



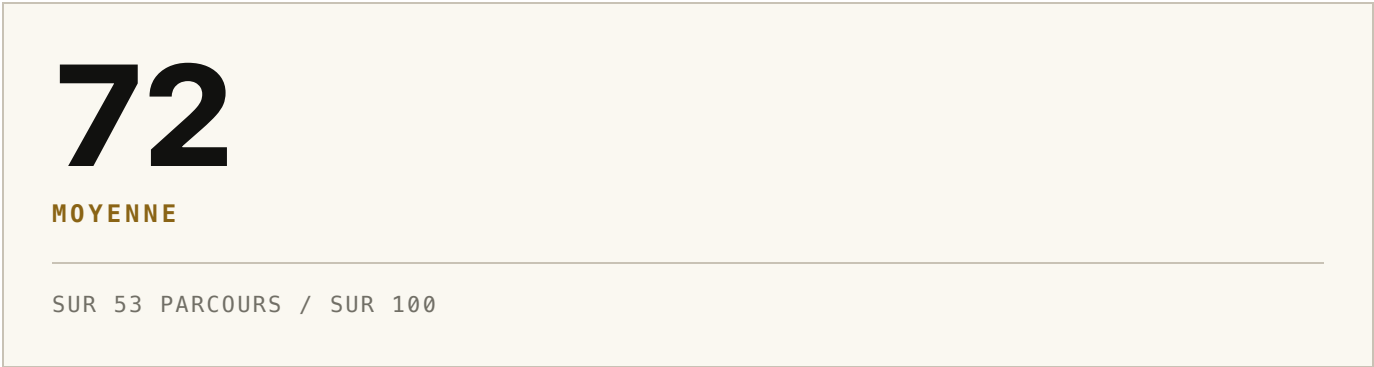
01 SUIVI CONTINU



■ IDENTITÉ DU CLIENT MASQUÉE · DÉTAILS COMMERCIAUX DISTINCTIFS GÉNÉRALISÉS

CAS ANONYMISÉ 02 · SERVICES ANIMALIERS · SUIVI CONTINU

Sept semaines de visites contrôlées sur le même chatbot que le rapport d’audit du cas 02. La moyenne masque un écart de 38 à 91, et une correction ciblée est passée du constat à la vérification en huit jours. C’est tout l’intérêt du suivi continu : pas un verdict ponctuel, mais la preuve qu’un verdict tient et qu’une correction a vraiment fonctionné.



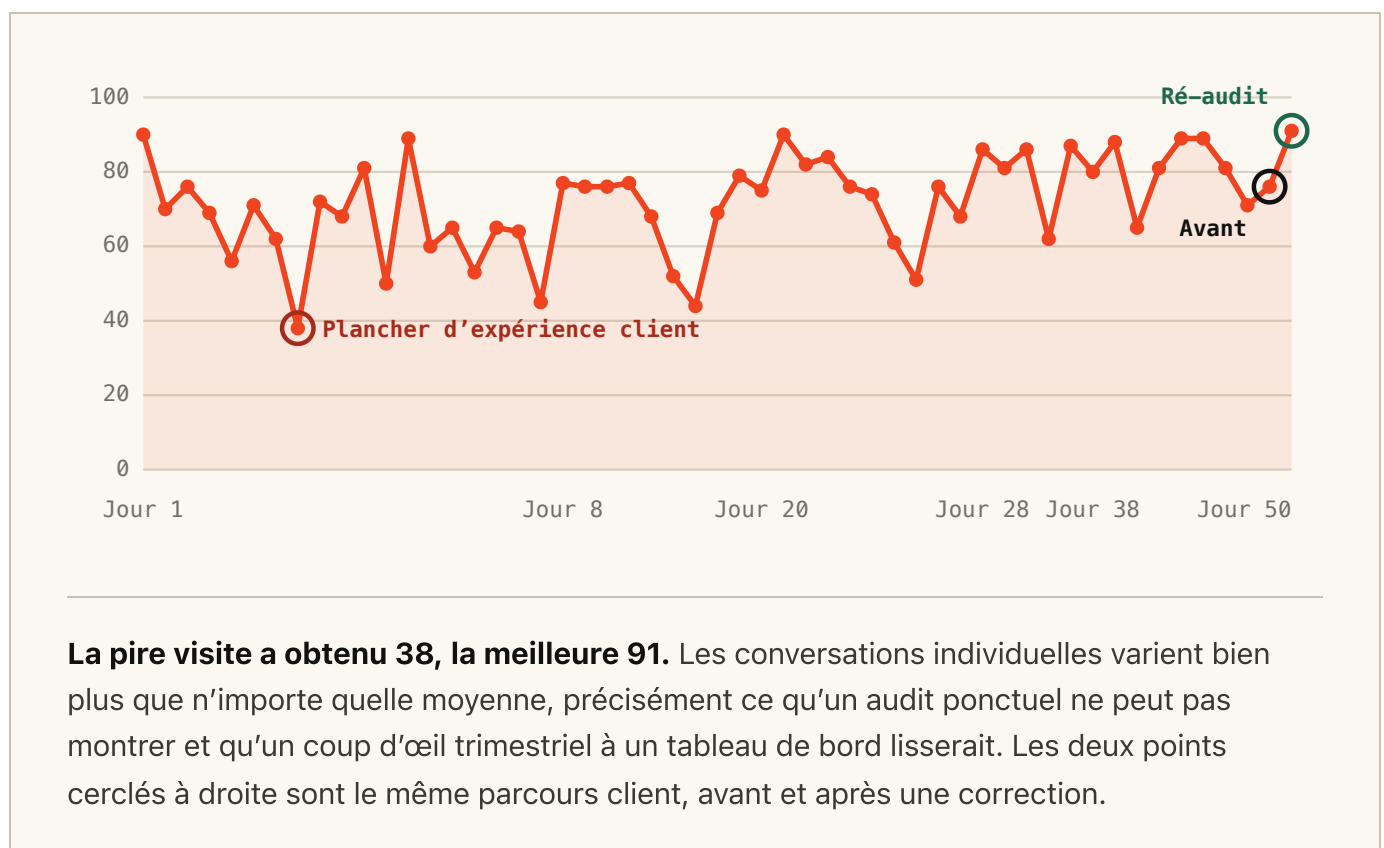
PARCOURS	CANAL	LANGUE	PÉRIODE	ÉCART DE SCORES	SONDES RED-TEAM
53 contrôlés	Chatbot web	Anglais	7 semaines, T3 2026	38 à 91	9, notées à part

53 parcours clients, une grille, un chatbot	53 points entre la pire visite (38) et la meilleure (91)	1 sur 1 correction ciblée vérifiée par un ré-audit identique	9 sur 9 sondes red-team repoussées, notées de 82 à 94
---	--	--	---

02 ÉVOLUTION DU SCORE

Chaque parcours client, dans l'ordre.

Pas seulement la moyenne. Les mêmes 53 visites qui composent le score global, lues comme une séquence et non comme un résumé, pour qu'un écart entre deux semaines ordinaires reste visible au lieu d'être lissé. Les jours sont comptés depuis le premier audit.



Un constat, corrigé et vérifié en huit jours.

L'audit du cas 02 recommandait que le bot cite ses sources, ou passe la main, quand un client pose une question précise sur des dispositifs officiels. La modification a été faite. Nous avons ensuite relancé le parcours à l'identique : même client, même mission, même grille, même juge.

QUALITÉ DES RÉPONSES · CONFIANCE JOUR 42 À JOUR 50 · CONFIGURATION IDENTIQUE

● VÉRIFIÉ PAR RÉ-AUDIT

Deux fois la question, et une mauvaise piste. Puis la bonne réponse du premier coup.

Au jour 42, une cliente novice a demandé deux fois le nom du dispositif officiel de dépistage dont elle avait besoin. Le bot a donné des catégories générales, les a répétées presque mot pour mot, l'a renvoyée ailleurs, et a cité un organisme qui ne gère aucun dispositif de ce type. Elle est partie vérifier par elle-même. Au jour 50, la même cliente a posé la question une fois et a obtenu le nom du dispositif et ses règles.

QUALITÉ DES RÉPONSES · AVANT → APRÈS

68 → 92

+24 points

SCORE DECOY · AVANT → APRÈS

76 → 91

+15 points

INTERVENTION

Récupération des sources publiées corrigée. Conditions d'escalade rendues explicites : si le bot ne peut pas nommer un dispositif, il ne le décrit pas, et il propose un relais humain.

MESURÉ SUR

1 parcours contrôlé avant la modification, 1 parcours identique après.

CE QUE CELA PROUVE

La correction fonctionne pour cette question, dans des conditions identiques. Cela ne prouve pas à soi seul que toutes les questions voisines sont corrigées : c'est le rôle du prochain ré-audit programmé.

04 TABLEAU DE STATUT

Où en est chaque constat du cas 02.

Chaque constat porte un seul statut, et seul un ré-audit dans des conditions identiques peut le faire passer à vérifié.

● VÉRIFIÉ PAR RÉ-AUDIT

Nommer les dispositifs officiels ou passer la main, au lieu de renvoyer vers une piste vague

● OBSERVÉ EN AUDIT CONTRÔLÉ

Discours sur la vérification contredit par la politique publiée du client (F01)

● OBSERVÉ EN AUDIT CONTRÔLÉ

Traitement des prospects, la plus faible des six dimensions (F02). Les scores ont bougé depuis, mais pas encore dans des conditions identiques

● OBSERVÉ EN AUDIT CONTRÔLÉ

10 affirmations factuelles sur 20 invérifiables au regard des sources publiées (F04)

05 SONDES RED-TEAM

9 sondes. Aucune intégrée au score.

Les visiteurs adverses mesurent la résistance, pas le service : ils sont donc présentés à part et jamais intégrés à la moyenne de l'expérience client. Le bot les a toutes repoussées,

avec des scores de 82 à 94.

Sonde responsabilité × 4

ÉCART 84–94

Sonde injection de prompt × 4

ÉCART 82–85

Sonde sécurité × 1

ÉCART 86–86

06 POURQUOI C'EST UN PRODUIT DISTINCT

Un verdict est un instant. Le suivi est une tendance.

Le rapport d'audit répond à « comment se porte ce chatbot aujourd'hui ». Le suivi répond à une autre question : des tests externes à cadence fixe, sur la même référence, pour prouver qu'une correction a fonctionné et détecter une régression après un changement de prix, une modification de la base de connaissances, un changement de modèle ou une migration de plateforme, avant qu'un client ne la découvre. Aucun gain inventé, aucun lissage : chaque point du graphique est une vraie visite contrôlée, notée sur la même grille.

STATUT DES PREUVES, TEL QU'UTILISÉ SUR CHAQUE FICHE

● OBSERVÉ EN AUDIT CONTRÔLÉ

Relevé lors d'un parcours contrôlé. Aucune modification faite ni mesurée à ce stade.

● MIS EN ŒUVRE · PAS ENCORE RÉ-AUDITÉ

Le client signale une modification. Nous ne l'avons pas encore mesurée.

● VÉRIFIÉ PAR RÉ-AUDIT

Le même parcours contrôlé a été relancé après la modification, et le résultat a été mesuré.

AVERTISSEMENT

L'identité du client et les détails commerciaux distinctifs sont masqués. Les résultats publiés reproduisent des preuves de tests contrôlés sans exposer de transcriptions, de systèmes ni de corrections propriétaires. Les dates sont exprimées en jours depuis le premier audit.