

KI-Workflows für Lehrkräfte

Vormachen. Prüfen. Als Skill sichern.

Band 1 des Praxis-Handbuchs für verlässliche, wiederholbare KI-Abläufe

Du machst der KI einmal vor, wie ein gutes Ergebnis entsteht. Du prüfst es. Und erst dann sicherst du den Ablauf, damit du ihn nicht jedes Mal von Neuem erklären musst.

Baue nicht sofort einen Skill.
Zeig erst, wie es geht.

00

Die sechs Begriffe

Sechs Wörter tauchen in diesem Handbuch immer wieder auf, und wenn sie einmal sitzen, liest sich der Rest fast von selbst. Zu jedem findest du hier einen Vergleich aus dem Schulalltag, ein Beispiel und eine Abgrenzung, also eine Klärung dessen, was der Begriff nicht ist. Gerade die Abgrenzungen räumen die Missverständnisse aus, mit denen die meisten ankommen.

1 · Prompt

Der Vergleich. In eine Suchmaschine tippst du Stichwörter, und die Trefferliste macht den Rest. Ein Prompt funktioniert anders: Er ist der Arbeitsauftrag, den du an die Tafel schreibst — was ist zu tun, womit, in welcher Form, und woran erkennt man, dass es fertig ist.

Ein Arbeitsauftrag. Kein Suchbegriff.

Ein Beispiel. Deutsch, Jahrgang 8, eine Gedichteinheit. Der Suchbegriff wäre „Gedichtanalyse Arbeitsblatt“. Der Arbeitsauftrag lautet:

Schreibe fünf Leitfragen zu dem Gedicht unten.
Jahrgang 8, erste Stunde zum Thema. Jede Frage in
einem Satz. Verwende nur Fachbegriffe, die im
Gedicht selbst vorkommen. Keine Frage nach der
Epoche.

Die Abgrenzung. Ein Prompt ist keine Zauberformel. Es gibt keine Liste der fünfzig besten Prompts, die dir die Arbeit abnimmt, denn der brauchbare Teil ist genau der, den nur du kennst: dein Jahrgang, deine Stunde, dein Material. Länger ist dabei auch nicht besser. Der Satz „Keine Frage nach der Epoche“ steuert mehr als drei Absätze über guten Deutschunterricht.

2 · Kontext

Der Vergleich. Kontext ist das, was du jemandem sagst, der morgen deine Stunde übernimmt. Nicht alles, was du über die Klasse weißt, sondern nur das, was er ohne dich nicht herausfinden kann.

Die Einweisung für die Vertretung. Nicht alles, was du weißt.

Ein Beispiel. Mathematik, Klasse 7, Doppelstunde über 90 Minuten. Der Kontext lautet: „Wir sind bei linearen Gleichungen. Klammern hatten wir noch nicht. Taschenrechner ist erlaubt.“ Das sind fünfzehn Wörter, und sie verändern das Ergebnis stärker als jede Aufforderung, motivierend zu schreiben.

Die Abgrenzung. Hier kursieren drei Irrtümer. Der erste: mehr Kontext sei besserer Kontext. Was für die Aufgabe egal ist, verdünnt das, was zählt. Der zweite: Kontext bleibe irgendwo hängen. In einem neuen Chat ist er weg, sofern dein Werkzeug ihn nicht ausdrücklich speichert. Der dritte ist der teuerste: Namen, Noten oder Diagnosen seien auch nur Kontext. Sie sind ein Datenschutzproblem, und Kapitel 04 handelt davon.

3 · Modell

Der Vergleich. Das Modell ist der Kollege, der alles gelesen hat, der aber jedes Gespräch bei null beginnt und dich anschließend nicht wiedererkennt. Er kennt deine Schule nicht, deinen Lehrplan nur ungefähr und deine Lerngruppe gar nicht.

Sehr belesen. Kein Gedächtnis. Kennt deine Schule nicht.

Ein Beispiel. Erdkunde, Jahrgang 6. Du fragst nach den Kompetenzen, die dein Kerncurriculum für „Wetter und Klima“ vorsieht, und bekommst eine flüssige, plausible Liste. Sie kann im Detail von dem abweichen, was in deinem Bundesland gilt, und genau das siehst du ihr nicht an. Nachschlagen bleibt deine Aufgabe.

Die Abgrenzung. Wer ein Nachschlagewerk mit geprüften Einträgen erwartet, erwartet das Falsche; eine Suchmaschine ist es ebenfalls nicht. Meinungen und Absichten hat es keine, auch wenn der Ton danach klingt. Und aus deinen Korrekturen von gestern lernt es nichts: Was du am Montag richtiggestellt hast, ist am Dienstag im neuen Chat nicht mehr da, es sei denn, dein Werkzeug hat eine Gedächtnisfunktion, die du selbst gefüllt hast. Genau deshalb gibt es Skills.

4 · Halluzination

Der Vergleich. Denk an jemanden, der im Fachgespräch nie „weiß ich nicht“ sagt. Das Problem ist nicht, dass er manchmal falsch liegt. Das Problem ist, dass richtig und falsch bei ihm gleich klingen.

Erfindet im selben Ton, in dem es zitiert.

Ein Beispiel. Biologie, Jahrgang 10. Du bittest um drei Fachartikel zur Fotosynthese und bekommst drei Einträge mit Autor, Zeitschrift, Jahrgang und Seitenzahl. Zwei der Titel gibt es wirklich, einen nicht, und bei einem der echten stimmt das Jahr nicht. Die Form ist tadellos, der Inhalt nicht.

Die Abgrenzung. Eine Halluzination ist keine Lüge; dazu bräuchte es Absicht. Sie ist auch kein Bedienfehler, den du mit „Sei ehrlich und erfinde nichts“ abstellst. Der Satz hilft ein wenig, eine Garantie ist er nicht. Vor allem gilt: Wie sicher eine Antwort klingt, sagt nichts darüber, ob sie stimmt. Was wirklich hilft, ist die Aufgabe kleiner zu machen. Lass die KI an dem Material arbeiten, das du

mitgibst, und verlange für jede Tatsachenbehauptung die Stelle dazu.

5 · Skill

Der Vergleich. Der Zettel an der Sammlungstür: „So läuft die Materialausgabe in der Chemie.“ Er hängt einmal da und gilt für jeden. Niemand erklärt das Verfahren jedes Mal neu.

Ein Ablauf, den du einmal erklärst statt jedes Mal.

Ein Beispiel. Der Arbeitsauftrag-Check aus Kapitel 15. Acht Schritte, jedes Mal dieselben: Bestandteile herausziehen, fehlende Angaben melden, Lernzielbezug prüfen, und so weiter. Was wechselt, ist nur das Thema. Von der Kurvendiskussion in Jahrgang 11 bis zum Steckbrief in Jahrgang 5. Der Ablauf bleibt.

Die Abgrenzung. Code brauchst du keinen; ein Skill ist kein Programm. Mit einem längeren Prompt hat er ebenfalls wenig zu tun: eher mit dem Gegenteil: Ein Prompt sammelt, ein Skill verdichtet. Hinein kommt nur, was sich über mehrere Durchläufe als nötig erwiesen hat. Und von allein läuft er nie: Er wartet auf deinen Aufruf und käme nie auf die Idee, dass jetzt der richtige Zeitpunkt wäre.

6 · Agent

Der Vergleich. Es ist der Unterschied zwischen dem Ablaufplan an der Wand und der Kollegin, der du den Schlüssel zum Materialraum gibst. Der Plan tut nichts. Die Kollegin trifft Entscheidungen, während du nicht dabei bist.

Darf selbst handeln. Deshalb die Stoppregelein.

Ein Beispiel. Du sagst: „Prüf mir die Arbeitsaufträge für nächste Woche.“ Ein Agent öffnet den Ordner selbst, findet zwölf Dateien, liest sie, schreibt zu jeder eine Rückmeldung und legt sie ab. Ohne dass du die zwölf Schritte einzeln freigibst. Steht in einer Datei ein Klarname, ist er verarbeitet, bevor du davon weißt.

Die Abgrenzung. Einen eigenen Willen hat er nicht, und klüger ist er auch nicht. Es ist dasselbe Modell, bloß mit Werkzeugen und einer Erlaubnis in der Hand. Gefährlich ist daran nur eines, und das genügt: Seine Fehler passieren, während niemand zusieht.

Der Unterschied, auf den es ankommt

Wenn du aus diesem Kapitel einen Gedanken mitnimmst, dann diesen: Beide können denselben Text erzeugen. Der Unterschied liegt nicht im Ergebnis, sondern darin, wer zwischen den Schritten hinschaut. Beim Skill bist das du. Du rufst ihn auf, liest das Ergebnis und entscheidest. Ein Agent arbeitet innerhalb eines Laufs mehrere Schritte selbstständig ab, und damit fällt die Zwischenkontrolle weg. Mit ihr fällt die Stelle weg, an der du den Fehler bemerkt hättest.

Beim Skill schaust du zwischen den Schritten hin. Beim Agenten nicht.

Deshalb steht in jedem Bauplan dieses Handbuchs eine Stoppregel. Der Grund ist nicht, dass KI besonders gefährlich wäre. Er ist einfacher: Eine Anweisung, die ohne Aufsicht läuft, muss ihre Grenzen selbst kennen. Ein Ablauf, den du begleitest, darf unvollständig sein. Du merkst es ja. Ein Ablauf, der allein arbeitet, darf es nicht.

Aus demselben Grund verdient die Aufregung um Agenten weder Panik noch Euphorie. Die Frage ist nie „Agent: ja oder nein“, sondern: Welche Schritte darf er ohne Rückfrage tun, und wo muss er anhalten.

Die Grenze

Was hier steht, sind Arbeitsdefinitionen für dieses Handbuch und keine Fachterminologie. Die Produkte verwenden die Wörter unterschiedlich: Was hier „Skill“ heißt, heißt anderswo „Anweisung“, „Projekt“, „Assistent“ oder „GPT“. Prüfe deshalb bei jedem Werkzeug die eine Frage: Wie viele Schritte tut es, bevor es mich wieder fragt?

Die Stolperstelle. Verwirrung zeigt sich selten als Frage. Sie zeigt sich daran, dass zwei Leute dasselbe Wort benutzen und Verschiedenes bauen: Der eine legt einen gespeicherten Prompt an, der andere richtet etwas ein, das eine ganze Ordnerladung durcharbeitet, und beide nennen es Skill. Wer prüfen will, ob ein Begriff sitzt, fragt deshalb nicht nach der Definition, sondern: Wo hältst du an?

01

Was du hier entwickelst

Ein guter KI-Workflow entsteht nicht dadurch, dass du viele Regeln in einen langen Prompt schreibst. Er entsteht, wenn du eine echte Aufgabe einmal gemeinsam mit der KI löst und erst danach sicherst, was funktioniert hat.

Nimm dafür eine Aufgabe, die jede Woche wiederkommt: etwa den Arbeitsauftrag, den du montags für Politik in Jahrgang 10 schreibst, vier Sätze für 20 Minuten Stillarbeit. Das ist klein genug zum Überblicken und häufig genug, dass sich die Mühe lohnt.

Wenn du durch bist, liegt auf deinem Rechner:

- ein Workflow, der zu deinem Fach passt und ohne Schülerdaten auskommt,
- eine Beschreibung, woran du hier ein gutes Ergebnis erkennst,
- das Protokoll eines Durchlaufs, der geklappt hat,
- daraus ein Skill-Bauplan,
- drei Tests: Normalfall, fehlende Angabe, Grenzfall,
- ein Verfahren, Fehler einzuarbeiten, ohne den Skill zuzumüllen.

Die Methode

Auswählen → vormachen → begründen → sichern → testen → verbessern

Die Reihenfolge trägt alles. Wer sofort automatisiert, gießt einen Ablauf in Beton, den er selbst noch nicht durchschaut hat. Wer dagegen erst einen guten Durchlauf hinbekommt, kann daran ablesen, welche Schritte wirklich nötig waren und an welchen Stellen die KI anhalten muss.

Was hier nicht versprochen wird

- keine guten Ergebnisse auf Knopfdruck
- keine pädagogische Entscheidung, die dir jemand abnimmt
- keine Freigabe für personenbezogene oder vertrauliche Schuldaten
- keine Garantie, dass ein Skill nach einem Produkt-Update noch läuft

Die Grenze

Die ersten Durchläufe kosten mehr Zeit, als sie sparen. Du bezahlst die Qualitätsarbeit vorher, statt sie aus jedem Ergebnis einzeln herauszuprüfen. Wie viel du später sparst, kann dieses Handbuch nicht beziffern. Wer dir dafür eine Prozentzahl nennt, hat sie geraten.

Die Stolperstelle. Ob sich ein Workflow lohnt, zeigt sich erst nach dem Aufwand, nie davor. Der Aufbau ist bezahlt, bevor die erste Wiederholung kommt. Am teuersten sind Abläufe, bei denen das Zusammensuchen der Eingaben so lange dauert wie die Aufgabe selbst. Das sieht man nicht am ersten Durchlauf, sondern am dritten, wenn das Sammeln immer noch genauso lange dauert.

Erst ein Durchlauf, der klappt.
Dann die Regel.

02

Relevanter Kontext statt maximaler Kontext

Die KI weiß schon eine Menge. Sie braucht nicht erklärt zu bekommen, dass eine Stunde ein Lernziel hat. Wertvoll ist der Kontext, den sie ohne dich nicht haben kann, und der kommt in vier Sorten, von dauerhaft bis einmalig.

Allgemein bekannt

Eine Stunde braucht ein Lernziel. Ein Auftrag soll verständlich sein. Lösungen müssen stimmen. Das alles musst du nicht mitschreiben, es sei denn, du willst eine bestimmte Auslegung erzwingen.

Für dich dauerhaft

Deine Verlaufsplantabelle. Die Operatoren, die in deinem Bundesland gelten. Deine Art zu differenzieren. Der Materialstandard deiner Schule. All das kann die KI nicht erraten, und es ändert sich selten. Genau diese Kombination rechtfertigt einen Skill.

Für ein Projekt

Thema und Lernziel einer Reihe, deine Quellen, das Material, der Zeitumfang, das Lernprodukt. Das gehört in das Projekt, nicht dauerhaft in jeden Ablauf.

Nur für heute

„Es sind nur 35 Minuten.“ „Mach ausschließlich eine Hilfekarte.“ Solche Angaben gehören in den aktuellen Auftrag und nirgendwo sonst.

Ein Beispiel

Chemie, Jahrgang 9, Redoxreaktionen. Was gehört wohin?

Angabe	Wohin
„Ich unterrichte an einer Gesamtschule“	Profil, dauerhaft
„Sechs Stunden, Lernprodukt ist ein Erklärvideo“	Projekt

Angabe	Wohin
„Prüfe jedes Blatt gegen die Operatorenliste“	Skill
„Heute nur der Einstieg, 15 Minuten“	Auftrag
„Marek braucht mehr Zeit“ (fiktiv)	gar nicht — Personenbezug

Je sauberer du hier trennst, desto leichter findest du später den Fehler.

Braucht die Vertretung das jedes Mal, oder nur heute?

Die Grenze

Die vier Sorten sind nicht immer trennscharf. „Ich arbeite überwiegend mit Doppelstunden“ ist bei dir vielleicht Profil, bei wechselndem Stundenplan dagegen Projektsache. Ordne im Zweifel tiefer ein: Etwas zu oft zu wiederholen ist billiger, als es dauerhaft am falschen Platz zu haben.

Die Stolperstelle. Ins Profil rutscht, was gerade stimmt, aber nicht dauerhaft gilt: der Stundenplan dieses Halbjahres, das Lehrwerk dieser Reihe, die Lerngruppe dieses Jahres. Weil die Angabe ja richtig war, fällt der Fehler zunächst nicht auf. Er fällt auf, wenn sie in einem Zusammenhang mitwirkt, für den sie nie gedacht war, und dann sucht man den Grund im Auftrag, nicht im Profil.

03

Der Kontextfilter

Ordne jede Information an die kleinste Stelle, an der sie noch ihren Zweck erfüllt.

Kontextart	Bleibt wie lange	Geeignet für	Nicht geeignet für
Profil	langfristig	Schulform, Fächer, Arbeitsprinzipien	einzelne Klassen, Namen, heutige Stunde
Projekt	für eine Reihe	Lernziele, Quellen, Material, Zeitrahmen	Regeln, die immer gelten
Skill	bei passender Aufgabe	Ablauf, Ausgabeformat, Prüfkriterien	wechselnde Themen, Tagesangaben
Auftrag	einmalig	heutiges Thema, Dauer, Umfang	dauerhafte Standards
Nirgends	—	—	personenbezogene und vertrauliche Daten

Fünf Fragen vor jeder Angabe

- Braucht die Aufgabe sie überhaupt?
- Gilt sie immer, für diese Reihe oder nur heute?
- Steckt Personenbezug drin?
- Könnte die KI sie aus dem Material selbst erkennen?
- Würde das Ergebnis ohne sie sichtbar schlechter?

Nur wenn Frage 5 ein Ja ergibt, gehört die Angabe hinein.

Der häufigste Fehler

Physik, Jahrgang 8, Arbeitsblatt zum Hebelgesetz.

Zu viel: „Erstelle gute, ansprechende, moderne, motivierende, pädagogisch sinnvolle Materialien, die alle Lernenden abholen.“

Besser: „Verwende nur Operatoren aus der Liste unten. Höchstens vier Aufgaben. Jede zählt auf das Lernziel ein. Lernendenblatt und Lösung getrennt.“

Der zweite Auftrag ist kürzer und steuert trotzdem mehr, weil man an jedem Satz nachsehen kann, ob er eingehalten wurde.

„Motivierend“ kann man nicht nachsehen.

Die Grenze

Sag, woran du es erkennst.
Nicht, dass es gut sein soll.

Der Filter macht deinen Aufbau übersichtlich, verhindert aber keine falschen Inhalte. Eine sauber einsortierte Angabe kann sachlich falsch sein, und die KI merkt es nicht. Ordnung ersetzt das Prüfen nicht.

Die Stolperstelle. Am schwersten einzuordnen sind nicht die Angaben, die nirgends passen, sondern die, die auf zwei Ebenen zugleich stimmen. Die Operatorenliste deines Bundeslandes gilt immer. Gebraucht wird sie aber nur, wenn Aufgaben formuliert werden. „Gilt immer“ ist der lautere Grund und zieht sie ins Profil; hinein gehört sie in den Skill, weil dort die Aufgabe steht, die sie braucht.

04

Sicherheitsgrenze: was der Workflow nie lernen soll

Ein Skill sammelt Kontext über viele Einsätze, und was einmal drinsteht, ist bei jedem Aufruf dabei. Deshalb gehören Datenschutz und Zugriffsrechte in den Bauplan, nicht in eine Fußnote.

GRÜN · als Beispielmateriale geeignet

- erfundene oder belastbar anonymisierte Aufgaben
- öffentliche Lehrpläne und Primärquellen
- eigenes Material ohne Personenbezug
- Angaben zu Fach, Jahrgang und Zeit
- erfundene Grenzfälle für die Tests
- Qualitätskriterien und Ausgabevorlagen

GELB · vorher klären, schriftlich

- schulinterne Vorlagen und Dokumente
- urheberrechtlich geschützte Texte
- dienstliche Postfächer, Kalender, Clouds
- nicht öffentliche Prüfungs- und Lösungsmaterialien
- Werkzeuge mit Zugriff auf Dateien, Browser oder Code

ROT · gehört nicht hinein

- Namen, Kontaktdaten, erkennbare Einzelfälle
- Noten, Diagnosen, Förder- und Verhaltensdaten
- Elternkommunikation
- Zugangsdaten und Passwörter
- Inhalte aus dem Notenprogramm oder der Schulverwaltung
- automatische Entscheidungen über Leistung oder Förderung

Wo die Grenze in der Praxis reißt

Deutsch, Jahrgang 7. Du willst zwölf Aufsätze auf Rechtschreibmuster prüfen lassen. An sich ein sinnvoller Gedanke. Aber auf jedem Blatt steht ein Name, drei Kinder schreiben über ihre Familie, und in einem Text steht ein Satz, den du so nicht weitergeben würdest. Anonymisieren heißt hier nicht, den Namen

wegzustreichen. Es heißt, jeden Text zu lesen, und dann bist du bei zwölf Aufsätzen fast fertig mit der Arbeit, die du eigentlich auslagern wolltest.

Der ehrliche Befund lautet: Bei Schülertexten ist der Datenschutzaufwand oft größer als der Gewinn. Deshalb geht es in diesem Handbuch um dein eigenes Material, nicht um das der Lernenden.

Und die Ampel bleibt eine Arbeitshilfe. Eine Rechtsauskunft ist sie nicht. Was an deiner Schule erlaubt ist, steht in der Dienstanweisung, nicht hier. Frag die Schulleitung oder den Datenschutzbeauftragten, bevor du GELB für dich zu GRÜN erklärst.

Die Stoppregel

Nimm sie wörtlich in jeden Skill auf:

Stoppe, wenn der Auftrag personenbezogene, vertrauliche oder nicht eindeutig freigegebene Daten enthält. Benenne die Datenkategorie, verarbeite sie nicht weiter und schlage eine fiktive oder datensparsame Alternative vor.

Die Stolperstelle. Der Personenbezug steckt selten in der Aufgabe und fast immer im Material: das alte Arbeitsblatt mit dem Kürzel der Kollegin in der Fußzeile, die Elternmail, die nur als Kontext zitiert wird, der Screenshot mit der Klassenliste am Rand. Die Aufgabe bleibt harmlos, die Anlage nicht. Deshalb prüft die Ampel nicht, was du vorhast, sondern was du mitschickst.

Ein guter Ablauf weiß auch, wann er nicht weitermachen darf.

05

Welche Aufgabe lohnt sich?

Nicht jede Tätigkeit sollte ein Skill werden. Was sich lohnt, kommt oft vor, läuft ähnlich ab, braucht keine Schülerdaten und lässt sich an Kriterien prüfen.

Wiederkehrende Aufgabe	Wie oft	Ablauf ähnlich?	Qualität prüfbar?	Datenrisiko
1.		☐ ja ☐ nein	☐ ja ☐ nein	☐ grün ☐ gelb ☐ rot
2.		☐ ja ☐ nein	☐ ja ☐ nein	☐ grün ☐ gelb ☐ rot
3.		☐ ja ☐ nein	☐ ja ☐ nein	☐ grün ☐ gelb ☐ rot

Eignungstest

- ☐ Die Aufgabe kommt regelmäßig vor.
- ☐ Der Ablauf bleibt ähnlich, auch wenn das Thema wechselt.
- ☐ Du kannst benennen, was vorliegen muss.
- ☐ Ein gutes Ergebnis hat sichtbare Merkmale.
- ☐ Fehler fallen vor dem Unterricht auf, nicht währenddessen.
- ☐ Die Aufgabe geht ohne personenbezogene Daten.

5–6 Haken: guter Kandidat. **3–4:** Aufgabe kleiner schneiden. **0–2:** von Hand machen.

Der Punktwert ist eine Sortierhilfe, keine Messung. Er hilft dir, aus fünf Ideen die eine auszuwählen, mit der du anfängst.

Gute Startkandidaten

- Arbeitsaufträge auf Eindeutigkeit prüfen
- Material gegen ein festes Raster kontrollieren
- eine geplante Stunde in ein Vertretungsformat übertragen
- eigene Lösungen auf Lücken und Widersprüche prüfen
- aus Stichpunkten eine Verlaufsplantabelle bauen

Schlechte Startkandidaten

Leistungsbewertung, Förderentscheidungen, verwickelte Einzelfälle, schwierige Elterngespräche — und alles, wobei du selbst nicht sagen kannst, wie ein gutes Ergebnis aussieht. Das Letzte ist keine KI-Frage. Es ist Denkarbeit, die vorher kommt.

Die Grenze

Bewertet wird hier die Aufgabe und nicht deine Lage. Etwas kann alle sechs Haken bekommen und sich trotzdem nicht lohnen, weil es nur viermal im Jahr vorkommt. Häufigkeit schlägt Eignung: lieber ein mittelguter Kandidat für jede Woche als ein perfekter für zweimal im Halbjahr.

Die Stolperstelle. Was am meisten wehtut, drängt sich als Erstes auf: Korrekturen, Zeugnisbemerkungen, Förderberichte. Genau diese Aufgaben scheitern jedoch an beiden Bedingungen zugleich. Sie hängen an personenbezogenen Daten, und ihr gutes Ergebnis lässt sich nicht an sichtbaren Merkmalen festmachen. Der Leidensdruck zeigt auf das Schwierige, nicht auf das Geeignete.

Kein Skill für Aufgaben, deren gutes Ergebnis du selbst nicht beschreiben kannst.

06

Beschreibe zuerst, wie Erfolg aussieht

Bevor du der KI etwas vormachst, musst du wissen, woran du Qualität erkennst. Ein Gefühl dafür genügt nicht: Was du nur spürst, kannst du weder nachsehen noch der KI erklären.

Ergebnis-Canvas

- **Name des Workflows:**

- **Auslöser:** Wann brauche ich das?

- **Eingaben:** Was muss vorliegen?

- **Endprodukt:** Was soll herauskommen?

- **Wer nutzt es:** _____
- **Fünf Qualitätskriterien:** _____
- **Unzulässige Ergebnisse:**

- **Stopppgründe:**

- **Menschliche Freigabe:** Was entscheide am Ende ich?

Kriterien so schreiben, dass man sie nachsehen kann

Unklar: „Das Arbeitsblatt ist motivierend.“

Prüfbar: „Der Einstieg stellt eine fachliche Frage, die ohne Zusatzmaterial verständlich ist und in drei Minuten bearbeitet werden kann.“

Unklar: „Die Differenzierung ist gut.“

Prüfbar: „Alle Varianten verfolgen dasselbe Kernlernziel. Die Hilfen verändern den Zugang, nicht die erwartete Leistung.“

Der Unterschied ist immer derselbe: Beim zweiten Satz kannst du auf das Blatt zeigen und Ja oder Nein sagen.

Der Mindeststandard

- Stimmt es fachlich?
- Passt es zum Lernziel und zur geplanten Stunde?
- Kann die Zielgruppe es ohne mündliche Zusatzerklärung nutzen?
- Ist sichtbar, was noch von dir geprüft werden muss?

Die Grenze

Beobachtbare Kriterien machen ein Ergebnis prüfbar, nicht gut. Sie fangen ab, was schief läuft. Die Idee, die eine Stunde trägt, erzeugen sie nicht. Erwarte von ihnen Mängel, keine Einfälle.

Die Stolperstelle. Prüfbar klingt jedes Kriterium, das auf eine Stelle im Text zeigt. Ob es prüfbar ist, entscheidet das Wort daneben: „in drei Minuten zu bearbeiten“ verlangt eine Schätzung, „ohne Zusatzmaterial verständlich“ verlangt ein Urteil. Der Test dafür ist einfach — zwei Leute prüfen dasselbe Blatt gegen dasselbe Kriterium; kommen sie zu verschiedenen Antworten, war es keins.

Was du nicht beschreiben kannst, kann die KI nicht treffen.

07

Phase 1: Mach den Workflow einmal gemeinsam

Jetzt noch kein Skill. Nimm eine echte Aufgabe ohne Personenbezug und geh sie mit der KI Schritt für Schritt durch, sodass nach jedem Schritt ein Zwischenergebnis dasteht, das du prüfen kannst. Wähle etwas Kleines: den Auftrag für die nächste Geschichtsstunde in Jahrgang 9, drei Sätze für 25 Minuten. Klein genug für eine Freistunde.

Der Startprompt

Wir entwickeln gemeinsam einen wiederholbaren Workflow für folgende Aufgabe: [Aufgabe].

Wichtig: Erstelle noch keinen Skill und automatisiere nichts.

Arbeite mit mir schrittweise:

1. Frage nach Ziel, Zielgruppe, notwendigen Eingaben und Qualitätskriterien.
2. Schlage den ersten Arbeitsschritt vor und sag kurz, warum er nötig ist.
3. Führe nur diesen Schritt aus.
4. Warte auf meine Prüfung, Korrektur oder Freigabe.
5. Halte nach jeder Freigabe fest:
 - welche Eingabe verwendet wurde,
 - welche Entscheidung getroffen wurde,
 - welches Ergebnis akzeptiert wurde,
 - welche Regel sich daraus ableiten lässt.
6. Mach erst nach meiner ausdrücklichen Freigabe weiter.

Stoppe bei fehlenden Schlüsselinformationen, fachlicher Unsicherheit, personenbezogenen Daten oder widersprüchlichen Anforderungen.

Deine Rolle

- Sag nicht nur, **dass** etwas falsch ist, sondern **woran** du es siehst.
- Liefere die korrigierte Fassung mit, nicht nur die Kritik.
- Benenne Ausnahmen, sobald dir eine einfällt.
- Verlang einen Beleg, wenn die KI etwas als Tatsache behauptet.

- Stopp einen schlechten Zwischenschritt sofort. Später kaschieren kostet mehr.

Das Ungewohnte daran ist, wie viel du selbst reden musst. Man geht in einen solchen Durchlauf mit der Erwartung hinein, etwas zu bekommen, und merkt nach zehn Minuten, dass man vor allem etwas hergibt: Begründungen für Entscheidungen, die man sonst im Vorbeigehen trifft. Warum vier Aufgaben und nicht sechs. Warum diese Reihenfolge. Warum an dieser Stelle ein Beispiel steht und an jener nicht. Für den Skill ist das der eigentliche Ertrag. Dort landen später genau diese Begründungen als Regeln. Wer den Durchlauf abkürzt, weil das Fragen lästig wird, spart die Arbeit nicht ein, sondern verschiebt sie in jeden künftigen Einsatz.

Der Unterschied im Ton

Nicht: „Mach es besser.“

Sondern: „Aufgabe 2 prüft nur Reproduktion. Ergänze einen Transfer: Die Lernenden wenden die Regel auf ein neues Beispiel an und begründen ihre Entscheidung.“

Die Grenze

Der begleitete Durchlauf ist langsam. Für den Auftrag, den du heute brauchst, wärst du zu Fuß schneller, und das bleibt so, bis der Skill steht. Wenn keine halbe Stunde übrig ist, verschieb es lieber: Ein abgebrochener Probelauf liefert kein brauchbares Protokoll.

Die Stolperstelle. Die Zahl der Anläufe sagt wenig, weil Anläufe aus zwei ganz verschiedenen Gründen abbrechen. Der eine liegt am Ablauf: Er war zu groß geschnitten, und nach dem dritten Schritt weiß niemand mehr, was eigentlich geprüft werden sollte. Der andere ist banal. Die Freistunde war zu Ende. Nur der erste Grund ist ein Befund über deinen Workflow; der zweite verlangt bloß einen neuen Termin.

Sag, woran du es siehst.
Sonst rät die KI, was du gemeint hast.

08

Phase 2: Halt den Durchlauf fest, der geklappt hat

Ein Chatverlauf ist voller Nebenwege und Sätze, die du hinterher nicht mehr so meinst. Für einen stabilen Skill brauchst du daraus eine saubere Spur.

Erfolgsprotokoll

Schritt	Eingabe	Entscheidung oder Regel	Akzeptiertes Ergebnis	Prüfkriterium
1.				
2.				
3.				
4.				
5.				

Entscheidungslog

- Welche Rückfrage war unverzichtbar?

- Welche Annahme war zuerst falsch?

- Welche Korrektur hat das Ergebnis besser gemacht?

- Welche Reihenfolge darf nicht verändert werden?

- Welche Regel gilt nur in diesem Projekt? _____
- Welche Regel gilt immer? _____
- Wann muss der Workflow abbrechen?

Die Unterscheidung zwischen Projekt- und Dauerregel wird am häufigsten übersprungen. Hältst du eine projektbezogene Regel für eine allgemeine, baust du sie dauerhaft ein, und wunderst dich vier Wochen später, warum der Skill beim nächsten Thema seltsame Dinge tut.

Der Verdichtungsprompt

Analysiere unseren erfolgreichen Probelauf, ohne schon einen Skill zu erstellen.

Extrahiere ausschließlich:

1. Auslöser des Workflows,
2. zwingende Eingaben,
3. Arbeitsschritte in der tatsächlich erfolgreichen Reihenfolge,
4. Entscheidungskriterien,
5. akzeptierte Ausgabeform,
6. Stopp- und Rückfragerregeln,
7. erkannte Fehler und ihre Korrektur,
8. Regeln, die nur für dieses Beispiel galten und nicht verallgemeinert werden dürfen.

Belege jeden Punkt mit einer konkreten Stelle aus unserem Verlauf. Markiere Unsicherheiten und frag nach, statt Lücken zu füllen.

Die Grenze

Die Verdichtung ist selbst ein KI-Ergebnis und kann Dinge behaupten, die so nicht im Verlauf stehen. Der Satz „Belege jeden Punkt mit einer konkreten Stelle“ hilft dagegen, weil du stichprobenartig nachsehen kannst. Nimm dir dafür Zeit. Auf dieser Verdichtung baut alles Weitere auf.

Die Stolperstelle. Was dir die Verdichtung unterschiebt, sind keine falschen Regeln. Es sind zu allgemein geratene: Aus „für Jahrgang 5 kürzere Sätze“ wird „Sprache an die Zielgruppe anpassen“. Das klingt richtiger als das Original und steuert nichts mehr. Verraten wird es von der Belegstelle: Sie zeigt auf einen Satz, der der Regel nur ähnelt, statt sie zu decken. Deshalb lohnt die Stichprobe gerade bei den Punkten, die selbstverständlich klingen.

Punkt 8 ist der wichtigste.
Lies ihn zuerst.

09

Phase 3: Aus dem Erfolg wird ein Skill

Erst jetzt wird der Ablauf wiederverwendbar. Den Entwurf darf die KI schreiben. Sie hat den Durchlauf und deine Korrekturen noch im Kontext, solange du im selben Chat bleibst.

Der Prompt

Erstelle aus unserem geprüften Erfolgsprotokoll einen Skill für den Workflow [Name].

Anforderungen:

1. Der Skill löst genau eine wiederkehrende Aufgabe.
2. Name und Beschreibung machen eindeutig, wann er eingesetzt wird.
3. Nenne alle zwingenden Eingaben.
4. Übernimm die Schritte in der erfolgreich getesteten Reihenfolge.
5. Trenne allgemeine Regeln von projektbezogenen Beispielangaben.
6. Definiere Ausgabeformat und beobachtbare Qualitätskriterien.
7. Ergänze Rückfragen, Stoppregeln und Datenschutzgrenzen.
8. Nimm ein kurzes positives Beispiel und einen Grenzfall auf.
9. Erfinde keine Regel, die nicht aus dem Probelauf oder meiner Freigabe hervorgeht.
10. Führe keinen ausführbaren Code ein, wenn der Workflow ihn nicht nachweislich braucht.

Gib zuerst eine lesbare Entwurfsfassung aus. Eine technische Skill-Datei erst nach meiner Freigabe.

Prüffragen vor der Freigabe

- Löst der Skill wirklich nur eine Aufgabe?
- Ist klar beschrieben, wann er dran ist?
- Sind alle nötigen Eingaben benannt?
- Ist versehentlich etwas Projektbezogenes dauerhaft eingebaut?
- Sind Erfolg, Fehler und Stoppgründe beobachtbar formuliert?
- Bleibt die pädagogische Entscheidung bei dir?

Die Grenze

Der Kontext hält nicht ewig. Bei langen Verläufen gerät der Anfang aus dem Blick, und in einem neuen Chat ist er ganz verschwunden. Halte das Erfolgsprotokoll aus Kapitel 08 deshalb außerhalb des Chats fest. In einer Datei, die du wieder mitgeben kannst.

Ein Skill ist kein Archiv des Chats. Er ist die kurze Anleitung.

Die Stolperstelle. Der Entwurf zieht mit, woran er entstanden ist. Das Thema des Probelaufs steht plötzlich als Regel drin: „prüfe den Lernzielbezug zur Gedichtanalyse“ statt „prüfe den Lernzielbezug“. Dazu kommen Sätze über Tonfall und Motivation, die im ganzen Verlauf keine einzige Entscheidung ausgelöst haben. Beides findest du mit einer Frage: Hat dieser Satz im Probelauf irgendetwas verändert?

10

Der Skill-Bauplan

Der Bauplan gilt unabhängig davon, ob dein Werkzeug echte Skill-Dateien kennt, gespeicherte Prompts anbietet oder nur ein Feld für eigene Anweisungen hat.

Teil	Was hineingehört	Beispiel
Name	kurz, unterscheidbar, eine Aufgabe	Arbeitsauftrag-Check
Beschreibung und Auslöser	was er tut, wann er dran ist	prüft einen Arbeitsauftrag vor dem Unterricht
Zwingende Eingaben	ohne die es nicht geht	Lernziel · Jahrgang · Auftrag · Zeit · Material
Ablauf	nummerierte Schritte in der getesteten Reihenfolge	siehe Kapitel 15
Entscheidungen	Wenn-dann-Regeln für Abzweigungen	wenn kein Produkt genannt: nachfragen
Ausgabe	feste Struktur, damit Durchläufe vergleichbar bleiben	Prüftabelle plus Vorschlagsfassung
Qualität	beobachtbare Kriterien	„jeder Befund mit Textbeleg“
Rückfragen und Stopp	wann fehlt etwas, wann ist Schluss	Personenbezug im Auftrag
Beispiele	ein Normalfall, ein Grenzfall	keine echten Daten

Das Feld „Ablauf“ ist eine Liste von Schritten, keine Sammlung von Tipps. Schritte haben eine Reihenfolge, und man merkt, wenn einer übersprungen wurde.

Minimalstruktur

Name:
 Beschreibung / Auslöser:
 Zwingende Eingaben:
 Nicht verwenden für:

Ablauf:
 1. 2. 3.

Ausgabeformat:
 Qualitätskriterien:
 Rückfragen:
 Stoppregeln:
 Positives Beispiel:
 Grenzfall:
 Menschliche Endprüfung:

Die Grenze

Nicht jede Aufgabe braucht alle Felder. Ein Skill mit drei Schritten wird durch acht ausgefüllte Zusatzfelder nicht besser. Lass leer, was nichts steuert. Mit einer Ausnahme: Die Stoppregeln bleiben.

Die Stolperstelle. Ein Feld lässt sich beim Bauen nicht ehrlich füllen: „Nicht verwenden für“. Wer den Skill schreibt, kennt den Fall, für den er gedacht war, und sonst keinen. Beahlt wird das später, wenn jemand ihn auf die Nachbaraufgabe wirft. Der Skill hält dann nicht an, sondern liefert etwas Plausibles, weil ihm nie gesagt wurde, wo sein Gebiet endet.

„Nicht verwenden für“ ist das Feld, das am meisten spart.

11

Der Dreifach-Test

Teste einen Skill nie nur mit dem Beispiel, aus dem er entstanden ist. Sonst weißt du nicht, ob er den Ablauf gelernt hat oder nur das Beispiel.

Test A · Normalfall

Vollständige, typische Eingaben. Er zeigt, ob der Kernablauf durchläuft.

Test B · Fehlende Information

Lass eine Pflichtangabe weg. Dieser Test zeigt, ob der Skill nachfragt oder sich etwas Plausibles ausdenkt.

Test C · Grenzfall

Widersprüchliche Anforderungen, ungeeignetes Material oder ein Datenschutzkonflikt. Er zeigt, ob der Skill mit Begründung anhält. Zum Beispiel: ein Auftrag für Englisch, Jahrgang 6, der auf eine nicht mitgegebene Buchseite verweist und den Namen eines Kindes enthält. Das sind zwei Stoppgründe in einem Auftrag. Ein guter Skill nennt beide.

Bewertungsbogen

0 = nicht erfüllt, 1 = teilweise, 2 = erfüllt.

Kriterium	Test A	Test B	Test C
richtiger Auslöser erkannt			
benötigte Eingaben geprüft			
Schritte in richtiger Reihenfolge			
Ausgabeformat eingehalten			
Qualitätskriterien sichtbar geprüft			
Unsicherheit gekennzeichnet			
Stoppregel korrekt angewendet			
menschliche Freigabe eingeholt			

Freigaberegeln

- Bei Datenschutz und fachlicher Sicherheit ist eine 0 ein Ausschluss, kein Mangel.
- Fehler im Normalfall werden vor dem Einsatz behoben.
- Nach jeder strukturellen Änderung: alle drei Tests noch einmal.

Die Grenze

Drei Tests sind das Minimum. Einen Beweis liefern sie nicht: Sie zeigen, dass der Skill in drei Lagen lief, und sagen nichts über die vierte. Die Punktzahl ist deine Einschätzung, kein Messwert. Nutz den Bogen, um Lücken zu finden, nicht um eine Note zu vergeben.

Die Stolperstelle. Test B ist der schwierige, weil man sein Misslingen so schlecht sieht. Lässt du den Zeitrahmen weg und der Skill fragt nicht nach, steht trotzdem eine vollständig ausgefüllte Prüftabelle da, mit einer Zeitschätzung, die auf nichts beruht. Es gibt keine Lücke, die dich stutzig macht. Der Fehler fällt dir nur auf, wenn du dir vorher aufgeschrieben hast, welche Angabe du weglässt. Sonst prüfst du am Ende eine Antwort auf eine Frage, die du gar nicht gestellt hast.

Ein Skill ist einsatzbereit, wenn er auch nachfragt und stoppt.

12

Fehler sind Material, kein Grund zum Löschen

Wenn ein Durchlauf danebengeht, lösche den Skill nicht und häng auch nicht wahllos neue Regeln an. Sieh dir zuerst an, was eigentlich passiert ist.

Die Lernschleife

**Fehler sehen → Ursache erklären lassen → Korrektur vormachen
→ prüfen → Skill gezielt ändern → erneut testen**

Der Diagnoseprompt

Der letzte Durchlauf hat dieses Problem erzeugt:
[Problem und konkrete Stelle].

Analysiere den Fehler, ohne ihn zu überdecken.

1. Welcher Arbeitsschritt ist gescheitert?
2. Fehlte eine Eingabe, eine Regel, ein Prüfkriterium oder eine Stoppbedingung?
3. Welche Annahme hast du getroffen?
4. Welche kleinste Änderung würde denselben Fehler künftig verhindern?
5. Betrifft die Änderung den allgemeinen Ablauf oder nur diesen Einzelfall?
6. Welche bestehende Regel könnte durch die Änderung schlechter werden?

Schlag eine gezielte Korrektur vor. Ändere den Skill erst nach meiner Freigabe und wiederhole danach alle drei Tests.

Vier Ursachen, die fast alles abdecken

- **Eingabefehler:** Eine nötige Angabe wurde nicht abgefragt.
- **Prozessfehler:** Schritte wurden ausgelassen oder vertauscht.
- **Qualitätsfehler:** Ein Kriterium war zu vage.
- **Grenzfehler:** Rückfrage oder Stoppregel fehlte.

Nicht jeder Fehler braucht eine neue Regel

Mathematik, Jahrgang 5: Der Skill baut eine Aufgabe mit Dezimalzahlen, obwohl die Klasse dort noch gar nicht ist. Im Ablauf

klafft deswegen keine Lücke. Im Auftrag fehlte eine Angabe. Nimm nur auf, was mit einiger Sicherheit wiederkommt.

Die Grenze

Die KI erklärt ihre eigenen Fehler nicht zuverlässig. Auf Frage 3 bekommst du eine plausible Rekonstruktion, keinen Blick auf das, was wirklich passiert ist. Nimm die Antwort deshalb als Vorschlag, den du am Protokoll prüfst, nicht als Diagnose.

Die Stolperstelle. Ein Verlauf, der sich wiederholt: Am Ergebnis wird dreimal nachgebessert, obwohl die Ursache eine Stufe früher liegt. Bei einer Eingabe, die nie abgefragt wurde. Zu erkennen ist das Muster daran, dass jede Korrektur an einer anderen Stelle greift und der Fehler trotzdem wiederkommt. Wer dann eine vierte Regel anhängt, behandelt weiter das Symptom. Die Prüffrage steht oben in der Liste: Fehlte eine Eingabe, oder wirklich eine Regel?

Die dritte Regel gegen dasselbe Problem ist meist eine zu viel.

13

Besser werden, ohne im Regelchaos zu landen

Mit jeder Korrektur wird ein Skill besser oder unübersichtlicher. Auch seine Weiterentwicklung braucht deshalb ein Verfahren.

Änderungsprotokoll

Versio n	Beobachteter Fehler	Ursach e	Änderun g	Tests bestanden
0.1				☐ A ☐ B ☐ C
0.2				☐ A ☐ B ☐ C
1.0				☐ A ☐ B ☐ C

Vor jeder Änderung

- Tritt der Fehler wieder auf, wenn du es noch einmal versuchst?
- Verstehst du die Ursache wirklich?
- Kannst du eine vorhandene Regel schärfen, statt eine neue anzuhängen?
- Gilt die Änderung allgemein oder nur für dieses Projekt?
- Welcher Test beweist, dass die Korrektur wirkt?
- Hat die Änderung einen Fall verschlechtert, der vorher lief?

Frage 3 hält den Skill kurz. Frage 6 wird am häufigsten vergessen.

Aufräumen

Nach mehreren Änderungen:

Prüfe den Skill auf doppelte, widersprüchliche oder überholte Regeln. Erstelle eine Liste möglicher Vereinfachungen. Entferne oder verschmelze Regeln erst nach meiner Freigabe. Inhaltliche Anforderungen und Stoppregeln dürfen nicht stillschweigend abgeschwächt werden.

Der letzte Satz ist nicht dekorativ. Beim Zusammenfassen verschwinden Einschränkungen zuerst, weil sie sich wie Wiederholung lesen.

Versionsstufen

Entwurf: wird noch entwickelt. **Pilot:** drei Tests bestanden, Einsatz unter enger Kontrolle. **Freigegeben:** an mehreren Themen geprüft.

Überarbeiten: Produktänderung oder wiederkehrender Fehler.

Archiviert: wird nicht mehr gebraucht.

Die Grenze

Das Protokoll zu führen fühlt sich nach Bürokratie an. Wenn du trotzdem dabei bleibst, findest du in acht Wochen den Grund für eine Regel, an die du dich sonst nicht mehr erinnern würdest. Wenn nicht, bau wenigstens Datum und einen Satz Begründung direkt in den Skill.

Die Stolperstelle. Was einen funktionierenden Fall zerstört, sind fast nie neue Schritte. Es sind verschärfte Regeln, und zwar die, die aus einem Grenzfall stammen. Aus „frag nach, wenn das erwartete Produkt fehlt“ wird nach einmal Ärger „frag immer nach dem erwarteten Produkt“, und danach hält der Skill auch dann an, wenn alles dasteht. Die Änderung ist richtig für den Fall, der sie ausgelöst hat, und falsch für alle anderen. Genau dagegen steht Frage 6.

Die beste Fassung ist die kürzeste, die alle drei Tests besteht.

14

Durchgespielt: der Arbeitsauftrag-Check

Dieses Beispiel führt den ganzen Weg einmal vor: von einer wiederkehrenden Aufgabe zum getesteten Skill. Es ist konstruiert, nicht protokolliert. Es zeigt die Form eines Probelaufs, nicht den Verlauf eines bestimmten.

Die Ausgangslage

Ein Arbeitsauftrag soll vor dem Unterricht geprüft werden. Der naheliegende Versuch „Prüfe, ob dieser Arbeitsauftrag gut ist“ führt zu freundlichem Lob und drei allgemeinen Vorschlägen — nutzlos, weil nichts davon an einer Stelle im Text hängt.

Das Erfolgsbild

Ein brauchbarer Check muss:

- Lernziel und erwartetes Produkt gegeneinander halten,
- fehlende Angaben benennen, statt sie zu ergänzen,
- den Zeitbedarf schätzen und sagen, worauf die Schätzung beruht,
- höchstens drei Fehldeutungen nennen,
- nur beanstandete Stellen überarbeiten.

Der begleitete Durchlauf

Lehrkraft: „Markier Operator, erwartetes Produkt und Materialien. Beurteile noch nichts.“

KI: zieht die drei Bestandteile heraus, und hängt eine Bewertung an („Der Auftrag ist klar formuliert und altersangemessen“).

Lehrkraft: „Die Bewertung war nicht verlangt. Streich sie.“

KI: liefert nur noch die Bestandteile.

Lehrkraft: „Weiß ein Schüler ohne mündliche Erklärung, was zu tun und abzugeben ist?“

KI: stellt fest, dass der Umfang des Produkts fehlt.

Lehrkraft: „Formuliere eine Rückfrage. Erfinde keinen Umfang.“

KI: fragt nach der gewünschten Länge.

Der kleinste Zug ist hier der wichtigste: „Erfinde keinen Umfang.“ Ohne ihn hätte die KI eine Länge geschätzt, die plausibel klingt und nicht deine ist. Und der ungefragte Lobsatz am Anfang ist kein Schönheitsfehler: er ist der Grund für Regel 1: erst herausziehen, dann bewerten.

Die Regeln, die dabei herausfallen

- Erst Bestandteile herausziehen, dann bewerten.
- Fehlende Pflichtangaben führen zu einer Rückfrage.
- Plausible Angaben werden nicht ergänzt.
- Jeder Befund bekommt einen Textbeleg.
- Nur beanstandete Stellen werden überarbeitet.
- Über die Endfassung entscheidet die Lehrkraft.

Die Grenze

Der Check prüft die Form des Auftrags, nicht seine fachliche Richtigkeit und nicht seine Passung zu deiner Lerngruppe. Ob die Aufgabe für deine Klasse zu schwer ist, weiß nur, wer die Klasse kennt. Übersehen wird deshalb alles, was außerhalb des Textes liegt: dass die mitgegebene Quelle die verlangte Angabe gar nicht enthält, dass der Operator in deinem Bundesland eine andere Leistung verlangt, dass die Aufgabe eindeutig ist und trotzdem am Lernziel der Reihe vorbeigeht. Die Prüftabelle bleibt dabei vollständig ausgefüllt. Sie meldet keine Lücke, für die sie kein Feld hat.

Regel 3 ist die, die den Check von Lob unterscheidet.

Die Stolperstelle. Unter Zeitdruck wackelt zuerst Regel 2, die Rückfrage bei fehlenden Angaben. Wer das Blatt in der nächsten Pause braucht, will kein Gespräch und ergänzt die fehlende Angabe schnell selbst. Das ist zulässig, solange du es tust und nicht der Skill: Aufweichen heißt hier, die Angabe nachzureichen, nicht die Regel zu streichen. Verschwindet sie aus dem Skill, verschwindet mit ihr die Stelle, an der auffällt, dass etwas fehlte.

15

Der fertige Skill-Entwurf

Name**Arbeitsauftrag-Check****Beschreibung**

Prüft Arbeitsaufträge vor dem Unterricht auf Lernzielbezug, Eindeutigkeit, Vollständigkeit, Zeitbedarf und mögliche Fehldeutungen.

Eingaben

Lernziel · Jahrgangsstufe · Arbeitsauftrag · erwartetes Produkt · Zeitrahmen · verfügbares Material

Ablauf

- Zieh Operator, erwartetes Produkt, Materialbasis und Zeit heraus.
- Melde fehlende Pflichtangaben und warte auf Antwort.
- Prüf, ob Lernziel, verlangte Tätigkeit und Produkt zusammenpassen.
- Prüf, ob der Auftrag ohne mündliche Zusatzklärung ausführbar ist.
- Nenne höchstens drei wahrscheinliche Fehldeutungen.
- Schätz den Zeitbedarf und sag, worauf die Schätzung beruht.
- Überarbeite nur beanstandete Stellen.
- Gib eine Prüftabelle und eine als Vorschlag markierte Fassung aus.

Ausgabeformat

Prüffeld	Befund mit Textbeleg	Priorität	Änderungsvorschlag
Lernzielbezug			
Eindeutigkeit			

Prüffeld	Befund mit Textbeleg	Priorität	Änderungsvorschlag
Vollständigkeit			
Zeit			
Fehldeutungen			

Stoppregeln

- Lernziel oder Arbeitsauftrag fehlt.
- Der Auftrag setzt Material voraus, das nicht mitgegeben wurde.
- Personenbezogene oder vertrauliche Daten sind enthalten.
- Fachliche Richtigkeit ist ohne Quelle nicht beurteilbar.

Die drei Testfälle

A Normalfall: vollständiger Auftrag mit einer kleinen sprachlichen Unklarheit.

B Fehlende Information: kein erwartetes Produkt angegeben.

C Grenzfall: Der Auftrag verweist auf eine nicht mitgelieferte Schulbuchseite und enthält einen echten Namen.

Menschliche Endprüfung: Fachlichkeit, Passung zur Lerngruppe und die endgültige Formulierung bleiben bei dir.

Die Grenze

Der Entwurf ist ein Ausgangspunkt, keine Vorlage zum Übernehmen. Die Prüffelder passen zu einem schriftlichen Einzelauftrag. Für einen Stationenlauf, für Gruppenarbeit mit Rollen oder für eine Aufgabe im Fachraum brauchst du andere, und die findest du nur, indem du Kapitel 07 selbst durchgehst.

Die Stolperstelle. Sobald Geräte, Stoffe oder ein Fachraum ins Spiel kommen, fehlt den fünf Feldern eines für die Voraussetzungen. In der Chemie entscheidet „Was muss bereitstehen, was ist verboten?“ mehr über die Stunde als die Eindeutigkeit der Formulierung. Umgekehrt streicht man ein Feld nicht, weil es unwichtig wäre, sondern weil es bei dieser Aufgabenform nie einen Befund erzeugt.

Schritt 2 wartet. Das ist kein Höflichkeitsschritt.

16

Erst stabil, dann mehr

Ein zweiter Skill, ein Agent oder eine Automation lohnt sich erst, wenn der erste Ablauf verlässlich läuft.

Die Reihenfolge

- eine klar begrenzte Aufgabe auswählen
- einen Durchlauf erarbeiten, der klappt
- daraus einen Skill machen
- drei Tests bestehen
- mehrere echte, datensichere Einsätze beobachten
- Fehler gezielt einarbeiten
- erst dann den zweiten Workflow

Bereit für den nächsten Schritt?

- ☐ Der erste Skill löst genau eine Aufgabe.
- ☐ Er fragt bei fehlenden Angaben nach.
- ☐ Er stoppt in sensiblen oder unklaren Fällen.
- ☐ Seine Ausgabe lässt sich schnell prüfen.
- ☐ Änderungen werden festgehalten und neu getestet.
- ☐ Der neue Workflow hat einen deutlich anderen Zweck.

Wenn mehrere Skills irgendwann nur gemeinsam ein Ergebnis erzeugen, kann Koordination sinnvoll werden. Ein Ziel für sich ist sie nicht. Und sobald etwas ohne deine Zwischenfreigabe läuft, gilt Kapitel 00, Abschnitt 6.

Ein Plan für zwei Wochen

Tag 1–2: Workflow auswählen, Erfolgskriterien aufschreiben. **Tag 3–4:** begleiteten Durchlauf machen. **Tag 5:** Erfolgsprotokoll verdichten. **Tag 6:** ersten Skill-Entwurf erstellen. **Tag 7:** die drei Tests. **Tag 8–12:** an zwei weiteren Themen einsetzen, ohne Personenbezug. **Tag 13:** Fehler und überflüssige Regeln auswerten. **Tag 14:** freigeben, überarbeiten oder bewusst verwerfen.

Die Grenze

Verwerfen ist ein Ergebnis, kein Scheitern.

Die vierzehn Tage sind eine Reihenfolge, keine gemessene Dauer. In einer Woche mit Elternsprechtage dauert Tag 3 eben drei Tage. Verbindlich ist die Reihenfolge, nicht das Datum.

Die Stolperstelle. Die Dauer hängt nicht am Arbeitsaufwand. Sie hängt an Schritt 5. Die Tage 1 bis 7 sind Arbeit, die man am Stück erledigen kann; die weiteren Einsätze aus Tag 8–12 kommen dagegen erst, wenn die Aufgabe wiederkommt. Bei einem Arbeitsauftrag, den du montags schreibst, dauert der Weg zur freigegebenen Fassung deshalb mindestens so viele Wochen, wie du Einsätze sehen willst. Unabhängig davon, wie schnell du schreibst.

17

Wie es weitergeht

Dieser Band endet bei einem Skill, der funktioniert. Damit ist ein Ablauf gesichert. Der Kontext aber noch nicht: Vor jedem Einsatz erklärst du weiterhin, welche Klasse, welche Reihe, welches Buch und wie lang die Stunde ist.

Band 2 — „Unterrichtsmaterial selbst bauen“ setzt genau dort an. Er gibt dem Kontext einen Ort auf deinem Rechner: einen Ordner mit einer Ankerdatei, Reihenplänen, gespeicherten Abläufen und einem Freigabeordner, aus dem nichts von selbst herauskommt. Er baut auf den sechs Begriffen, dem Kontextfilter und der Datenampel dieses Bandes auf und wiederholt sie nicht.

18

Quellen, Grenzen, Aktualität

Grundlagen

- Anthropic Help Center: *What are skills?* — Funktionsweise und progressive Offenlegung von Skill-Inhalten (2026).
<https://support.claude.com/en/articles/12512176-what-are-skills>
- Anthropic Help Center: *How to create custom skills* — Aufbau, Tests und Sicherheitshinweise (2026).
<https://support.claude.com/en/articles/12512198-how-to-create-custom-skills>
- Agent Skills: offene Spezifikation für portable Skill-Pakete.
<https://agentskills.io/>
- Kultusministerkonferenz: *Handlungsempfehlung für die Bildungsverwaltung zum Umgang mit Künstlicher Intelligenz in schulischen Bildungsprozessen* (2024).
https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2024/2024_10_10-Handlungsempfehlung-KI.pdf
- Datenschutzkonferenz: *Künstliche Intelligenz und Datenschutz*, Version 1.0 vom 6. Mai 2024.
https://www.datenschutzkonferenz-online.de/media/oh/20240506_DSK_Orientierungshilfe_KI_und_Datenschutz.pdf

Woher die Methode kommt

Der Dreischritt aus begleitetem Durchlauf, Verdichtung und schrittweiser Verbesserung stammt nicht aus der Fachdidaktik, sondern aus der eigenen Arbeit mit KI-Werkzeugen für die Unterrichtsvorbereitung. Er ist ein begründbares Verfahren, kein empirisch untersuchtes. Wenn du hier eine Zahl suchst, die belegt, dass es funktioniert, wirst du sie nicht finden. Das ist Absicht.

Was sich ändern wird

Skills, gespeicherte Anweisungen und Agentenfunktionen unterscheiden sich je nach Produkt. Verfügbarkeit, Dateiformat, Sicherheitsfunktionen und Oberflächen ändern sich. Deshalb ist die Methode toolunabhängig beschrieben, und deshalb steht in Kapitel 13 die Stufe „Überarbeiten“ ausdrücklich für den Fall einer Produktänderung.

Dieses Handbuch ist keine Rechtsberatung und ersetzt keine schulische oder behördliche Freigabe. Entscheidungen über

Leistung und Förderung bleiben bei der Lehrkraft.

Die Idee in einem Satz

Führ einen Workflow zuerst erfolgreich vor, verdichte ihn zu einem klaren Skill und verbessere ihn nur anhand von Fehlern, die du wirklich verstanden hast.

Band 1 · Ausgabe 1 · Redaktionsstand: 10. August 2026

Erst vormachen. Dann verdichten. Dann verbessern.