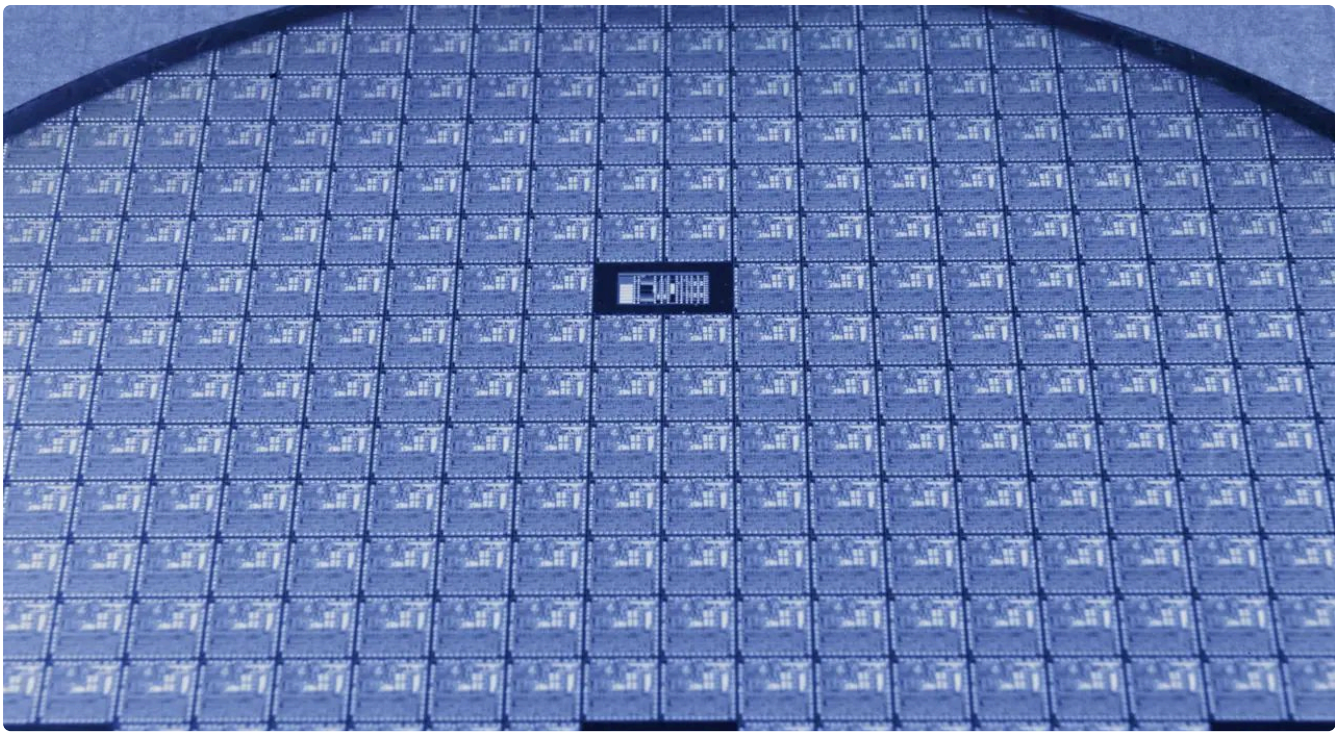


PERSPEKTIVE · NOVEMBER 2026 · 6 MIN. LESEZEIT

Prototypen, die Entscheidungen erzwingen

Vier Regeln für Experimente, die mit einer klaren Entscheidung enden



Philipp Pöschke
Founder, Paice

AUF EINEN BLICK

- Ein Prototyp hat seinen Zweck erfüllt, wenn er zu einem vorher vereinbarten Termin eine Entscheidung möglich macht: weiterbauen, anpassen oder beenden.
- Erfolgskriterien, Messmethode und Entscheidungstermin gehören vor den ersten Arbeitstag, weil sich jedes Ergebnis nachträglich als interessant deuten lässt.
- Aussagekräftig wird ein Prototyp erst unter echten Bedingungen, und über das Ergebnis entscheiden Messungen, weil der Eindruck der Beteiligten nachweislich täuschen kann.

Prototypen sind in vielen Unternehmen Routine geworden. Sie zeigen, dass etwas technisch geht. Ob es sich lohnt und was als Nächstes passiert, bleibt danach oft offen, obwohl genau diese Entscheidung ihr Zweck ist.

Das Analystenhaus Gartner hatte im Juli 2024 vorhergesagt, dass bis Ende 2025 mindestens 30 Prozent der Vorhaben mit generativer AI nach dem Proof of Concept aufgegeben werden. Als Gründe nannte es schlechte Datenqualität, unzureichende Risikokontrollen, steigende Kosten und unklaren Geschäftsnutzen. Dass Vorhaben enden, halten wir für kein Problem. Teuer wird es, wenn die Entscheidung spät fällt und niemand vorher festgelegt hat, woran sie sich bemisst.

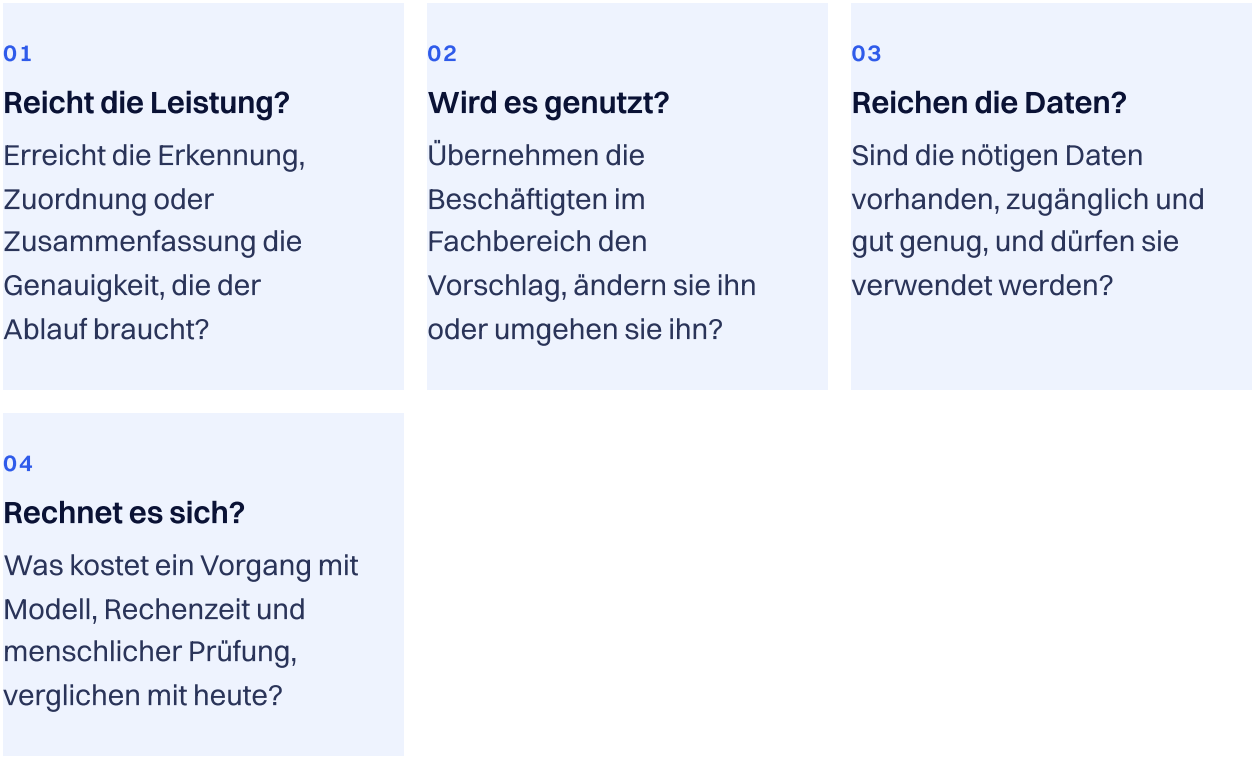
Im Modell des AI-Deltas gehört ein Prototyp zum Erkunden: Er klärt eine einzelne Frage und bereitet die nächste Entscheidung vor. Damit das gelingt, helfen vier Regeln und eine Vereinbarung, die auf eine Seite passt.

Regel 1: Mit einer Frage beginnen

Ein guter Prototyp beginnt mit einer Frage, die man ohne ihn nicht beantworten kann. Die Frage bestimmt, was gebaut wird, und meist ist das weniger als gedacht. Vier Arten von Fragen lassen sich unterscheiden.

ABBILDUNG 1

Vier Arten von Fragen, die ein Prototyp beantworten kann



Darstellung: Paice

Steht die Frage fest, lässt sich der Prototyp auf das beschränken, was sie beantwortet. Ein vereinfachtes Beispiel zeigt den Unterschied. Lautet die Frage, ob ein Modell die Positionen aus den Rechnungen der häufigsten Lieferanten zuverlässig genug ausliest, braucht der Prototyp weder eine Oberfläche noch eine Anbindung an die Buchhaltung. Er braucht einen Stapel echter Rechnungen, die richtigen Werte dazu und eine Auswertung. Erst wenn die Antwort positiv ausfällt, folgt die nächste Frage, etwa ob die Buchhaltung die Vorschläge im Alltag übernimmt.

Regel 2: Kriterien vor dem Bau festlegen

Vor dem ersten Arbeitstag steht fest, wann der Prototyp als erfolgreich gilt, wie das gemessen wird, wer entscheidet und an welchem Tag. Ohne diese Vereinbarung wird aus fast jedem Ergebnis ein „interessant, machen wir weiter“. Mit ihr gibt es drei klare Ausgänge: weiterbauen, anpassen oder beenden.

Die Kriterien sollten aus dem Geschäft kommen. Eine Genauigkeit über einem bestimmten Wert ist nur dann ein sinnvolles Ziel, wenn jemand weiß, ab welchem Wert sich der Ablauf tatsächlich ändern lässt. Deshalb legt der Fachbereich die Schwellen fest, und das Entwicklungsteam sagt, wie sie gemessen werden.

Auch die Dauer gehört in die Vereinbarung. Sie folgt aus der Frage und wird vorher begrenzt. Ist der Termin erreicht, wird entschieden, auch wenn das Ergebnis unvollständig ist. Dann lautet die Entscheidung

womöglich „anpassen“, mit einer neuen, engeren Frage und einem neuen Termin. Ein Prototyp ohne Enddatum wächst dagegen still zu einem Projekt, das niemand beschlossen hat.

Bei generativer AI kommt hinzu, dass dieselbe Eingabe unterschiedliche Ergebnisse liefern kann. Ein Prototyp braucht deshalb einen festen Satz von Testfällen mit erwarteten Ergebnissen, an dem jede Version gemessen wird. Die Auswertung erfasst auch die Art der Fehler, denn eine erfundene Angabe wiegt in einem Vertrag schwerer als eine umständliche Formulierung.

Regel 3: Unter echten Bedingungen testen

Prototypen auf Beispieldaten überzeugen leicht. Aussagekräftig werden sie erst mit echten Daten, echten Nutzern und echten Ausnahmen. Zwei Studien zeigen, warum.

Die Forschung zur „gezackten Grenze“ von AI zeigt, wie trügerisch Beispiele sind. In einem Experiment mit 758 Beschäftigten einer Unternehmensberatung arbeiteten Teilnehmende mit GPT-4 bei passenden Aufgaben deutlich schneller und besser. Bei einer Aufgabe, die ähnlich aussah, aber außerhalb der Stärken des Modells lag, lieferten sie 19 Prozentpunkte seltener eine richtige Lösung als die Gruppe ohne AI. Ein Prototyp, der nur an den leichten Fällen getestet wird, sagt über die schwierigen nichts.

Die zweite Studie zeigt, warum gemessen werden muss. Erfahrene Open-Source-Entwickler brauchten mit AI-Werkzeugen 19 Prozent länger für ihre Aufgaben. Hinterher schätzten sie, sie seien 20 Prozent schneller gewesen. Der Eindruck der Beteiligten ist deshalb kein Ergebnis. Gemessen werden Zeit, Fehler und Kosten, am besten im Vergleich mit dem heutigen Ablauf.

Regel 4: Beenden ist ein Ergebnis

Ein Prototyp, der zeigt, dass sich etwas nicht lohnt, hat Geld gespart. Damit das so gesehen wird, braucht das Beenden eine Form. Eine kurze Dokumentation hält fest, was getestet wurde, was herauskam und unter welchen Bedingungen sich ein neuer Versuch lohnen würde, etwa mit einem besseren Modell oder neuen Daten. So bleibt das Gelernte erhalten, und das nächste Team fängt nicht bei null an.

Entscheiden sollte die Person, die auch das Budget für den nächsten Schritt verantwortet. Liegt die Entscheidung beim Team, das den Prototyp gebaut hat, entscheidet es über die eigene Arbeit, und das macht das Beenden unnötig schwer.

Unternehmen, die das Beenden anerkennen, trauen sich mehr Experimente. Viele kleine Entscheidungen auf Grundlage von Messungen sind robuster als wenige große Wetten auf Grundlage von Vorführungen.

Gängige Sicht, unsere Sicht

Die vier Regeln verschieben, woran ein Prototyp gemessen wird. Die Übersicht stellt die gängige Sicht unserer gegenüber.

ABBILDUNG 2

Prototypen: gängige Sicht, unsere Sicht

GÄNGIGE SICHT	UNSERE SICHT
Ein Prototyp soll zeigen, dass etwas technisch geht.	**Ein Prototyp soll eine Entscheidung möglich machen.**
Erfolg heißt, der Prototyp wird weitergebaut.	**Erfolg heißt, zum vereinbarten Termin wird entschieden, auch wenn die Entscheidung Beenden lautet.**
Getestet wird mit Beispieldaten, bewertet nach Eindruck.	**Getestet wird mit echten Daten und Ausnahmen, bewertet nach vorher festgelegten Messungen.**
Ein abgebrochener Prototyp ist ein Misserfolg.	**Ein abgebrochener Prototyp hat eine Investition verhindert, die sich nicht gelohnt hätte.**

Darstellung: Paice

Am deutlichsten zeigt sich der Unterschied bei der Frage, was als Erfolg gilt. Zählt nur der weitergebaute Prototyp, hat jedes Team einen Grund, Ergebnisse freundlich zu lesen und Termine zu verschieben. Zählt die rechtzeitige Entscheidung, lohnt es sich, früh ehrlich zu sein. Für ein Portfolio von Experimenten ist das der bessere Maßstab, weil er Budget dorthin lenkt, wo die Messungen tragen.

Die Prototyp-Vereinbarung

Die vier Regeln lassen sich in einer Vereinbarung festhalten, die vor dem Start von der Person unterschrieben wird, die entscheidet, und vom Team, das baut. Sie passt auf eine Seite und enthält sechs Punkte:

01 **Frage:** Welche Frage beantwortet der Prototyp, und welche Entscheidung hängt daran?

02 **Annahme:** Was müsste stimmen, damit sich das Vorhaben lohnt?

03 **Kriterien:** Ab welchen Werten gilt die Frage als beantwortet, und wie wird gemessen?

04 **Bedingungen:** Mit welchen Daten, Nutzern und Fällen wird getestet, und wie lange?

05 **Entscheidung:** Wer entscheidet, und an welchem Tag?

06 **Ausgänge:** Was passiert beim Weiterbauen, beim Anpassen und beim Beenden, einschließlich Budget und Zuständigkeit?

> Ein Prototyp ist fertig, wenn entschieden ist, was mit ihm passiert.

Quellen

- Gartner Predicts 30% of Generative AI Projects Will Be Abandoned After Proof of Concept By End of 2025, Gartner, 29.07.2024
- Navigating the Jagged Technological Frontier, Dell'Acqua u. a., Harvard Business School Working Paper 24-013, 22.09.2023
- Measuring the Impact of Early-2025 AI on Experienced Open-Source Developer Productivity, Becker u. a., METR, 12.07.2025



AUTOR

Philipp Pöschke

Founder, Paice

Philipp Pöschke hat Paice 2026 in Bremen gegründet. Er verantwortet jedes Vorhaben persönlich und arbeitet selbst an Analysen, Prototypen und Lösungen mit. Die Beiträge unter Insights verantwortet er als Autor.