multiplie l'assiette.**

Regardez ce qui augmente. Le volume de code que personne ne relit ligne à ligne. Le nombre de décisions prises dans une chaîne automatisée dont aucune trace lisible ne subsiste. Les adoptions d'outils que personne n'observe, dans des équipes où plus personne n'a le temps de regarder comment le travail se fait réellement. Chacune de ces choses était déjà partiellement invisible il y a trois ans. Chacune l'est massivement aujourd'hui, et le sera davantage demain.

Or nous avons montré, tout au long de cette étude, ce qui arrive à ce qui n'est pas vu : ce n'est pas volé, ce n'est pas négligé, ce n'est pas maltraité. C'est simplement **absorbé** — parce que rien ne le retient, et parce que personne n'a d'endroit où en rendre compte.

La question qui organise les vingt prochaines années n'est donc probablement pas celle que posent les comités de direction aujourd'hui. Ce n'est pas comment produire davantage : cette question-là est en train de se résoudre toute seule, et plus vite que prévu.

C'est celle-ci, et nous la laissons ouverte parce qu'elle mérite mieux qu'une réponse écrite en fin de document :

Qui valide ce qu'on ne voit plus ?

Bibliographie

Une convention, et elle vaut avertissement. Les entrées marquées * n'ont pas été lues intégralement : ce que l'étude en cite provient des communications officielles de leurs auteurs. Ce n'est pas un défaut caché, c'est une limite déclarée — et elle porte toujours sur la **méthode** de production des chiffres, jamais sur les chiffres eux-mêmes.

Ouvrages et articles

Ashby, W. Ross. An Introduction to Cybernetics, Chapman & Hall, 1956, 156 p. Diffusion non lucrative autorisée par la succession ; le texte est librement accessible en ligne. → La loi de la variété requise. Le chapitre 11 est celui dont l'étude tire son palier théorique : un régulateur ne peut maintenir un système que s'il dispose d'au moins autant de variété que ce système.

Collins, Jim. Built to Last (avec Jerry Porras), 1994 ; Good to Great, 2001. → Deux enquêtes sur ce qui distingue les entreprises qui durent. L'étude ne les mobilise que pour rétablir une attribution : l'épigramme du principe 9 de SAFe est de lui, non de Drucker.

Cunningham, Ward. Debt Metaphor, vidéo, 14 février 2009 ; transcription publique sur le wiki c2 et sur cmdev.com. La métaphore elle-même date du début des années 1990. → La définition d'origine de la dette technique, et le contresens que son auteur écarte lui-même : la dette n'est pas du code mal écrit, c'est l'écart entre le code et la compréhension acquise du problème. Douze minutes, et elles valent le détour.

DeMarco, Tom. Slack: Getting Past Burnout, Busywork, and the Myth of Total Efficiency, Broadway Books, 2001. → Le mou n'est pas un gaspillage résiduel : c'est la capacité d'adaptation elle-même. Du même auteur, avec Timothy Lister, Peopleware (1987), sur l'idée que les échecs de projet sont sociologiques avant d'être techniques.

Drucker, Peter. The Practice of Management, Harper & Row, 1954, chapitre « Management by Objectives and Self-Control ». → Le travailleur du savoir se définit par son asymétrie d'information, et le seul contrôle qui opère sur lui est l'auto-contrôle. Lire le titre entier du chapitre : la seconde moitié est celle qui a disparu de la postérité du management par objectifs.

Hamant, Olivier ; Enlart, Sandra ; Charbonnier, Jean-Marc. L'Entreprise Robuste, Odile Jacob, 2025. → La performance est la stratégie d'un monde stable ; la robustesse, la condition de persistance d'un monde fluctuant. Du même auteur, plus court et plus accessible : Antidote au culte de la performance. La robustesse du vivant, coll. Tracts, Gallimard.

Holling, C. S. « Resilience and Stability of Ecological Systems », *Annual Review of Ecology and Systematics*, vol. 4, 1973, p. 1-23. → L'article fondateur du mot « résilience » tel qu'on l'emploie aujourd'hui, et le cas empirique des pêcheries des Grands Lacs : optimiser pour la stabilité use la résilience sans aucun signe visible avant l'effondrement.

Hollnagel, Erik ; Woods, David. Resilience Engineering: Concepts and Precepts, Ashgate, 2006 ; et Hollnagel, Safety-I and Safety-II, Ashgate, 2014. → La sécurité n'est pas l'absence d'incident mais une capacité d'adaptation : une organisation qui n'a plus rien à signaler peut simplement avoir cessé de regarder.

Lehman, Meir M. « Programs, Life Cycles, and Laws of Software Evolution », *Proceedings of the IEEE*, vol. 68, n° 9, septembre 1980, p. 1060-1076. DOI 10.1109/PROC.1980.11805. Copie libre sur la page de cours de Dewayne Perry, UT Austin. → La complexité d'un logiciel qui évolue croît et sa qualité perçue décline, sauf si un travail actif est mené pour l'en empêcher. Quarante-cinq ans, et rien à retirer.

Martin, Roger L. * « The High Price of Efficiency », *Harvard Business Review*, janvier-février 2019 ; version développée en ouvrage : *When More Is Not Better*, Harvard Business Review Press, 2020. → L'optimisation poussée d'un système le couple plus étroitement, et un incident local s'y propage au lieu de s'y amortir.

Reinertsen, Donald G. *The Principles of Product Development Flow: Second Generation Lean Product Development*, Celeritas Publishing, 2009. → Le coût du délai : combien coûte, en euros par semaine, qu'une capacité arrive plus tard qu'elle n'aurait pu. C'est le vocabulaire qui permet de parler de qualité en comité de direction sans prononcer le mot.

Ridgway, V. F. « Dysfunctional Consequences of Performance Measurements », *Administrative Science Quarterly*, vol. 1, n° 2, 1956, p. 240-247. → Sept pages, et vingt ans avant Goodhart : les mesures de performance produisent des conséquences dysfonctionnelles, d'autant plus sûrement qu'elles sont mieux conçues. Peu citée, et c'est dommage.

Strathern, Marilyn. * « "Improving ratings": audit in the British University system », *European Review*, vol. 5, n° 3, 1997, p. 305-321. → C'est d'elle qu'est la formule qu'on attribue à Goodhart : quand une mesure devient un objectif, elle cesse d'être une bonne mesure. Article payant ; l'histoire de la formule est retracée dans Chrystal & Mizen, « Goodhart's Law: Its Origins, Meaning and Implications for Monetary Policy », Bank of England, 2001.

Taleb, Nassim Nicholas. *Antifragile: Things That Gain from Disorder*, Random House, 2012. → Une seule idée est retenue ici : l'optimisation n'est pas un état neutre, elle est elle-même un facteur de fragilisation.

Weick, Karl E. ; Sutcliffe, Kathleen M. *Managing the Unexpected: Resilient Performance in an Age of Uncertainty*, 2^e éd., Jossey-Bass, 2007, 208 p. → Les cinq principes des organisations à haute fiabilité, dont deux sont mobilisés ici : la préoccupation pour les défaillances, et la déférence à l'expertise. Attention à l'édition : les cinq principes sont formulés différemment en 2001, 2007 et 2015.

Rapports et enquêtes

CISQ, *The Cost of Poor Software Quality in the US: A 2022 Report*, Consortium for Information & Software Quality, 2022, 61 p. → Le coût annuel de la non-qualité logicielle aux États-Unis, et — **comptés séparément** — le stock de dette technique accumulée. La distinction entre le flux et le stock est ce qui se perd le plus souvent dans les reprises.

DORA / Google Cloud, *Accelerate State of DevOps Report 2024*, dixième édition annuelle. Téléchargement libre, sans formulaire. → La corrélation entre débit de livraison et stabilité (p. 79), et l'effet mesuré de l'adoption de l'IA sur la livraison (figure 10, p. 38-39). Le rapport signale lui-même que ce second résultat va contre ses attentes, ce qui est rare et mérite d'être salué.

GitClear / Alloy.dev Research, *AI Copilot Code Quality: Evaluating 2024's Increased Defect Rate via Code Quality Metrics*, v2025.2.5, données 2020-2024, 211 millions de lignes. Inscription requise. → La chute des lignes déplacées, la montée du copié-collé, la hausse du churn, et le facteur ×8 sur la fréquence des blocs de cinq lignes dupliquées ou plus. Lire la section méthodologique avant les chiffres : le rapport ne compare à aucun moment du code écrit avec assistant à du code écrit sans. Une édition 2026 existe, *The Maintainability Gap*, non instruite ici.

OX Security, * The Army of Juniors: The AI Code Security Crisis, 23 octobre 2025, plus de 300 dépôts libres analysés. → La distribution de dix anti-patterns dans le code généré. Ce qui n'est pas public : la méthode de détection, la sélection des dépôts, et la façon dont le code « généré par IA » est identifié.

Splunk / Oxford Economics, * The Hidden Costs of Downtime, vague de mai 2026, 2 000 dirigeants interrogés dans 20 pays. → Le coût de l'indisponibilité non planifiée pour les grandes entreprises. **Ce sont des coûts estimés par les répondants**, non des pertes constatées.

PMI, Pulse of the Profession, édition 2025. → Premier recul enregistré du nombre d'organisations déclarant employer des méthodes agiles.

Un dossier de méthode accompagne cette étude. Il détaille le dispositif de preuve, le régime de vérification de chaque source et les erreurs de sourcing corrigées en cours de route. Il n'a pas sa place dans ce document et se lit à part : **roquesci.com/etudes/plus-le-monde-fluctue-methode.pdf**