



Dominik Maximini

Agiler Sparringspartner für Entwicklungsleiter
dominik.maximini@valuerise-consulting.de

Skyler – dein virtueller agiler Coach



Ausgangslage

Im Jahr 2020 während der globalen Corona-Pandemie machten wir uns bei ValueRise Gedanken darüber, wie wir unseren Kunden in diesen schwierigen Zeiten helfen könnten. Unsere Dienstleistungen (Agile Coaching und Trainings zu agilen Themen) konnten wir zwar über digitale Kanäle erbringen, jedoch fehlte vielen Kunden das Budget, um uns zu beauftragen. Die größten Kostentreiber bei Agile Coaching sind dabei die Tagessätze der Berater. Wir starteten unsere Ideenfindung also mit dem Ziel, diese Tagessätze abzuschaffen und uns damit ultimativ als Agile Coaches überflüssig zu machen. Am Ende des Kreativprozesses stand die Idee, eine künstliche Intelligenz (KI) zu schaffen, die ähnlich wie ein Agile Coach aus Fleisch und Blut Fragen stellt und beantwortet. Hierüber sollten 80% der Leistungen externer Berater eingespart werden und so Liquidität bei unseren Kunden erhalten bleiben.

Die Entwicklung dieses Produktes erfolgte streng nach agilen Prinzipien und führender Prozess war dabei „Lean Startup“. Bei diesem Ansatz stellt man Hypothesen auf, validiert diese am Kunden durch Produkte mit dem kleinstmöglichen Funktionsumfang und nähert sich so Schritt für Schritt einem marktfähigen Produkt an.

Virtueller Agile Coach

Eine künstliche Intelligenz (KI), die uns als Agile Coaches überflüssig macht.



Das Produkt

Wir nannten unser Produkt „Skyler“, was so viel wie „Schüler“ oder auch „Gelehrter“ bedeutet und sowohl für Männer als auch Frauen genutzt werden kann. Mit dem Namen drückten wir einerseits die Idee von Skyler aus (eine lernende KI, die anderen etwas beibringt) und bereiteten andererseits ein Feature vor, mit dem wir später die Nutzer auswählen lassen könnten, ob sie mit einem weiblichen oder männlichen Avatar im Produkt interagieren wollten.

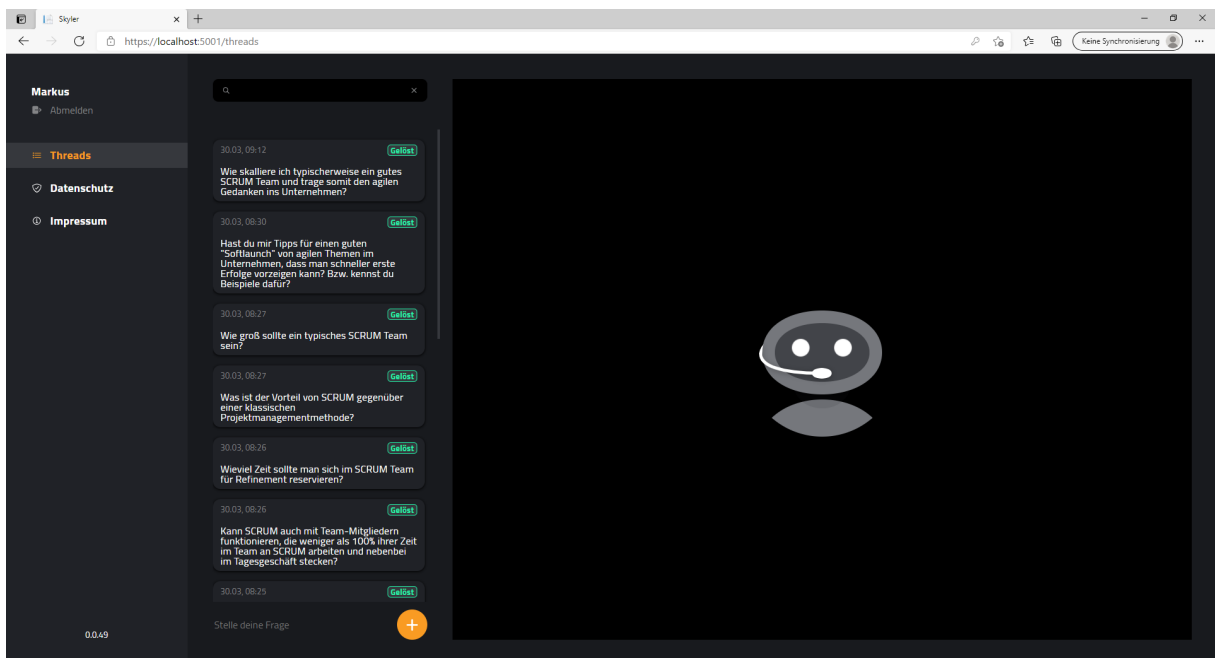


Abb. 1: Skyler aus Sicht der Nutzer (Übersicht)

Die erste zu testende Hypothese war: „Unsere Zielgruppe möchte mit einer künstlichen Intelligenz im Bereich des Agile Coachings interagieren.“

Dabei hatten wir als Zielgruppe vor allem Scrum Master und Product Owner mit geringer Erfahrung ausgewählt. Also genau die Zielgruppe, mit der wir auch in unserer Arbeit als Agile Coaches am häufigsten interagieren. Eine spezielle Branche oder Firmengröße war für uns nicht im Fokus.

Um diese Hypothese zu validieren, führte unser Product Owner zunächst Interviews mit etwa 15 Personen aus der Zielgruppe durch. Das Ergebnis war überwältigend: Nahezu alle gaben an, die Idee großartig zu finden. Allerdings wurde in den Befragungen auch deutlich, dass die Befragten sich nur schwer vorstellen konnten, wie Skyler funktionieren würde.

Daraus ergab sich die zweite zu testende Hypothese: „Unsere Zielgruppe interagiert gerne und häufig mit einer künstlichen Intelligenz im Bereich des Agile Coachings.“

Wir benötigten also einen Prototyp. Zusammen mit einem Partner (TZM GmbH in Göppingen) entwickelten wir einen „Wizard of Oz“-Prototyp. Dabei handelt es sich um ein Produkt, bei dem der Nutzer das Gefühl hat, er arbeite mit einer künstlichen Intelligenz. In Wahrheit sitzt am anderen Ende aber ein Mensch, der alle Interaktionen manuell steuert. Für Skyler bedeutete dies, dass der Nutzer glaubte, dass eine KI seine Fragen beantworten würde, dies in Wahrheit aber durch einen menschlichen Agile Coach erfolgte.

Für unseren Prototyp entwickelten wir eine einfache Webseite (mobile-optimiert) mit folgenden Hauptfunktionen:

- Ein User Interface, das annehmbar gut aussah, ohne übertrieben aufwändig gestaltet zu sein.
- Öffnen, schließen und beantworten von Threads, also Fragen des Nutzers an Skyler.
- Initiative Fragen durch Skyler, um den Nutzer zu eigenen Fragen zu animieren („Triggerfragen“).
- Einholen von Nutzerfeedback mittels Sterne-Bewertung und Freitext.
- Ein Begrüßungsworkflow, mit dem neuen Nutzern die Funktionsweise von Skyler nahegebracht wurde.
- Einen Demo-Modus, bei dem man automatische Antworten auf zuvor definierte Fragen hinterlegen konnte. So entstand bei Vorführungen der Eindruck, die KI würde eigenständig antworten.
- Einfache Telemetrie, also die Erfassung von Nutzungsdaten. Im Wesentlichen listeten wir auf, welche Nutzer Skyler wie häufig verwendeten.

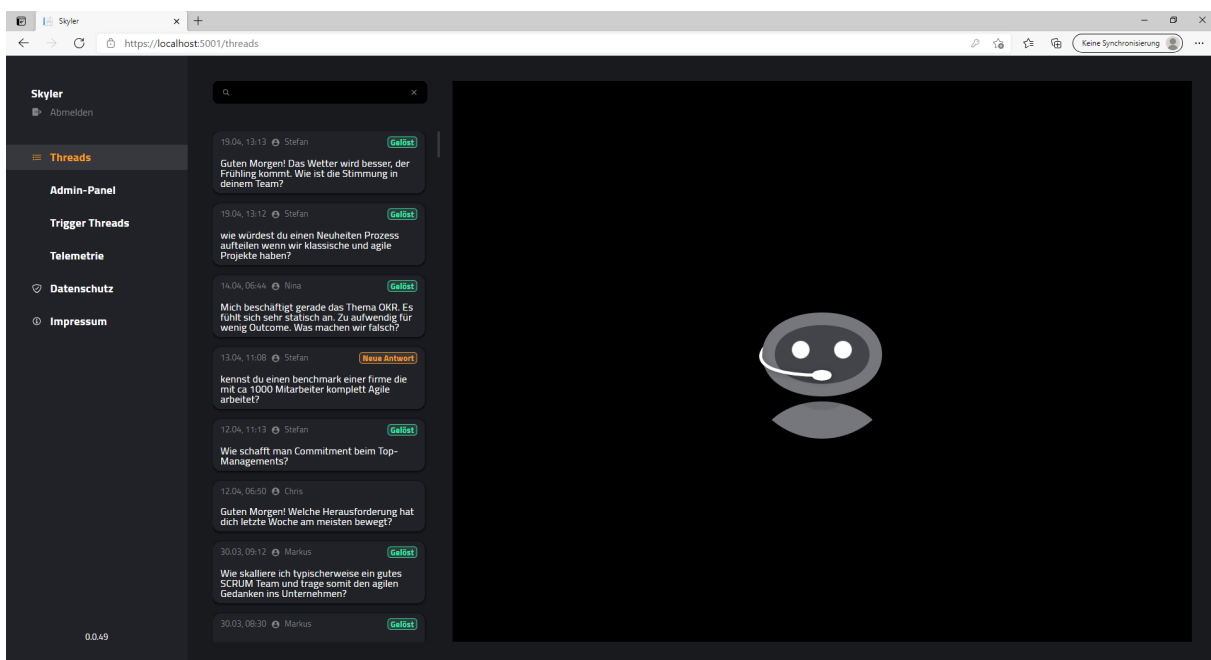


Abb. 2: Skyler aus Sicht des Experten



Die erste Testreihe

Für die Testreihe wählten wir aktuelle Consulting-Kunden und Schulungsteilnehmer aus. Wir boten etwa 40 Personen in persönlichen Gesprächen einen kostenfreien Testzugang an und demonstrierten ihnen die Funktionsweise. Etwa die Hälfte von ihnen schaltete ihren Zugang auch frei. Danach warteten wir einerseits auf initiativ gestellte Fragen der Nutzer und beantworteten diese manuell meist wenige Minuten nachdem die Frage gestellt worden war. Andererseits stellten wir einmal pro Woche eine „Triggerfrage“, um die Nutzer an Skyler zu erinnern und ihnen bei der Reflektion zu helfen. Beispielsweise fragten wir unsere Nutzer über Skyler, was diese Woche ihre größte Herausforderung gewesen sei.

Diese Testreihe dauerte drei Monate, in denen wir immer wieder andere Triggerfragen ausprobierten. Außerdem nahmen wir auch Personen mit in die Testreihe auf, die nicht zu unserer Zielgruppe gehörten. Dies betraf erfahrene Scrum Master, Product Owner und Agile Coaches sowie Führungskräfte. Wir erhofften uns dadurch, die Nebenhypothese („Unsere Zielgruppe sind unerfahrene Scrum Master und Product Owner“) ebenfalls zu validieren.

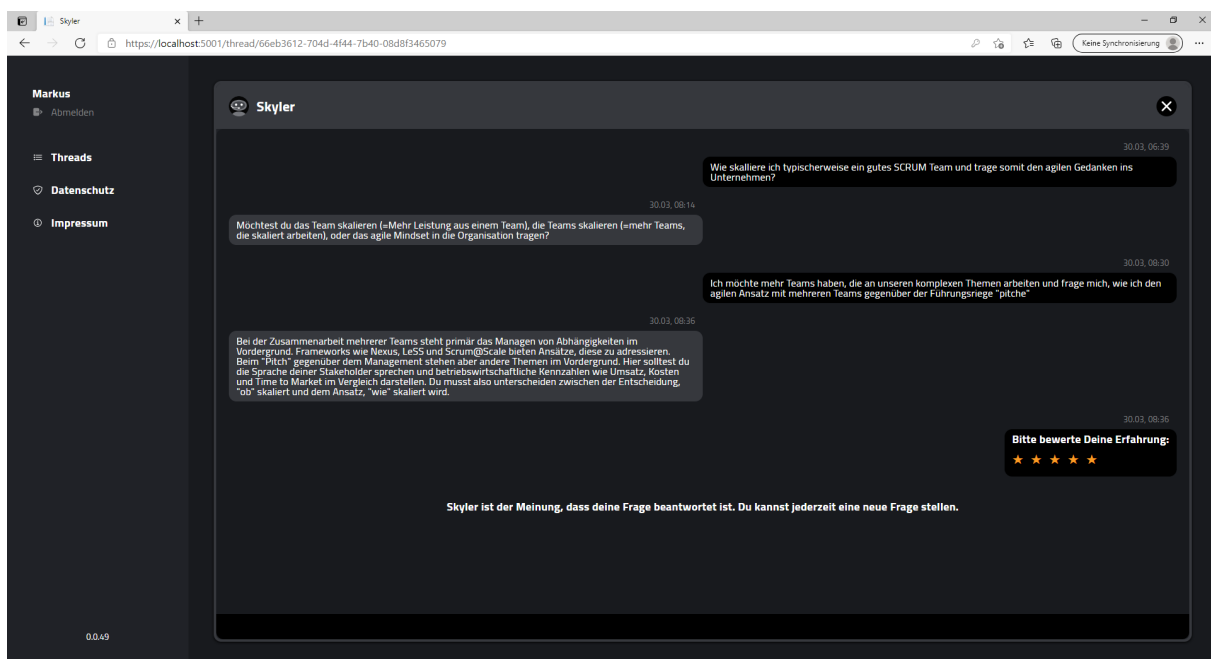


Abb. 3: Ein konkreter Thread aus Nutzersicht

Das Ergebnis der Testreihe entsprach in keiner Weise unseren Erwartungen. Viel weniger Personen als erwartet schalteten ihren Zugang frei und noch weniger nutzten ihn aktiv. Von den knapp 20 aktivierten Testnutzern waren nur fünf wirklich aktiv. Hiervon nutzten vier Personen Skyler nur, um uns einen Gefallen zu tun und die Qualität des Produktes zu testen.

Nur eine Person nutzte das Produkt, um reale Praxisprobleme zu lösen. Alle fünf Nutzer zeigten darüber hinaus das gleiche Nutzungsverhalten: Initial wurden ein paar Fragen kurz hintereinander gestellt. Danach brach die Nutzung ab und konnte nur durch unsere Triggerfragen kurzzeitig reaktiviert werden. Die Begrüßungsworkflows wurden kaum genutzt.

Am Ende der ersten Testphase bemühten wir uns um Gespräche mit allen Testbenutzern. Nur sechs der 20 widmeten uns die entsprechende Zeit. Das Ergebnis war, dass alle von der hohen Qualität der Antworten beeindruckt waren. Zwei Nutzer empfahlen die Weiterentwicklung des Produktes und die anderen vier gaben an, sich selbst alle Fragen beantworten zu können. Bei allen zeigten die Nutzungszahlen, dass sie, unabhängig von ihrer Empfehlung, Skyler nicht intensiv genutzt hatten. Vereinzelt kam auch der Hinweis, dass ihr Grund für die Nutzung des Produktes das Vertrauen war, dass ein von ihnen geschätzter Agile Coach (Dominik Maximini) die Qualität der Antworten garantierte.

„Klingt jetzt ein bisschen plump, aber: Es ist gut Zugriff auf dich zu haben (faktisch bestimmst du ja die Qualität der Antworten). Meine persönliche Nutzung ist aber zu niedrig, dass für mich irgendeinen paid Plan in Frage kommt. Außerdem bilde ich mir ein, viele alltägliche Fragen schon selbst beantworten zu können“

Pilotnutzer

Wir entschieden uns für eine zweite Testreihe mit einem unveränderten Produkt, um die Hypothese zu validieren, dass Menschen, die keinen direkten Bezug zu Dominik haben, das Produkt intensiver nutzen würden. Zugegebenermaßen lag in dieser Hypothese mehr Hoffnung als Glaube.



Die zweite Testreihe

Für die zweite Testreihe wählten wir Nutzer aus, die keinen Bezug zu Dominik hatten. Wir gewannen acht Personen aus der Zielgruppe für eine dreiwöchige Testphase. Zeitgleich ließen wir die Nutzer der ersten Testreihe aktiv im System und beobachteten deren Nutzungsverhalten weiter. Allerdings stellten wir diesen keine Triggerfragen mehr und prüften so die Hypothese: „Die Aktivität wird durch Triggerfragen signifikant gesteigert.“

Wir stellten fest, dass die im System verbliebenen Nutzer der ersten Testphase das System mit Ausnahme von einer Person nicht mehr nutzten. Interessanterweise war diese Ausnahme keine der Personen, die uns das Feedback gegeben hatten, Skyler weiterzuentwickeln. Die neu hinzugekommenen Testbenutzer zeigten ein identisches Nutzungsverhalten wie die Teilnehmer des ersten Testdurchgangs.

„Ich würde ziemlich vermessen als PO und ausgebildeter SCRUM Master von mir selbst behaupten, dass ich rund um SCRUM nicht wirklich viele Blankspots besitze. Meine Fragen an die KI dienten lediglich dazu, sie zu challengen“

Pilotnutzer

Im Feedbackgespräch gaben auch diese Nutzer an, dass sie ihren eigenen Bedarf nicht sähen. Sie hätten keine Fragen gehabt, die sie stellen müssten. Ein Use Case im eigenen Bereich wurde nicht erkannt, jedoch wurde die Vermutung geäußert, dass Menschen mit noch weniger Erfahrung als sie den Bedarf haben könnten.



Interpretationen

Wenn wir unsere Hypothesen und das Ergebnis der Experimente betrachten, entsteht folgendes Bild:

Hypothese	Interviews vor Prototyp	Testreihe 1	Testreihe 2
Unsere Zielgruppe möchte mit einer künstlichen Intelligenz im Bereich des Agile Coachings interagieren	validiert	-	-
Unsere Zielgruppe interagiert gerne und häufig mit einer künstlichen Intelligenz im Bereich des Agile Coachings	-	widerlegt	widerlegt
Unsere Zielgruppe sind unerfahrene Scrum Master und Product Owner	-	validiert	-
Der direkte Kontakt zu Dominik beeinflusst das Nutzerverhalten signifikant	-	-	widerlegt
Die Aktivität wird durch Triggerfragen signifikant gesteigert.	-	-	validiert

Tabelle 1: Hypothesen

Der angegebene Hauptgrund für die geringe Nutzung war das Nichtvorhandensein von Fragen und Lernbedarf. Unserer Experteneinschätzung nach sind aber sowohl Fragen offen als auch Lernbedarf vorhanden, was wir stichprobenartig auch überprüft haben. Das Experiment zeigt, dass der eigentliche Mehrwert eines Agile Coaches nicht die Methodenkompetenz ist, sondern die Fähigkeit live vor Ort die richtigen Fragen zu stellen und durch persönliche Anwesenheit die Haltung des Coachees zu verändern. Ein Chatbot erfordert, dass der Nutzer seine Fragen selbst formulieren kann. Dies ist im Regelfall nicht gegeben. Ein Agile Coach vor Ort erkennt Handlungsfelder, stellt Fragen, die den Coachee zu einem Erkenntnisgewinn führen und bleibt auch manchmal hartnäckig bei der eigenen Expertenmeinung, während der Impuls einer künstlichen Intelligenz schon längst ignoriert wurde. Auch kann ein realer Agile Coach vor Ort vorleben, welche Haltungen und Aktionen erwartet werden. Dieses Vorleben gelingt einer Künstlichen Intelligenz nicht. Hinzu kommt eine gewisse Scheu der Anwender davor, einer KI delicate Themen anzuvertrauen. Dies mag sich in den nächsten Jahren möglicherweise ändern – Stand heute ist das aber für einige Nutzer ein Hinderungsgrund für eine wirklich offene Interaktion mit Skyler.

Die Fakten sprechen unterm Strich eine deutliche Sprache: Es ist uns nicht gelungen, eine aktive Nutzerbasis zu entwickeln. Die Kunden bevorzugen den persönlichen Kontakt zu uns als Agile Coaches aus Fleisch und Blut. Dabei schätzen sie unsere Fachexpertise und Gespräche auf Augenhöhe, in denen wir innovativen Ideen mit Offenheit und Lösungsorientierung begegnen, statt sie mit Bedenken zu überschütten. Darauf fokussieren wir uns auch weiterhin bei ValueRise.



Die Rolle des Product Owners

Dominik musste als Product Owner sicherstellen, dass alle Hypothesen mit dem geringstmöglichen finanziellen Aufwand validiert werden. Diese Aufgabe umfasste vier Phasen.

Phase 1: Vision und Strategie

Noch bevor es einen Prototyp gab, musste die Vision entwickelt werden. Dies taten wir bei ValueRise im Team und sprachen auch mit Personen außerhalb unserer Organisation darüber. Durch das positive Feedback ermutigt, führten wir eine intensive Wettbewerbsanalyse durch. Welche Apps und Produkte gab es noch am Markt? Was machten andere Branchen, zum Beispiel im Bereich der Depressionstherapie?

Interessanterweise entging uns einer der größten Wettbewerber: Internetsuchmaschinen. Erst im Verlauf der ersten Testreihe wurden wir durch einen Tester darauf aufmerksam gemacht, stellten aber glücklicherweise schnell fest, dass wir mit Skyler in puncto Qualität besser und im Bereich Geschwindigkeit schlechter waren, was für die Nutzer aber akzeptabel war.

Zur Vision gehörte es auch, einen Prototypen des User Interfaces zur Visualisierung darzustellen. Dies erfolgte sehr leichtgewichtig mit Standardgrafiken, was den Stakeholdern enorm half, das Vorhaben zu verstehen. Fast alle Diskussionen und Interviews wurden anhand dieses grafischen Prototypen geführt. Das gilt auch für die Kooperationsgespräche mit unserem Partner, der TZM GmbH.

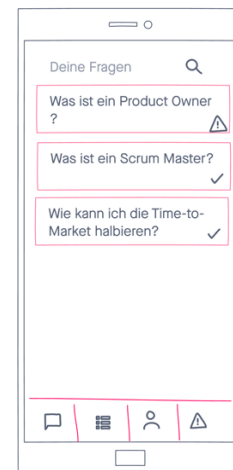


Abb. 4: Erster UI-Prototyp zur Visionsvermittlung

Phase 2: Prototypenbau

Nachdem wir beschlossen hatten zusammenzuarbeiten, begann die TZM damit, den Wizard-of-Oz-Prototypen zu entwickeln. Dies erfolgte in kurzen Sprints nach Scrum. Die Aufgabe des Product Owners war es dabei, alle Featurewünsche zu priorisieren. Das war besonders schwierig, weil Dominik die Vision maßgeblich mitentwickelt hatte und dafür brannte. Jetzt die eigenen Ideen zu streichen, war schmerzlich. Die allermeisten Featurewünsche wurden auf „nicht in Version 1“ verschoben, so dass wir nur die essenziellen Funktionen umsetzten. Perfekt waren wir natürlich nicht: So setzten wir beispielsweise den Begrüßungsworkflow um, den am Ende nur wenige Nutzer wirklich nutzten, der uns aber technisch vor große Herausforderungen stellte und viel Zeit kostete. In Summe kamen wir aber zügig voran und blieben in etwa im vorgegebenen Budget.

Phase 3: Pilotphase

Mit der Veröffentlichung des Prototypen änderte sich auch die Rolle des Product Owners wieder. Während es zuvor um die Erstellung der Vision und die technische Umsetzung ging, standen jetzt betriebswirtschaftliche Überlegungen im Vordergrund. Nur wenn wir eine breite Nutzerbasis gefunden hätten oder wenigstens einzelne Nutzer, die für das Produkt brannten, hätten wir weiter in die technische Entwicklung, insbesondere eine funktionierende KI, investiert. Der Product-Owner-Alltag war daher geprägt von enger Interaktion mit den Nutzern und der Auswertung von Kennzahlen. Die technische Produktentwicklung stand in diesem Zeitraum still.

Phase 4: Das Ende

Nachdem es uns in zwei Testreihen nicht gelungen war die Haupthypothese zu validieren, sahen wir nach einer Kosten-/Nutzen-Abwägung, dass dieses Produkt keinen wirtschaftlichen Erfolg haben würde. Daher musste Dominik als Product Owner entscheiden, ob eine Änderung der Zielgruppe, des Produktes oder des Testaufbaus Sinn machen könnte. Andernfalls wäre die Entwicklung einzustellen. Nach vielen Gesprächen und Analysen war klar: Die Idee war zwar hervorragend – aber leider nur aus unserer Sicht. Ein funktionierendes Geschäftsmodell fanden wir nicht. Dominik entschied daher konsequent unter Berücksichtigung aller Erkenntnisse, das Produkt einzustellen. Das war eine sehr emotionale Entscheidung, denn es steckten bereits jede Menge Herzblut und viele intensive Arbeitsstunden im Produkt. Trotzdem war es die richtige Entscheidung, denn die weitere Entwicklung bis zur Marktreife hätte einen hohen sechsstelligen Betrag an Investitionen erfordert, deren Amortisation bestenfalls unwahrscheinlich gewesen wäre.



Ergebnis

Wir betrachten das Ergebnis mit einem lachenden und einem weinenden Auge. Einerseits sind wir traurig, dass unsere Idee im Markt nicht so angenommen wurde, wie wir das erwartet hatten. Andererseits sind wir froh, dass wir durch die agile Vorgehensweise schon früh gemerkt haben, dass unsere Produktinnovation nicht funktioniert. So konnten wir eine Menge Zeit und Geld sparen. Nach nur sechs Monaten wussten wir sehr genau, wo wir richtig und falsch lagen und ob weitere Investitionen Sinn machen würden. Außerdem konnten wir die Aufgaben eines Product Owners im Lebenszyklus eines Produkts demonstrieren und durchleben, was eine wertvolle Erfahrung war. Wir sehen es so: Wir haben mit Skyler eine gute Investition in unsere persönliche Entwicklung getätigt.

Außerdem konnten wir bestätigen, dass unsere Kunden den persönlichen Austausch mit uns sehr schätzen und Mehrwerte erzielen, die sie selbst zunächst gar nicht erkennen. Erst durch die persönliche und individuelle Interaktion auf Augenhöhe entstehen Erkenntnisgewinne und Optimierungen, die zu zufriedeneren Mitarbeitern und wertvolleren Ergebnissen führen.

Unsere **Ergebnisse:**

Vom Markt abgelehnte Produktidee
Frühes Scheitern
Voller Lebenszyklus eines Product Owners, komprimiert in 6 Monaten
Bestätigung des restlichen Geschäftsmodells

Die Schlüssel für diesen Erfolg waren:

- Der Fokus auf schnelle Prototypen mit minimalem Funktionsumfang.
- Schnelles Feedback aus dem Markt.
- Das Beibehalten eines hohen Fokus.
- Angeregte Diskussionen auf Augenhöhe zwischen Kunden, TZM und ValueRise Consulting.
- Die Fähigkeit, loszulassen.