

CROspector : le moteur déterministe que nous avons construit, et les juges-machines pour lesquels il a été conçu

Ceci est la Partie 3 d'une histoire en trois parties. La [Partie 1](#) était le pari. La [Partie 2](#) était le verdict. Ceci est la réponse.

L'audit de 3 h 47 du matin

Chaque nuit, pendant que vous dormez, un robot d'exploration que vous n'avez pas invité ouvre vos pages produit (PDP) et audite trois prix pour chacune : **(1)** le prix que Googlebot lit sur la page, **(2)** le prix que vous avez envoyé à Merchant Center dans votre flux (qui porte déjà votre `sale_price` actif), et **(3)** le prix que votre caisse facturerait réellement si un acheteur ajoutait l'article au panier. Vous vous attendez à ce que ces trois prix concordent, parce que vous avez informé Google de la promotion via le flux et que votre caisse est cohérente avec lui. Parfois, ils ne concordent pas — à 3 h 47, du JavaScript injecte votre bannière promotionnelle une fraction de seconde après que Googlebot a consommé le HTML initial, et le prix que le robot enregistre est le prix avant promotion, alors que votre flux et votre caisse portent déjà le prix réduit. C'est un mode d'échec que Google lui-même [documente](#) : « Googlebot explore les données présentes dans le HTML renvoyé par votre serveur web » — et si votre prix final est rendu après ce HTML initial, « cela déclenchera une erreur ». Au moment où vous vous réveillez, chaque navigateur piloté par un humain sur la planète a rendu votre page correctement, aucun acheteur n'a rencontré le moindre état défaillant, et le rendu partiel que Googlebot a consommé n'est pas reproductible. Mais Google détient l'audit sur votre compte ; vous n'avez aucune copie des octets que le robot a lus, ni aucun moyen de les reproduire — et une machine a déjà décidé de ce qui arrive ensuite à cette référence (SKU), sur des preuves que vous ne pouvez pas inspecter, pendant qu'aucun humain dans votre entreprise ne regardait.

À quelle fréquence cela arrive-t-il à votre catalogue en ce moment ? Personne ne peut répondre aujourd'hui. Pas Google, qui vous montre le verdict mais pas l'état de la page qui l'a produit ; pas votre équipe, qui ne voit que l'avis de désapprobation. Rien dans votre pile n'enregistre ce que le robot a vu à cet instant. Le mécanisme que Google appelle « [désapprobation préventive de l'article](#) » — retrait sur suspicion d'incohérence, avec une fenêtre de réexamen de 28 jours au niveau du compte pendant laquelle les articles restent suspendus — est le texte littéral de la politique, et ses conséquences sont sans ambiguïté : la politique de Google est que « [les produits désapprouvés cessent de s'afficher sur l'ensemble de Google](#) ». La référence concernée disparaît des fiches gratuites (du revenu que vous ne gagnez plus) et des inventaires payants Shopping et Performance Max sur lesquels votre budget média enchérit (de la dépense

désormais soit perdue, soit réorientée à un coût plus élevé vers une référence moins pertinente qui ne convertit pas au même taux). La fréquence est inconnue ; l'asymétrie, non.

Ce n'est pas une anecdote sur la sévérité de Google. C'est la première apparition d'une espèce : le **juge-machine** — un auditeur algorithmique qui lit les surfaces du commerce référence par référence, rend des verdicts sur une horloge mesurée en temps machine, pas le vôtre, et attache la sanction directement au revenu. Le robot Shopping de Google est le juge en place. Les agents acheteurs que l'on raccorde désormais aux protocoles de paiement et de passage en caisse sont la prochaine instance : en [octobre 2025](#), Visa a annoncé son *Trusted Agent Protocol*, un cadre piloté par l'écosystème permettant à des agents IA de transiger sur les sites marchands, avec Adyen, Checkout.com, Cloudflare, Mastercard, Microsoft, Shopify, Stripe et Worldpay parmi les collaborateurs de lancement nommés. Les machines décident de plus en plus si vos produits peuvent être trouvés, montrés et achetés — et elles lisent votre boutique comme les machines lisent : littéralement, de façon répétée, sans pardon.

La Partie 1 de cette série était un pari. La Partie 2 était le verdict. Cette partie est la réponse : ce que nous avons construit, et le marché que les juges-machines créent pour cela.

Ce que le verdict a tranché

La Partie 1 soutenait que la conversion n'est pas un problème d'optimisation mais un problème de **faisabilité** : un parcours ne convertit que s'il vaut la peine d'être commencé, qu'il est possible de le terminer, et qu'il est fiable dans des conditions réelles. La Partie 2 a publié le verdict de notre première tentative d'instrumenter cette faisabilité à l'échelle : 32 341 \$ de dépenses AWS pour faire tourner un agent LLM de pointe sur 10,5 millions de pages produit chez 1 056 enseignes françaises. Taux de complétion : 9 %. Pas parce que les sites nous bloquaient — nous avons franchi environ 92 % des systèmes anti-bot qui protègent ces sites : Cloudflare, Akamai, DataDome, et le reste du parcours du combattant « prouvez que vous êtes humain ». Les 9 % représentent ce qu'une IA bâtie sur un modèle de langage de pointe a réellement produit en essayant de lire des pages d'enseignes réelles de bout en bout. La raison est inscrite dans la technologie : l'IA repart de zéro à chaque page, ses connaissances figées à l'instant où son entraînement s'est arrêté, et rien de ce qu'elle apprend de votre catalogue ne se reporte sur la page suivante. Le même plafond apparaît dans chaque benchmark académique publié cité par la Partie 2 — WebArena, OSWorld, et les autres. La limite n'est pas un bug que nous aurions pu contourner par l'ingénierie. C'est une propriété du fonctionnement de ce type d'IA.

Un détail de cette expérience compte davantage aujourd'hui qu'à l'époque. Les deux anomalies de panier que nous avons conçu l'agent pour détecter étaient une **rupture de stock** et un **écart de prix où le prix du panier dépasse celui de la page produit**. Nous les avons choisies avant de savoir ce que nous construirions ensuite. Il se trouve que ce sont exactement les deux états que les juges-machines sanctionnent et sur lesquels le trafic payant meurt. L'expérience a échoué. Sa cible a survécu.

Ce que nous avons construit

Re-juger, voilà ce qui échoue. L'IA que nous avons fait tourner dans la Partie 2 re-devinait le sens de chaque page, à chaque fois, parce que c'est ainsi que fonctionne ce type d'IA. Un humain relit jusqu'à ce que l'attention s'effondre — et à des dizaines de milliers de pages, aucune équipe n'est assez nombreuse ni assez attentive pour lire chaque page à la main. Nous avons donc supprimé la répétition elle-même.

Le moteur de SHORA enregistre le **gabarit** de la page d'une enseigne une seule fois — quel genre de page c'est — et non les pages individuelles elles-mêmes. Une personne l'enregistre une fois, avec toute son attention, sur une page en direct exactement telle qu'un acheteur la voit ; le moteur rejoue ensuite cet enregistrement sur chaque page produit du site, peu importe le nombre de produits qu'il vend. Le rejeu est mécanique, pas interprétatif — le même gabarit, lu de la même manière, aussi souvent que vos opérations l'exigent, sans aucune IA qui réfléchit lorsque le moteur lit vos pages, et sans relecteur humain dans la boucle.

L'enregistrement du gabarit survit aux **changements de contenu** — un prix change, un niveau de stock change, un nouveau produit est ajouté, une bannière promotionnelle se charge tardivement, un test A/B tourne sur la page — sans réenregistrement. Il survit aussi aux **refontes** — nouveau look, nouvelles couleurs, nouvelle mise en page, nouveau framework sous le capot. Aucun de ces changements ne déclenche un réenregistrement, parce que l'enregistrement est lié à l'intention structurelle de la page, pas à son apparence de surface. (La manière dont ce lien tient à travers ces mutations est la part que nous gardons.)

Il existe exactement un événement qui exige un nouvel enregistrement : un **changement de l'intention structurelle elle-même**. Si l'enseigne re-conçoit ce qu'est sa page produit — le prix cesse d'être un prix et devient « à partir de X », « ajouter au panier » devient « demander un devis » — l'enregistrement existant échoue bruyamment, preuve à l'appui, dès le tout premier rejeu. Vous payez un nouvel enregistrement pour ce gabarit, et les rejeux reprennent sur l'ensemble du site. C'est le coût honnête de l'architecture : vous le payez une fois, quand la page devient une page fondamentalement différente.

Voici comment le moteur répond à chacun des modes d'échec de la Partie 2 :

- **Pourquoi notre moteur ne commet pas les erreurs de l'IA.** Une personne porte le jugement minutieux sur ce qu'est une page — une fois par gabarit, couvrant chaque produit que le site vend — et le moteur le lit ensuite mécaniquement pour toujours. L'IA le re-devinait à chaque page ; le moteur, jamais.
- **Ce qui se passe quand quelque chose change réellement sur une page.** Le moteur n'improvise pas. Si la page change d'une manière qui compte, le rejeu s'arrête et vous le signale, la page sous les yeux comme preuve. L'échec de l'IA était l'inverse — elle continuait à produire des réponses, avec assurance, et vous n'appreniez qu'elles étaient fausses que par un symptôme en aval que vous risquiez de ne pas détecter.

- **Le seul type d'échec qu'une décision financière ne peut pas survivre.** Silencieux, sûr de lui, et faux. C'est le schéma d'échec dans lequel l'IA de la Partie 2 est tombée — des sorties fausses qui se sont composées tout au long du parcours jusqu'à ce que quelque chose casse en aval, ce qui explique pourquoi le taux de complétion s'est arrêté à 9 % au lieu de la barre de 80 % que nous nous étions engagés à atteindre. Le moteur ne peut pas produire cet échec — quand il ne peut pas lire une page correctement, il s'arrête et le dit. Nous n'avons pas construit un système qui ne casse jamais ; personne d'honnête ne peut le prétendre. Nous avons construit un système qui ne peut pas mentir.
- **La différence de coût.** Par tâche web réalisée, faire tourner le moteur coûte environ un dixième du coût d'un agent IA sur le même catalogue. C'est la différence entre surveiller votre catalogue complet à la cadence dont vos opérations ont réellement besoin, et être contraint d'en échantillonner une fraction parce que le coût unitaire de l'IA rend la couverture complète prohibitive.

Pourquoi maintenant : trois forces

Les juges se multiplient. L'application de Merchant Center audite déjà l'accord page-flux-caisse article par article, avec la désapprobation préventive et le réexamen au niveau du compte comme barème de sanction. Les agents acheteurs étendent l'instance : dès qu'un agent transige pour le compte d'un acheteur via le Trusted Agent Protocol de Visa ou ses pairs, « achetable par une machine » cesse d'être un raffinement SEO et devient une condition préalable au revenu.

Le média payant est devenu une boîte noire, et la page est la dernière surface que vous possédez. Le marché français de la publicité digitale a atteint **12,4 milliards d'euros** en 2025, en hausse de 11 % sur un an, avec huit plateformes captant 76 % de la dépense totale et 83 % de la croissance de l'année. À l'échelle unitaire, le Digital Experience Benchmark 2025 de **Contentsquare** — tiré de plus de 90 milliards de sessions sur 6 000 sites — rapporte un *coût de la visite en ligne en hausse de 9 % sur un an et de 19 % sur deux ans*, face à un taux de conversion global en baisse de 6,2 % sur un an (une chute de 7,4 % pour les nouveaux clients en particulier) ; les marques qui ont davantage misé sur le paid social ont vu les conversions baisser de 10,6 % sur un an sur ce trafic. Les deux indicateurs se recourent : davantage d'euros affluent vers les mêmes enchères, chaque visite coûte plus cher à gagner, et moins de ces visites convertissent. Le secteur débat des causes — inflation des enchères, perte de signal, demande molle — et la réponse honnête est que l'agrégat ne peut pas être décomposé depuis un tableau de bord. Mais une tranche est entièrement auditable : quelle part du trafic que vous payez atterrit sur des produits qui ne peuvent pas être achetés. Avec les enchères déléguées à Performance Max et Advantage+, l'état de destination est la seule variable encore entièrement entre les mains de l'enseigne — et la moins instrumentée.

L'Europe a aveuglé sa propre analytique. Les bannières de consentement ont structurellement amputé la mesure par balises : lorsque « Tout refuser » apparaît à égalité avec

« Tout accepter », environ 60 % des utilisateurs de l'UE refusent — de sorte que les sessions que vous avez le plus besoin de comprendre sont en grande partie celles que vous n'êtes plus autorisé à observer depuis l'intérieur de votre propre site. Les robots côté serveur qui sautent la bannière vont trop loin dans l'autre sens : accepter les politiques de confidentialité révèle jusqu'à 70 traqueurs supplémentaires que la vue sans consentement n'observe jamais — aucune des deux vues ne correspond à ce que voit un véritable acheteur. Le bon instrument n'est ni l'un ni l'autre : il visite votre boutique comme un véritable acheteur, rencontre la bannière, prend une décision de consentement, et enregistre tout ce que l'acheteur voit à partir de là — y compris les différences entre les parcours acceptés et refusés. Dans l'UE, ce n'est pas une astuce de contournement. C'est la seule façon de lire votre vitrine complète telle que vos clients la voient réellement, une fois leur décision de consentement prise.

Trois forces, une conclusion : **l'état achetable des produits promus est devenu la surface la plus à enjeu et la moins instrumentée du commerce.**

Si vous agissez sur des données concurrentielles, la dérive de votre outil est votre perte

Si vous prenez des décisions à partir de données concurrentielles — alignement de prix, alignement de promotions, surveillance des stocks — lisez ceci deux fois.

Toute comparaison concurrentielle est une soustraction : votre prix moins le leur, votre disponibilité moins la leur, votre promotion contre la leur. Les soustractions amplifient l'erreur. Un outil qui lit mal le prix de votre concurrent de quelques pour cent — parce que sa page a été refondue, parce que son indicateur de rupture a changé, parce qu'un défi anti-bot a été pris pour un panier vide — vous donne une comparaison fautive d'une manière que vous ne pouvez pas mesurer. Vous n'obtenez pas une comparaison bruitée. Vous en obtenez une qui est fautive avec assurance, sans aucun signal que quelque chose a cassé.

Là où une machine juge votre boutique référence par référence et où votre budget média repose sur la comparaison, un instrument qui lit de la même manière à chaque fois cesse d'être une préférence d'ingénierie. Il devient la barre que vos données doivent franchir avant que quiconque ne les valide.

Il y a une deuxième raison pour laquelle cela compte — celle qu'un service achats remarque. Un système qui produit la même sortie à partir de la même entrée est quelque chose que votre service achats peut spécifier, auditer et placer sous SLA : **signer pour**. La Partie 2 se concluait sur le constat que ni le modèle ni l'humain ne fournit une couche de mesure qu'un acheteur peut signer. Celle-ci, oui.

Ce que cela vous permet de mesurer et que rien d'autre ne fait

Mettez les quatre problèmes ci-dessus côte à côte — des juges-machines qui sanctionnent les incohérences, des budgets pub qui meurent sur des produits cassés, des défaillances de concurrents que vous ne pouvez pas voir, des outils qui dérivent en silence — et le produit se dessine presque de lui-même : une couche de mesure pour les enseignes qui achètent du trafic, bâtie sur un moteur d'enregistrement qui ne dérive pas.

Concrètement, faire tourner le moteur sur votre catalogue promu produit trois chiffres que personne n'a de façon fiable aujourd'hui :

1. **Votre chiffre de dépense gaspillée.** Chaque euro de trafic payant que vous dépensez sur des produits que les acheteurs ne peuvent pas acheter — en rupture derrière un badge « en stock » périmé, prix du panier supérieur au prix annoncé, variante promue refusée à la caisse. Pas modélisé. Rejoué, horodaté, preuve attachée pour que vos équipes puissent reproduire et auditer.
2. **Les fenêtres de défaillance de votre concurrent.** Quand le best-seller promu d'un rival devient inachetable, ce rival sort de l'ensemble de considération pour ces requêtes — les acheteurs ne disparaissent pas, ils se redistribuent vers ceux qui restent debout, et votre fiche en convertit une part plus élevée parce qu'un substitut direct vient de quitter l'étal. Et si la référence du rival est désapprouvée plutôt que simplement en rupture, elle quitte aussi l'enchère publicitaire : moins d'enchérisseurs sur ces requêtes, un coût par clic plus bas pour vous. Le hic, c'est que la fenêtre ne vaut la peine d'être enchérie que si vous pouvez la voir tant qu'elle est encore ouverte et faire suffisamment confiance aux données pour y déplacer de la dépense. La soustraction de la section précédente est ce qui rend cette confiance possible ; un outil qui dérive vous donnerait de fausses fenêtres, et vous brûleriez du budget à courir après des fantômes.
3. **Accord flux-page-caisse, avant que l'horloge ne démarre.** C'est le problème de 3 h 47, attrapé selon vos conditions plutôt que celles de Google. Le moteur lit chaque page comme le robot la lit — le HTML brut que le serveur renvoie avant l'exécution de tout JavaScript — et comme un acheteur la lit — la page entièrement rendue — et compare les deux à votre flux et à votre caisse. Là où le prix vu par le robot et le prix vu par l'acheteur divergent se trouve exactement l'écart qui déclenche une désapprobation ; le moteur fait remonter ces références, la page capturée comme preuve, aussi souvent que vous le demandez, *avant* que l'horloge de 28 jours de Google ne démarre — pas après, quand la fiche a déjà disparu et que la page qui l'a causée est irreproductible.

« On ne peut pas simplement demander à une IA d'écrire l'outil d'extraction pour nous ? »

L'automatisation écrite par une IA reste de l'automatisation qui casse quand les pages changent — vous avez juste changé qui a écrit le code, pas ce qui se passe quand il échoue. Les casses

que vous détectez sont bon marché ; les silencieuses drainent votre budget média pendant que tout le monde fait confiance au tableau de bord. Les détecter est une tâche de vigilance — exactement le type de travail de surveillance répétitive dont la Partie 2 a montré que les humains échouent de façon mesurable.

Ce que vous obtenez si vous faites tourner ceci sur votre catalogue

Si vous pilotez le média payant ou l'e-commerce au sein d'une enseigne qui achète du trafic, vous avez maintenant vu les trois chiffres. Ce que vous obtenez, ce sont ces chiffres pour votre catalogue et vos principaux concurrents, avec des preuves que vos équipes peuvent reproduire et auditer. Ce que nous demandons, c'est un accès en lecture à GA4, Google Ads et aux diagnostics Merchant Center — pour que nous calculions votre chiffre de dépense gaspillée à partir de vos données, et non de nos hypothèses.

Le calendrier compte le plus dans le trimestre que vous vous apprêtez à planifier. Chaque heure où le best-seller d'un concurrent reste mort entre le Black Friday et les fêtes est la demande la moins chère de votre année — mais seulement si votre mesure tourne déjà, et fait déjà l'objet de confiance, avant que la saison ne commence. La personne qui entre dans la prochaine revue budgétaire avec le chiffre de dépense gaspillée n'argumente pas pour obtenir du budget ; le chiffre le fait.

Chaque mois où ceci reste non mesuré, le trafic payant continue d'atterrir sur des produits que personne ne peut acheter pendant que les fenêtres de défaillance de vos concurrents se referment sans être captées. Si votre trafic payant n'atterrit jamais dans une impasse et que vos concurrents ne sont jamais en rupture, vous n'avez pas besoin de nous. Sinon, ce sont **trente minutes** avec votre catalogue.

Conçu à l'INRIA. Aucun modèle dans le chemin de données. Aucun relecteur dans la boucle. Aucune exception.
