

A/B Tests - Testimonials Identifikation

warum braucht man noch ab tests, wenn wir ki haben, und erfahrungswerte aus mehr als 20 jahren?

ChatGPT said:

Sehr gute Frage – und absolut berechtigt.

Warum braucht man A/B-Tests, wenn wir bereits eine KI mit Erfahrungswerten und fundierten Methoden einsetzen?

Hier ist eine präzise Antwort aus Sicht der **Product Ownership, KI-Orchestrierung und Enterprise-Skalierung**:

Warum A/B-Tests trotz KI notwendig bleiben

1. KI ≠ Allwissenheit – sondern probabilistische Systeme



Selbst modernste KI-Modelle wie Claude 3.5 oder Gemini Ultra basieren auf Wahrscheinlichkeiten, Trainingsdaten und statischen Snapshots. Sie **kennen keine echte Echtzeit-Wirkung** in deinem konkreten Anwendungskontext – z. B.:

- Wie deine Zielgruppe **reagiert** (Klicks, Engagement, Absprünge)
- Welche Prompt-Variante **messbar mehr Conversions** erzeugt
- Ob eine Claude-Version „logisch klingt“, aber **nicht zur Marke passt**

Nur A/B-Tests können das im **echten Live-Kontext** empirisch validieren.

2. Feintuning & Performance-Maximierung

Auch wenn ein Prompt grundsätzlich funktioniert, kann A/B-Testing zeigen:

Variante A	Variante B
 3.8s Response, 65% Click Rate	 2.5s Response, 78% Click Rate

 Das macht im großen Maßstab **Tausende Euro Differenz pro Monat** aus – insbesondere bei AI-basierten Landingpages, Calls-to-Action oder DSGVO-Flows.

3. Business- & Persona-Adaption

Du willst nicht nur wissen, ob ein Prompt "funktioniert", sondern **für wen**:

- Persona A (Agenturinhaber) bevorzugt sachliche Analyse
- Persona B (Kochinfluencerin) reagiert besser auf emotionale Sprache

→ A/B-Tests liefern Daten für **persona-spezifische Optimierung** – auch in Kombination mit KI-basiertem Targeting.

4. Qualitätssicherung & Regressionsvermeidung

Bei KI-Updates, z. B. Claude Sonnet 3.5 → Claude Haiku 4.0, kann es **zu Leistungsver schlechterungen kommen**, die nicht sofort auffallen.

A/B-Tests fungieren als:

- **Frühwarnsystem**
- **Regression-Detector**
- **Qualitätsbeweis** gegenüber Stakeholdern & Partnern

5. Kontrollierte Innovation & Risikominimierung

Du kannst neue Prompt-Ideen oder Claude-Strategien **kontrolliert ausrollen**, z. B.:

- 10 % der Nutzergruppe erhalten Version B
- Nur wenn Conversion ↑ → Ausweitung auf 100 %

👉 Das schützt dein System vor Reputationsverlust oder Performance-Drops durch zu aggressive KI-Änderungen.

Fazit

| Künstliche Intelligenz erzeugt Vorschläge – aber A/B-Tests messen Wirkung.

Sie sind wie ein GPS für dein Prompt-System:

- **Klarheit statt Bauchgefühl**
- **Empirie statt Halluzination**
- **Wachstum statt Rätselraten**

Wenn du möchtest, kann ich ein kurzes **Visual für cChat oder Stakeholder** dazu erstellen („Warum A/B-Tests trotz KI?“) – z. B. als Markdown-Tabelle oder Chart. Gib mir einfach das Format.

You said:

und wer erstellt und setzt die tests in matbakh.app ein?

ChatGPT said:

Sehr gute Anschlussfrage. Die Verantwortlichkeit für A/B-Tests in **Matbakh.app** ist **klar verteilt**, aber erfordert Zusammenarbeit zwischen mehreren Komponenten und Rollen:

Wer erstellt und setzt A/B-Tests in Matbakh.app ein?

1. Prompt Engineers / Template Owner (Kiro)

- **Erstellen neue Prompt-Varianten** mit unterschiedlichen Zielsetzungen (z. B. emotional vs. sachlich)
- **Versionieren** die Prompts in `PromptTemplateManager`
- **Markieren** eine Variante als „A“, eine andere als „B“
- **Pflegen Metadaten:** Zielpersona, Version, Zweck, Beschreibung

📁 Ort:

```
src/ai/templates/
```

2. VC Generator / Claude Lambda

- **Wählt Prompt-Variante** basierend auf dem Rollout-Mechanismus (z. B. 50/50 oder persona-weighted)
- **Sendet Prompt & Version ID** mit zur API-Response
- **Loggt** Prompt ID + Output + Userpersona + Score + Feedback in `prompt_logs`

 Ort:

`infra/lambdaas/vc-generator/src/`

3. Analytics & QA Layer

- Nutzt `prompt_logs` + `vc_results` zur **Auswertung von Wirkung & Qualität**
- Metriken z. B.:
 - 🕒 Response-Zeit
 - ✅ Response-Güte (Scoring)
 - 🎯 Recommendation-Clickrate
 - 🧑 Persona-Match-Score
- Führt automatisierte Reports durch, z. B. „Prompt A performt 17 % besser als B bei IG-Influencerinnen in München“

 Ort:

`src/analytics/prompt-testing/
docs/reports/prompt-performance-2025-*.md`

4. Admin / Product Owner (cChat)

- Entscheidet auf Basis der Reports:
 - ⚖️ Variante B übernehmen?
 - 🗑️ Variante verwerfen?
 - ↺️ Neuen A/B-Zyklus starten?
- Kann über UI oder Flags **Rollout-Ratio ändern**

Zusammenspiel der Komponenten (End-to-End)

```
graph TD
  A[Prompt Engineer (Kiro)] --> B[PromptTemplateManager]
  B --> C[VC Generator Lambda]
  C --> D[Claude Output + Prompt Metadata]
  D --> E[Prompt Logs & VC Results]
  E --> F[Analytics Engine]
  F --> G[Performance Reports]
  G --> H[Product Owner (cChat) Entscheidung]
```

H → B

✓ Wer setzt die Tests konkret ein?

Komponente	Verantwortlich für	Realisiert durch
PromptTemplateManager	Erstellung + Versionierung	Kiro
vc-generator Lambda	Rollout + Logging	AI Orchestrator (Kiro)
prompt_logs + vc_results	Speicherung der Ergebnisse	Backend + DB
analytics/prompt-testing/	Auswertung + Reports	QA + Kiro
Admin UI / Feature Flags	Steuerung	cChat (PO)

Wenn du willst, kann ich ein Starter-Template für die `usePromptTesting.ts` Hook + `prompt_logs` Schema mitliefern. Damit kann cChat direkt erste Tests starten oder bewerten. Bereit?

You said:

da diese auf response basieren, ist eine klare definition der empfänger/testpersonen erforderlich. wir wissen was wir senden wollen, jetzt müssen wir herausfinden wer die empfänger sind, um dann zu ermitteln was bei denen besser ankommt. wie können wir die einzelnen personas runterbrechen auf die gruppe early adapters?, vllleicht vorher noch testimonials. wie werden sie ermittelt? ihr wert ist sicherlich materiell sehr positiv, aber immateriell noch wertvoller da sie als leitwölfe in der industrie oder als vorreiter oder als jemand den mann nachmachen will. wie identifiziere ich diese gruppe für matbakh.app

Exzellenter Punkt – du bringst das Thema auf die **strategisch entscheidende Ebene**:

Nicht *was* getestet wird, sondern *für wen* – und *warum gerade sie* das Spiel entscheiden.

Ziel: Identifikation der Early Adopters & Leitwölfe (Personas mit Signalwirkung)

Diese Gruppe ist **nicht nur Testgruppe**, sondern:

- Meinungsführer in ihrer Nische
- Multiplikatoren für deine Plattform
- Inspirationsquelle für andere Gastronomen („Wenn Sapralott es nutzt, probiere ich es auch“)
- **Proof of Value** für Investoren („Hier funktioniert es bereits“)

1. Zieldefinition: Wer sind *Early Adopters* in deinem Kontext?

Kriterium	Typische Merkmale
Offen für Neues	Probieren gerne Tools & Plattformen aus
Sichtbar & vernetzt	Große IG-Reichweite, Events, Interviews, Medienpräsenz
Experimentierfreudig	Machen „Pop-ups“, „Food Festivals“, spezielle Menüs
Lokal oder branchenspezifisch relevant	Werden oft nach Empfehlungen gefragt
Haben bereits <i>Pain Points</i> erkannt	Wissen, dass Online-Sichtbarkeit entscheidend ist

Kriterium	Typische Merkmale
<i>Influence by being first</i>	Andere machen es ihnen nach („Trendsetter“)

2. Verbindung zu bestehenden Personas (Matbakh MVP)

Du hast bereits 4 Personas definiert:

Persona	Frühadopter-Potenzial?	Warum / Warum nicht
Anna (gesundheitsbewusste Genießerin)	✓ Hoch	Achtet auf Nachhaltigkeit, folgt „clean food“-Trends, stark visuell
Sophie (trendbewusste Studentin)	✓ Sehr hoch	TikTok-affin, entdeckt Neues sofort, teilt aktiv
Jürgen (geselliger Best Ager)	● Mittel	Eher Nachzügler, aber empfiehlt gern weiter
Markus (erlebnisorientierter Familienvater)	● Gering	Nutzt funktionale Tools, aber testet ungern Neues

➡ Deine besten **Early-Adopter-Testgruppen** sind:

1. Restaurants, die auf **Anna & Sophie** abzielen
2. Orte, die **bereits von diesen Personas empfohlen werden**

3. Identifikationsmethoden für Leitwölfe & Early Adopters

A. Qualitatives Screening deiner bisherigen Kontakte

Nutze deine bestehenden Netzwerke in München (z. B. Giesinger Garten, Sapralott, Lost Weekend etc.):

Frage	Erkenntnis
Wer hat regelmäßig neue Ideen oder Formate?	→ Innovator
Wer hat große Sichtbarkeit auf IG/Google?	→ Meinungsführer
Wer bekommt viele Empfehlungen oder Bookings über digitale Kanäle?	→ Erfolgsbeweis
Wer hat Lust auf Tools / spricht dich auf neue Sachen an?	→ Testbereit

➡ Ergebnis: Erste Liste von 10–20 potenziellen Pilotkunden/Testpersonen

B. Matbakh-spezifische Indikatoren (Datenbasiert)

Du kannst folgende Kriterien aus der `visibility_check_leads` Tabelle verwenden:

Datenpunkt	Signal
Google Reviews > 50	Hohe öffentliche Sichtbarkeit
Instagram mit > 1k Follower	Potenzieller Multiplikator
Sehr hohe oder sehr schlechte Sichtbarkeit	Entweder „Vorreiter“ oder „will besser werden“
Hohe Engagementrate auf Posts	Community vorhanden
Events / Kooperationen / Food Pop-ups	Innovationstreiber

C. Testimonials gezielt gewinnen

Bevor du testest, könntest du:

1. 3–5 Leitwölfe persönlich ansprechen („Wir wollen dich als **erste*r Pilotkunde** gewinnen...“)
2. Ihnen kostenlos den **Visibility Check + individuelle Analyse** anbieten
3. Im Gegenzug ein Statement oder kurze Bewertung anfragen

🗨️ Z. B.:

„Seit ich mit Matbakh arbeite, sehen doppelt so viele Leute mein Lunch-Menü. Und ich muss mich um nichts kümmern.“

→ Du brauchst nur **2–3 echte Stimmen mit Wiedererkennungswert** (Name, Gesicht, Marke).

📌 Fazit & nächste Schritte

🎯 Zielgruppe für A/B-Tests:

- Restaurants mit folgenden Merkmalen:
 - Innovativ
 - Sichtbar
 - Bereits digital aktiv
 - Influencer-Wirkung in der Zielgruppe Anna/Sophie

🔪 Methoden zur Identifikation:

- Analyse bestehender Leads
- Manuelles Screening deiner Kontakte
- Sichtbarkeitsdaten aus Google/Instagram
- Direct Outreach & Angebot von kostenlosen VC + personalisierter Analyse

„ich habe im b2b bereich andere personas identifiziert (siehe dokument), aber ausgangspunkt waren sie unternehmertypen. jetzt geht es um was bewegt diese unternehmertypen die in dieser industrie aktiv sind, wie bekommt man ihre aufmerksamkeit? oder was brauchen sie? welche strategische vorgehensweise ist korrekt oder ist für matbakh.app die richtige? denn wenn sie was brauchen, habe ich die lösung für ihr problem in meiner kommunikation. wenn sie aber auf dem ohr neugier und interesse, nicht auf dem ohr habgier, angesprochen werden sollen, muss eine andere sprache her.“

Danke, sehr wertvoller Input – und dein Personas-Dokument zeigt klar, dass du im **B2B-Bereich** mit sehr unterschiedlichen Unternehmertypen arbeitest.

Wir gehen jetzt *strategisch systematisch* vor:

🎯 Ziel: Die richtige strategische Kommunikationslogik für matbakh.app B2B-Early-Adopters finden

📍 1. Ausgangslage: Deine identifizierten Unternehmertypen






Basierend auf deinem Personas-Dokument sind zentrale Typen:

Unternehmertyp	Merkmale	Beispiele
🧠 Visionär	Denkt langfristig, sucht Differenzierung	Franchisegeber, Erlebnisgastronom

Unternehmertyp	Merkmale	Beispiele
 Macher	Operativ stark, digital überfordert	Inhaber kleiner Restaurants
 Markenbauer	Design- und imagefokussiert	Hippe Cafés, Szene-Gastronomie
 Kommunikator	Stark in Community & Story	Veranstalter, Streetfood
 Routineunternehmer	Bestandskunden-orientiert, scheut Neues	Familienbetrieb, Wirtshaus

2. Was brauchen sie wirklich – und wie hören sie zu?

 Unterschiedliche „psychologische Kanäle“ (basierend auf Schulz von Thun):


Ohr-Typ	Was sie hören wollen	Tonalität
 Neugier / Innovations-Ohr	„Was ist neu? Wie funktioniert das?“	Explorativ, inspirierend
 Nutzen- / Effizienz-Ohr	„Wie spare ich Zeit, Geld, Aufwand?“	Direkt, pragmatisch
 Kompetenz-Ohr	„Was macht mich besser als andere?“	Strategisch, analytisch
 Werte-/Image-Ohr	„Passt das zu meinem Stil & Anspruch?“	Ästhetisch, wertebasiert
 Social Proof-Ohr	„Wer hat das schon? Was sagen andere?“	Testimonials, Storytelling

3. Strategischer Schlüssel für matbakh.app: „Jobs-to-be-Done + Matching Ohr“



Wir verbinden:

- Was sie wirklich wollen (Job-to-be-Done)
- mit dem richtigen Ohr, das sie aktiviert


 Beispiel: Der „Macher“ (Inhaber, operativ stark)

Job	Ohr	Ansprache
„Ich will mein Lokal voll bekommen, ohne mich selbst darum zu kümmern.“	 Effizienz-Ohr	„Matbakh bringt dir neue Gäste – ohne dass du dich um irgendwas kümmern musst.“

 Beispiel: Der „Kommunikator“ (Storyteller, Veranstalter)

Job	Ohr	Ansprache
„Ich will, dass meine Story gesehen wird.“	 Neugier +  Werte	„Matbakh macht deine Vision sichtbar – auf Google, Instagram & überall da, wo deine Community ist.“

 Beispiel: Der „Visionär“ (Franchise, Mehrstandort-Betreiber)

Job	Ohr	Ansprache
„Ich will datenbasiert wachsen und strategisch die besten Standorte skalieren.“	 Kompetenz-Ohr	„Matbakh zeigt dir, wo dein nächster Top-Standort ist – auf Basis echter Sichtbarkeitsdaten.“

4. Strategische Kommunikationsmatrix (Empfehlung für matbakh.app)

Zielgruppe	Haupt-Ohr	Sprache	Kanal
Kleine Gastronomen (1-Standort)	💰 Nutzen	Einfach, konkret, zeitersparend	WhatsApp, Direktmail, Beratung
Markenbewusste Konzepte	❤️ Werte + 🧠 Kompetenz	Ästhetisch, strategisch, visuell	IG DM, Landingpage, Testimonial
Franchises / Ketten	🧠 Kompetenz + 🎧 Neugier	Skalierbarkeit, Insights, Charts	LinkedIn, Slides, Case Study
Veranstalter / Community-Orte	🎧 Neugier + 💬 Social Proof	Sichtbarkeit, Eventwirkung, echte Stimmen	Video, Shownotes, Co-Kampagnen

✅ Handlungsempfehlung (Next Steps für dich)

1. 🎯 Persona-Ohr-Zuordnung vollziehen

→ Für jede B2B-Persona: Haupt-Ohr + Job-to-be-Done definieren

(Wenn du willst, helfe ich dir beim Ausfüllen einer Matrix.)

2. 🗣️ Messaging-Framework pro Persona erstellen

→ 3 Sätze:

- Was ist ihr Schmerz?
- Was ist ihre Hoffnung?
- Was löst Matbakh genau?

3. 📢 Landingpages & Ads daran anpassen

→ Je nach Ohr + Trigger kannst du sogar mit der gleichen Leistung

– einmal als „dein Laden füllt sich“,

– einmal als „dein Google-Ranking steigt“ kommunizieren.

Bonus: Wie du Testimonials gezielt nutzt

- **Nicht überreden, sondern Positionieren:**

„Wir suchen Vordenker aus der Branche – du bekommst alle Tools kostenlos + eine personalisierte VC-Analyse.“

- **Story, nicht Technik:**

Lass sie erzählen, was sich durch Matbakh geändert hat (mehr Buchungen, mehr Besucher, weniger Stress)

- **Formate:**

- 1-Min Reel
- Quote auf der Landingpage
- Interview in der App
- Fallstudie als LinkedIn-Post (von dir)

ich möchte gerne noch etwas hinzufügen: ich will ergebnisse, bzw. the done job zeigen, nicht die technik, oder was sie alles kann. was ist mein erfolg? was bringt matbakh.app in mein leben? und ja, bitte erstelle mir die matrix/tabelle mit bezug auf die von mir identifizierten personas. wir fangen zuerst mit dem typ der persona, und was bringt matbakh.app für ihn/sie in ihr leben, welcher job wird erledigt, was kommt dabei raus? denn wir wissen, nur gastronomie kommen infrage (und ihre mitarbeiter) es sei denn, ich will gleich mit den großen ketten und hotels loslegen. was denkst du?

Perfekt – du sprichst exakt die **entscheidende Perspektive** an:

Nicht *was Matbakh kann*, sondern **was es im Leben des Gastronomen verändert**.

💡 "Show me the steak, not the grill."

Die Technik ist unsichtbar – das **Ergebnis** ist sichtbar, spürbar, begehrenswert.

✅ Zielsetzung für die folgende Matrix:

Wir bauen ein **strategisches Messaging-Modell**, das zeigt:

- **Wem** du hilfst (Persona-Typ)
- **Welchen „Job“** du für sie erledigst
- **Welches konkrete Ergebnis** du bringst („The Done Job“)
- **Wie sich das im Leben des Gastronomen anfühlt oder auswirkt**
- **Welches „Ohr“** du dabei ansprechen musst (für Kommunikation)

📦 Struktur der Tabelle:

| Persona-Typ | Matbakh bringt in ihr Leben... | Der Job, den Matbakh erledigt | Ergebnis (the done job)
| Gefühl / Wirkung | Haupt-Ohr für Kommunikation |

🔧 1. Persona-Typ: **Der Macher (klassischer Lokalinhaber)**

Kategorie	Inhalt
Matbakh bringt...	Weniger Stress, mehr Gäste, ohne sich mit Technik befassen zu müssen
Job-to-be-done	Gäste erreichen, sichtbar sein, ohne Zeit oder Digitalwissen
Ergebnis	Google & Instagram-Profil sind top, Gäste buchen direkt, Umsatz steigt
Wirkung im Alltag	Weniger Frust mit leeren Tischen, weniger Kopfweh wegen Social Media
Kommunikations-Ohr	💰 Effizienz-Ohr – direkt, konkret, problemlösend

🎨 2. Persona-Typ: **Die Markenbauerin (stylisches Café, Konzeptgastronomie)**

Kategorie	Inhalt
Matbakh bringt...	Sichtbarkeit, die dem eigenen Markenimage entspricht
Job-to-be-done	Ästhetisch & professionell auf Google & Instagram wirken
Ergebnis	Online-Auftritt sieht so gut aus wie das Café selbst, Kunden kommen „wegen Insta“
Wirkung im Alltag	Stolz, Wiedererkennungswert, mehr Likes & echte Buchungen

Kategorie	Inhalt
Kommunikations-Ohr	❤️ Werte-/Image-Ohr – ästhetisch, identitätsbasiert, Community-nah

📈 3. Persona-Typ: **Der Visionär (Franchise, Expansion, Premiumkonzepte)**

Kategorie	Inhalt
Matbakh bringt...	Daten, um die besten Entscheidungen für Standorte & Marketing zu treffen
Job-to-be-done	Digitales Performance-Monitoring für mehrere Locations
Ergebnis	Standort A performt 35 % besser auf Google, Standort C braucht mehr Sichtbarkeit
Wirkung im Alltag	Klare Entscheidungen treffen, Investitionen gezielt steuern
Kommunikations-Ohr	🧠 Kompetenz-Ohr – datengetrieben, strategisch, kontrollierbar

👤 4. Persona-Typ: **Die Gastgeberin / Veranstalterin (Event- & Community-orientiert)**

Kategorie	Inhalt
Matbakh bringt...	Sichtbarkeit für Events, themenbasierte Besucheransprache
Job-to-be-done	Veranstaltungen digital sichtbar & ausbuchbar machen
Ergebnis	Events schneller voll, mehr Anmeldungen ohne extra Werbeaufwand
Wirkung im Alltag	Planungssicherheit, Freude am Austausch, Community wächst
Kommunikations-Ohr	👂 Neugier-/Erlebnis-Ohr – emotional, lebendig, mit Storytelling arbeiten



Gesamtüberblick als Tabelle

Persona-Typ	Matbakh bringt...	Job-to-be-done	Ergebnis (Done Job)	Wirkung	Haupt-Ohr
Der Macher	Mehr Gäste ohne Technikstress	Sichtbar sein ohne Aufwand	Mehr Buchungen, bessere Präsenz	Weniger Frust, mehr Umsatz	💰 Nutzen
Die Markenbauerin	Sichtbarkeit im eigenen Stil	Ästhetisch sichtbar sein	Kunden kommen wegen des Looks	Stolz, Markenaufwertung	❤️ Werte
Der Visionär	Skalierbare Sichtbarkeitskontrolle	Standort-Performance messen	Klarheit über Stärken & Schwächen	Strategische Kontrolle	🧠 Kompetenz
Die Gastgeberin	Sichtbarkeit für Events	Veranstaltungen bewerben	Mehr Gäste & Reichweite	Community wächst, Planbarkeit	👂 Neugier

📢 **Empfehlung zur Kommunikation**

👉 Du zeigst nicht *Features*, sondern:

- eine **Veränderung im Alltag**

- eine **Entlastung**
 - ein **Gefühl von Kontrolle, Wirkung oder Stolz**
-

Und deine Frage:


| „Soll ich mit kleinen Betrieben oder direkt mit Ketten starten?“

Meine klare Empfehlung:

 **Start mit Einzelgastronomen + inspirierenden Marken (Markenbauerinnen)**

Denn:

- Du hast direkten Zugang (lokales Netzwerk)
- Du bekommst echte Testimonials („Seitdem buchen 30 % mehr Gäste über Google“)
- Du schaffst Proof of Value mit minimalem Overhead
- Du sammelst qualitative und quantitative Daten für spätere B2B-Scale (Franchises)

 Ketten/Hotels kommen ins Spiel, **sobald du datenbasiert zeigen kannst**, dass Sichtbarkeit → Umsatz führt. Dann kannst du dieselbe Logik auf größere Systeme skalieren.
