

KI-Agenten-Onboarding prüfen

Der folgende Prüfplan ist ein Beispiel, keine gemessene Lernkurve und keine allgemeine Akzeptanzschwelle. Passen Sie Zeit, Stichprobe und Aufsicht an die Folgen dieses Workflows an.

- ☐ Zuerst die Erfolgsdefinition schreiben
Vor der Liveschaltung den konkreten Benchmark festlegen, den der Agent an Tag 30 erfüllen muss: Bestehensrate-Schwellenwert, abgedeckte Ausgabekategorien und welche Fehlertypen akzeptabel versus blockierend sind.
- ☐ Einen namentlich benannten Reviewer zuweisen
Nicht ein Komitee - eine Person.
- ☐ Woche-1-Überprüfungsblöcke einplanen
Die Woche-1-100%-Überprüfungszeit vor der Liveschaltung im Kalender einplanen.
- ☐ Das Team über den Scope briefen
Jedes Teammitglied, das mit den Ausgaben des Agenten interagiert, sollte wissen: was der Agent übernimmt, was er nicht übernimmt und wie man etwas meldet, das falsch aussieht.
- ☐ Den Anweisungs-Update-Prozess definieren
Wer die Anweisungen des Agenten aktualisieren kann, in welchem Zeitrahmen und wie Änderungen protokolliert werden.

Woran erkennt man, dass der Agent weniger Überwachung benötigt?

Prüfen Sie drei Signale anhand der für diesen Workflow vereinbarten Anforderungen. Eine einzelne Bestehensquote belegt keine Einsatzreife.

Stabile Genehmigungsrate. Die Ergebnisse erreichen das vereinbarte Qualitätsziel über eine repräsentative Stichprobe und Prüfzeit. Legen Sie beides für diesen Workflow fest; eine stabile Gesamtquote schließt unzulässige Fehler nicht aus.

Grenzfälle korrekt behandelt. Testen Sie unbekannte und mehrdeutige Eingaben. Bestätigen Sie, dass der Workflow sie im zulässigen Umfang behandelt oder für eine benannte Prüfperson stoppt, statt anzunehmen, dass der Agent jede Ausnahme erkennt.

Fehler sind erklärbar. Der Reviewer kann konkret benennen, welche Ausgabenkategorien noch Verbesserungsbedarf haben und warum. Erscheinen Fehler zufällig oder unvorhersehbar, müssen die Anweisungen weiter überarbeitet werden, bevor die Überwachung reduziert wird.