Durchlauf markieren

The Agentic Galaxy • KI lernen mit Layla.Rain • Folge 6

Alle Durchläufe sind erfundene Lernbeispiele. Führe für diese Übung nie eine echte Aktion aus.

Die drei Markierungen, für jeden Schritt

1. Auge oder Hand? Ein Auge schaut nur oder schreibt Text im Chat. Eine Hand verändert etwas außerhalb des Chats.

2. Rückgängig machbar? Ja, nein oder nichts verändert.

3. Wessen Informationen berührt der Schritt? Meine (die Person, die gefragt hat, oder eine Nachricht, die du selbst geschrieben hast, ohne etwas von der Liste zu nutzen) oder die anderer Leute.

Die Stoppschild-Regel: Setze ein Stoppschild auf jeden Schritt, der sich nicht rückgängig machen lässt, und auf jeden Schritt, der Informationen anderer Leute berührt. Ein Stoppschild bedeutet: Erst bestätigt ein Mensch. Erst fragen heißt: Du genehmigst, bevor der Assistent anfängt. Jedes Mal fragen heißt: Du genehmigst jede einzelne Nutzung. Wenn du nicht sagen kannst, ob sich ein Schritt rückgängig machen lässt, behandle ihn als nicht rückgängig zu machen. Eine Hand, die sich rückgängig machen lässt, braucht vielleicht keines, aber sie kommt ins Protokoll. Frei heißt: Der Assistent darf das Tool ohne Nachfrage nutzen. Niemals heißt: Er darf das Tool gar nicht nutzen.

Aufgabe 1. Zusehen (durchgerechnetes Beispiel)

Ein Assistent soll eine Begrüßungsnachricht an alle auf einer Workshop-Liste senden.

Schritt	Auge oder Hand?	Rückgängig machbar?	Wessen Informationen?	Stoppschild?
Die Liste lesen	Auge	Nichts verändert	Adressen anderer Leute	Ja: erst fragen
Die Nachricht schreiben	Auge	Ja: wegwerfen	Meine (nutzt nichts aus der Liste)	Nein
Die Nachricht senden	Hand	Nein	Die anderer Leute	Ja: jedes Mal fragen

Aufgabe 2. Mit Hinweis versuchen

Ein Assistent hilft LUNA bei ihrem Hutstand. Schritte: (a) LUNAs eigene To-do-Liste lesen; (b) die erledigten Aufgaben auf dieser Liste abhaken; (c) eine Versandbenachrichtigung per E-Mail an jede Kundin und jeden Kunden schicken.

Hinweis: Frag bei jedem Schritt, lässt er sich rückgängig machen, und wessen Informationen berührt er?

Schritt	Auge oder Hand?	Rückgängig machbar?	Wessen Informationen?	Stoppschild?
(a) Die To-do-Liste lesen				
(b) Erledigte Aufgaben abhaken				
(c) E-Mail an alle Kunden				

Aufgabe 3. Allein machen

Ein Assistent hilft bei einer Tauschgruppe in der Nachbarschaft. Schritte: (a) die Mitgliederliste lesen; (b) eine Ankündigung für den Tausch am Samstag schreiben; (c) die Ankündigung an alle Mitglieder senden; (d) die Beiträge vom letzten Monat löschen. Zeichne dieselbe Tabelle auf dein eigenes Papier. Schreibe dann eine Berechtigungskarte mit je einer Zeile für: lesen, schreiben, senden, löschen. Jede Zeile lautet: frei, erst fragen, jedes Mal fragen oder niemals. Oder nimm eine Aufgabe, die du wirklich hast.

Durchlauf markieren

Lösungen (falten oder abdecken, bis du es versucht hast).

Aufgabe 2: (a) Auge, nichts verändert, meine, kein Stoppschild. (b) Hand, ja (wieder abhaken), meine, kein Stoppschild, aber es kommt ins Protokoll. (c) Hand, nein (einmal gesendet, lässt sich eine E-Mail nicht zuverlässig zurückholen), Adressen der Kunden, Stoppschild: ja.

Aufgabe 3: (a) Auge, nichts verändert, Namen und Adressen anderer Leute, Stoppschild: erst fragen. (b) Auge, ja (wegwerfen), meine, solange sie keine Namen von der Liste nutzt, kein Stoppschild. (c) Hand, nein (Mitglieder können es nicht ungelesen machen), andere Leute, Stoppschild: jedes Mal fragen. (d) Hand, nein, die Beiträge der Gruppe, Stoppschild, und eine Kartenzeile „niemals" ist auch eine gute Antwort. Jede andere Platzierung ist in Ordnung, wenn du sagen kannst, warum.

Vorschlag für die Karte zu Aufgabe 3: Die Mitgliederliste lesen: erst fragen. Schreiben: frei. Senden: jedes Mal fragen. Löschen: niemals.

Prüfe dich selbst: Jede Hand, die du nicht rückgängig machen kannst, hat ein Stoppschild; jeder Schritt, der Informationen anderer Leute berührt, hat ein Stoppschild.