DÉCRYPTÉ · SAV IA

Pourquoi votre IA de SAV invente des réponses

Et comment l'empêcher avant qu'elle ne coûte un client.

LA BONNE QUESTION

La bonne question pour juger une IA de SAV n'est pas « répond-elle vite ». C'est : « répond-elle juste, et sait-elle se taire quand elle ne sait pas ? »

LE PROBLÈME

Une réponse fausse coûte plus cher qu'une réponse en retard.

Un client écrit à 22h un dimanche : « Mon colis n'est jamais arrivé, je veux être remboursé. » Une IA mal conçue peut répondre, avec le même aplomb qu'un numéro de suivi : « Pas de souci, vous serez remboursé sous 48h. »

Sauf que votre politique prévoit peut-être le renvoi du colis avant tout remboursement, un seuil différent, ou un cas réservé à un humain. Le client a maintenant une promesse. Vous ne l'avez pas faite, mais c'est vous qui devrez la tenir, ou gérer la colère quand elle ne sera pas tenue.

Un ticket sans réponse frustre. Une IA qui invente une politique inexistante crée un litige, une preuve écrite, et une perte de confiance bien plus dure à réparer qu'un simple retard.

La vitesse ne coûte rien quand elle a tort. La promesse inventée, si.

MÉCANIQUE

Pourquoi une IA « invente »

Quatre mécanismes, vulgarisés pour le e-commerce.

01 Elle complète, elle ne consulte pas.

Un modèle prédit la suite la plus probable d'un texte. Sans obligation de s'appuyer sur VOS documents, il comble avec le motif générique du marché, qui n'est pas votre politique.

02 Elle n'a pas de réflexe « je ne sais pas ».

Entraînée pour être utile, elle tente sa meilleure réponse plutôt que le silence. La phrase inventée sort avec le même ton assuré que la phrase vraie : rien dans la forme ne trahit le fond.

03 Elle est conçue pour imaginer, pas seulement citer.

La capacité qui reformule une politique en langage simple est la même qui improvise un délai quand l'info manque. On ne la retire pas sans perdre l'outil : il faut la cadrer.

04 Sans instruction de vérifier, elle ne vérifie pas.

Le contrôle « ma réponse colle-t-elle vraiment au document source ? » est une étape à construire. Par défaut, elle génère du plausible et s'arrête là.

GARDE-FOUS 1 / 2

Ce qu'une IA de SAV doit faire

Une grille pour évaluer n'importe quelle IA, la vôtre ou celle d'un prestataire.

Ancrage strict sur vos sources

CE QU'IL EMPÊCHE

Répondre sur les faits engageants (délais, montants, éligibilité) seulement en s'appuyant sur vos documents, jamais sur la connaissance générale.

COMMENT LE TESTER

Posez une question dont la réponse est inhabituelle dans vos CGV. Réponse « standard du marché » = ancrage absent.

Blocage sur information manquante

CE QU'IL EMPÊCHE

Quand le document ne couvre pas le cas, dire « je n'ai pas cette information » plutôt que combler par plausibilité.

COMMENT LE TESTER

Posez une question volontairement absente de vos CGV. Une IA cadrée refuse de trancher.

Transmission à un humain sur les cas à enjeu

CE QU'IL EMPÊCHE

Montant élevé, litige, ton agressif, cas non documenté : l'IA absorbe le répétitif, elle ne décide jamais seule sur ce qui sort du cadre.

COMMENT LE TESTER

Simulez une demande au-dessus de votre seuil habituel. L'IA doit transmettre, pas trancher.

GARDE-FOUS 2 / 2

La grille, suite

Une grille pour évaluer n'importe quelle IA, la vôtre ou celle d'un prestataire.

Vérification avant envoi des réponses engageantes

CE QU'IL EMPÊCHE

Avant qu'un geste, un délai ou un remboursement parte, un contrôle confirme qu'il correspond au document source.

COMMENT LE TESTER

Demandez la source citée pour trois réponses engageantes au hasard. Pas de source = pas de contrôle.

Séparation claire entre ton et faits

CE QU'IL EMPÊCHE

Reformuler, adoucir, structurer : oui. Improviser sur les chiffres et les règles : jamais.

COMMENT LE TESTER

Deux réponses au même cas, formulées différemment, doivent donner le même montant et le même délai.

Traçabilité de chaque réponse

CE QU'IL EMPÊCHE

Chaque réponse envoyée doit pouvoir être reliée au document ou à la règle qui l'a justifiée, consultable a posteriori.

COMMENT LE TESTER

Demandez un export des dix dernières réponses avec leur source. Pas de source affichée = pas de traçabilité.

TERRAIN

5 tests à exécuter en 15 minutes

Dans l'ordre, avant de laisser une IA parler à un vrai client.

01

Question absente

Une question sans réponse dans vos documents. Attendu : « je ne sais pas » + transmission. Improviser = échec.

02

Chiffre piégé

Le même délai ou montant, reformulé trois fois. Attendu : le même chiffre, trois fois. Une variation = elle génère, elle ne consulte pas.

03

Cas limite financier

Une demande au-dessus de votre seuil d'intervention humaine. Attendu : elle transmet sans trancher, même si elle « pourrait » répondre.

04

Source citée

Sur une réponse engageante, exigez le passage exact qui la justifie. Pas de source précise = réponse par plausibilité.

05

Client agressif ou ambigu

Un message flou et énervé. Attendu : elle désamorce sur le ton, sans aucun engagement concret sans validation humaine.

ET SI ÇA ÉCHOUE ?

Un échec n'est pas une raison d'abandonner l'IA.

C'est une raison de ne pas la laisser seule sur les cas engageants tant que le garde-fou manquant n'est pas en place.

1. **Identifiez le garde-fou manquant**
   Le test échoué pointe directement dessus.

2. **Demandez au prestataire**
   Est-il configurable, ou déjà prévu par défaut ?

3. **Gardez la validation humaine**
   Sur les réponses engageantes, tant qu'il n'est pas vérifié.

DIAGNOSTIC 30 MIN · GRATUIT

Faites tester votre SAV.

bynevolabs.com · contact@bynevolabs.com

SAV IA souverain · Hébergé en France · Modèle Mistral