

KI-gestützte Implementierung für Legacy-Modernisierung

Nativ oder über ein Fachmodell: Kosten, Aufwand, Qualität und Risiken im Vergleich

DIE KERNAUSSAGE

Beide Wege nutzen KI für die Umsetzung. Der Unterschied liegt in der Eingabe: Bei der nativen KI-Umsetzung gehen Anforderungen direkt in die Umsetzung, bei der KI-gestützten Umsetzung aus einem Fachmodell gehen geprüfte, belegte Vorgaben hinein. Das Fachmodell verlagert Aufwand nach vorn – in Fachprüfung und Modellaufbau – und senkt dafür Missverständnisse, Wiederholungsaufwand, Testschleifen und Risiken. Weil die IT schon beim Aufbau des Soll-Modells Prototypen für ausgewählte Bereiche baut und jedes Ergebnis ins Modell zurückfließt, ist das Vorgehen kein Wasserfall. Zusätzlich verbessert es das fachliche Know-how, die Qualität der Anwendung und ihre fachliche wie technische Dokumentation. Bei größeren Vorhaben spricht fast alles für das Fachmodell; bei kleinen, gut dokumentierten, kurzlebigen Vorhaben bleibt native KI oft die wirtschaftlichere Wahl.

01 Vorbemerkung: Richtungen, keine Messwerte

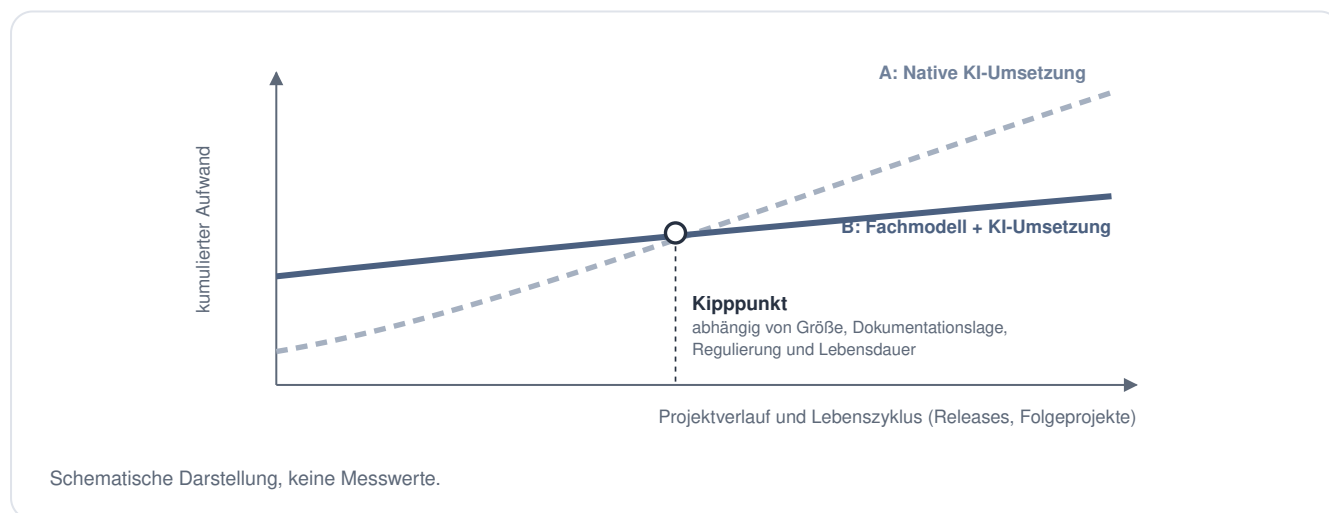
Verglichen werden zwei Wege, die beide KI nutzen. Weg A, native KI-Umsetzung: Änderungen werden als Anforderungen direkt in KI-Werkzeuge gegeben, die Code, Tests und Dokumentation erzeugen. Weg B, KI-gestütztes Fachmodell mit daraus KI-gestützter Umsetzung: Zuerst entsteht ein Ist-Modell, das die verfügbaren Quellen erschließt. Daraus wird ein verdichtetes Ist-Modell, reduziert um Überflüssiges und Redundantes. Dieses wird zum Soll-Modell erweitert, indem die neuen Anforderungen hinzukommen. Schon beim Aufbau des Soll-Modells baut die IT für ausgewählte Bereiche KI-gestützt Prototypen; deren Ergebnisse fließen als Delta ins Modell zurück und werden von einem Fachmitarbeiter bestätigt. Aus dem Soll-Modell werden Spezifikation, Tests und Dokumentation abgeleitet, und die IT setzt KI-gestützt um.

Eine belastbare vergleichende Studie zu genau diesen beiden Ansätzen gibt es nicht. Die folgenden Aussagen sind begründete Wirkungsrichtungen auf Basis der Funktionsweise beider Ansätze; Zahlenwerte entstehen erst mit eigenen Daten, am besten in einem Proof of Value (Abschnitt 7). Externe Anker gelten für KI-Einsatz allgemein, nicht für das Fachmodell: Laut IBM und BIAN überschreiten 94 % der Core-Banking-Modernisierungen den Zeitrahmen. McKinsey berichtet von 40 bis 50 % kürzeren Modernisierungszeiten mit generativer KI, warnt aber vor reinem „code and load“. Feldstudien zur Entwicklerproduktivität zeigen je nach Aufgabe und Kontext Gewinne und Verluste. KI-Wirkung hängt also am gesamten Arbeitssystem, nicht am Modell allein.

02 Wo im Projektverlauf die Kosten entstehen

Phase	A: Native KI-Umsetzung	B: Fachmodell + KI-Umsetzung
Start	Gering: Prompting, Kontextaufbereitung, schnelle erste Ergebnisse.	Höher: Quellenaufnahme, Scope- und Prüfumfang festlegen, erste Reviews.
Analyse und Klärung	Klärung geschieht verteilt in Chats und Gesprächen, oft erst bei Rückfragen aus der Umsetzung.	Klärung ist gebündelt vor der Umsetzung; Widersprüche sind als offene Punkte sichtbar.
Umsetzung	Schnell. Fachliche Interpretation übernimmt implizit Entwickler oder KI.	Ebenfalls schnell und KI-gestützt; schon beim Aufbau des Soll-Modells klären Prototypen für ausgewählte Bereiche fachliche Fragen früh am lauffähigen Ergebnis, gegen eine freigegebene Spezifikation mit Zusammenhang und Herkunft. Weniger IT-Implementierung, weil nur das verdichtete Soll umgesetzt wird.
Test und Abnahme	Fachliche Abweichungen werden spät entdeckt; zusätzliche Test- und Korrekturschleifen sind die Folge.	Testgrundlage folgt aus Regeln und Anforderungen; weniger fachliche Fehler in der Abnahme. Weniger Tests, weil nur das verdichtete Soll getestet wird.
Folgeänderungen	Kontext wird je Änderung neu aufgebaut; Wissen aus früheren Chats ist kaum wiederverwendbar.	Modell wird fortgeschrieben, auch aus Prototyp-Ergebnissen; Impact-Analyse zeigt betroffene Regeln, Komponenten und Tests.
Dokumentation	Nachträglich und manuell, oft unterlassen.	Ableitung aus dem Modell; fachliche Release-Notes als Differenz der Modellstände.

Schematisch: Weg A startet niedriger, Weg B verläuft später flacher; wo sich die Kurven kreuzen, hängt von Größe, Dokumentationslage, Regulierung und Lebensdauer ab.



03 Aufwand nach Rolle

Bereich	A: Native KI-Umsetzung	B: Fachmodell + KI-Umsetzung	Richtung
Fachaufwand	Wenig sichtbar am Anfang, dann punktuell, schwer planbar und hoch: Rückfragen, Klärungsrunden, Abnahme. Experten erklären ad hoc und oft mehrfach; Verständnis bleibt bei einzelnen Personen.	Früher und planbarer: Experten prüfen vorbereitete Aussagen, Widersprüche und Prüffragen statt zu erklären. Die Fachseite versteht dadurch nachhaltig, was im Projekt geschieht und entschieden wird. Das reduziert Missverständnisse und damit Wiederholungsaufwand. Der Prüfaufwand ist real und muss priorisiert werden.	Nach vorn verschoben; Gesamtaufwand geringer, weil Missverständnisse und Wiederholungsschleifen sinken. Das Review ist unverzichtbar, damit nichts Falsches im Modell stehen bleibt; ein abgegrenzter Scope macht seinen Aufwand planbar.
Testaufwand	Testfälle entstehen aus Anforderungen oder Code; fachliche Fehler zeigen sich in Testzyklen. Regressionsumfang schwer eingrenzbar.	Testfälle sind mit Regeln und Anforderungen verknüpft; Impact-Analyse grenzt Regressionstests ein. Weniger Tests, weil nur das verdichtete Soll getestet wird.	Geringer durch kleineren Testumfang, weniger fachliche Fehler und eingegrenzte Regression. Testaufwand entfällt nicht.
IT-Aufwand	Code entsteht schnell, doch die IT übernimmt fachliche Klärung, Interpretation und spätere Korrekturen.	Weniger Interpretation und Rückfragen; die KI-gestützte Umsetzung arbeitet gegen klare, freigegebene Vorgaben. Weniger IT-Implementierung, weil Überflüssiges und Redundantes vorher aus dem Modell entfällt.	Weniger Implementierungs-, Klärungs- und Nacharbeit.
Steuerung	Fortschritt wird an Code und Funktionen gemessen; Wissensstand bleibt unsichtbar.	Zweite Steuergröße Knowledge Coverage; Reife je Modul vor Migrationsbeginn prüfbar.	Bessere Steuerbarkeit, Aufwand für Kennzahl und Scope-Definition.

04 Gesamtkosten (TCO) über drei bis fünf Jahre

Für einen fairen Vergleich zählen nicht nur Projektkosten, sondern alle Kosten über den Lebenszyklus plus die erwarteten Risikokosten. Als Rechenrahmen genügt: $TCO = \text{Einmalkosten} + \text{Summe der Jahreskosten} + \text{erwartete Risikokosten}$, wobei sich Letztere aus Eintrittswahrscheinlichkeit mal Schadenshöhe der wesentlichen Risiken ergeben.

Kostenblock	A: Native KI-Umsetzung	B: Fachmodell + KI-Umsetzung
KI-Nutzung	Kontext wird je Aufgabe neu aufgebaut, daher wiederkehrender Verbrauch.	Kontext wird einmal aufgebaut und fortgeschrieben; je Arbeitsschritt passendes Modell.
Prüfung und Pflege	Gering geplant, später in Form von Korrekturen.	Fachliches Review und Abgleich nach Änderungen als geplanter Aufwand.
Nacharbeit und Testschleifen	Höher und spät sichtbar.	Geringer, auch durch weniger Implementierungs- und Testumfang.
Dokumentation und Nachweise	Nachträglich oder gar nicht; Audit-Aufwand entsteht punktuell.	Laufende Ableitung aus dem Modell.
Wissensverlust, Folgeprojekte	Wissen steckt in Chats und Köpfen; jedes Folgeprojekt beginnt weitgehend neu.	Wissen bleibt im Modell der Organisation und wird wiederverwendet.

Wann kippt die Rechnung? Zugunsten des Fachmodells sprechen große Codebasis, schwache oder widersprüchliche Dokumentation, regulatorische Nachweispflichten, lange Lebensdauer, mehrere Releases und Folgeprojekte sowie hohe Personenabhängigkeit. Zugunsten nativer KI sprechen kleiner Scope, gute Dokumentation, kurzlebige Lösungen, fehlende Nachweispflichten und verfügbare Experten.

05 Fachliches Know-how, Anwendungsqualität und Dokumentation

Neben Aufwand und Kosten entscheidet die Qualität dessen, was am Ende steht – und was die Organisation danach weiß.

Aspekt	A: Native KI-Umsetzung	B: Fachmodell + KI-Umsetzung	Tendenz
Fachliches Know-how	Wissen entsteht in Prompts, Chats und Köpfen einzelner; die Fachseite sieht Ergebnisse, nicht den Weg dorthin.	Wissen sammelt sich im geprüften Modell der Organisation; die Fachseite versteht und verantwortet, was umgesetzt wird.	Fachmodell
Fachliche Qualität der Anwendung	Regeln und Ausnahmen werden übernommen, soweit sie im Prompt oder im gelesenen Code auftauchen; Altentscheidungen bleiben unbewusst.	Jede bestehende Funktion erhält eine bewusste Entscheidung (beibehalten, ändern, entfallen); Regeln und Ausnahmen sind belegt übernommen.	Fachmodell
Technische Qualität	Schnell erzeugter Code gegen unscharfe Vorgaben; Nacharbeit bei Interpretationsfehlern.	KI-gestützte Umsetzung gegen eine konsistente, freigegebene Spezifikation; Architekturentscheidungen bleiben bei der IT.	Fachmodell
Fachliche Dokumentation	Nachträglich und manuell; veraltet mit dem nächsten Release.	Fachkonzept, Regeln, Prozesse und fachliche Release-Notes als Ableitung aus dem Modell.	Fachmodell
Technische Dokumentation	Generierte Beschreibungen und Kommentare; Bezug zu Anforderungen fehlt.	DV-Konzept, Schnittstellen, Spezifikation und Testgrundlage aus demselben Modell, mit Herkunft bis zur Quelle.	Fachmodell
Weiterentwicklung	Jede Änderung beginnt beim Verstehen neu.	Änderungen setzen auf geprüftem Wissen auf; Impact-Analyse zeigt Betroffenes.	Fachmodell

Dieser Vorteil entsteht nur, wenn der Prüfumfang realistisch gewählt wird und der Abgleich nach Änderungen im Änderungsprozess verankert ist.

06 Projektrisiken

Risiko	A: Native KI-Umsetzung	B: Fachmodell + KI-Umsetzung	Tendenz
Fachliche Fehlinterpretation	Hoch: Lesart wird implizit gewählt, Widersprüche werden geglättet.	Geringer: Widersprüche sind sichtbar und fachlich entschieden.	Fachmodell
Unentdeckte Abhängigkeiten	Hoch: nur das im Prompt Genannte wird beachtet.	Geringer: Beziehungen sind modelliert und abfragbar.	Fachmodell
Wissensverlust, Personenabhängigkeit	Hoch: Wissen in Chats und Köpfen.	Geringer: Wissen im Modell.	Fachmodell
Nachweis und Audit	Schwach: keine durchgängige Herkunftskette.	Stark: Quelle bis Test und zurück.	Fachmodell

Risiko	A: Native KI-Umsetzung	B: Fachmodell + KI-Umsetzung	Tendenz
Späte Erkenntnis (Wasserfall-Risiko)	Fachliche Erkenntnisse entstehen oft erst in Test und Abnahme.	Prototypen klären schon beim Aufbau des Soll-Modells fachliche Fragen früh; das Ergebnis fließt als Delta ins Modell zurück.	Fachmodell
Prototyp ersetzt Entscheidung	Der Code entscheidet implizit.	Leitplanken: Prototyp beantwortet eine Frage, die Entscheidung fällt im Modell; Rückgeführtes gilt als „exploriert“, bis der Fachbereich freigibt.	Neutral
Drift zwischen Modell und Realität	Nicht anwendbar, da kein Modell; Dokumentation fehlt ohnehin.	Real, aber durch Abgleich von Code und Modell mit fachlicher Bestätigung beherrschbar.	Neutral
Prüfumfang und Prüfspur	Geprüft wird später, in Test und Abnahme, ohne durchgängige Prüfspur; Falsches bleibt länger unentdeckt.	Geprüft wird früher und gezielt am Modell, mit Reviewstatus und Prüfspur; Falsches wird früher erkannt. Der Prüfumfang muss abgegrenzt und priorisiert werden.	Fachmodell
Scheinsicherheit durch Freigabe	Plausible Antworten ohne Prüfspur.	Falsch freigegebene Objekte sehen belegt aus; Reviewstatus und offene Punkte sind eigene Attribute.	Neutral
Anbieterabhängigkeit	Abhängig vom gewählten KI-Anbieter und -Modell.	Methode und Ergebnisse unabhängig vom KI-Modell; Komponenten innerhalb der vereinbarten Architektur austauschbar.	Neutral
Datenschutz, Datenabfluss	Abhängig von Werkzeug und Konfiguration.	Betrieb in Kundenumgebung; lokale oder kundenseitig gewählte KI.	Je Konfiguration

07 Mit eigenen Zahlen prüfen: der Proof of Value

Entscheidend ist, die Hypothesen dieses Papiers an einem eng begrenzten Ausschnitt zu messen, mit vorab vereinbarten Abnahmekriterien und einer Baseline, also einem Vergleichslauf mit dem bisherigen oder dem Prompt-Vorgehen. Sinnvolle Messgrößen sind:

- **Aufwand:** Stunden je Rolle (Fachbereich, Analyse, IT, Test) einschließlich Review.
- **Qualität:** bestätigte Befunde, nutzbare Ergebnisse, Rückfragen je Anforderung, Rückführungs-Delta aus Prototypen.
- **Test:** Testzyklen bis zur Abnahme, Fehler fachlicher Ursache, Regressionsumfang.
- **Wissensstand:** Knowledge Coverage (Belegquote, Verknüpfungsgrad, Konfliktklärung, Durchgängigkeit) und Zahl ungeklärter kritischer Abhängigkeiten als eigene Größe.
- **Verständnis und Dokumentation:** Verständnis der Fachseite, Vollständigkeit und Aktualität der fachlichen und technischen Dokumentation.
- **Wirtschaftlichkeit:** Hochrechnung auf den Gesamtumfang und TCO-Vergleich über drei bis fünf Jahre inklusive Risikokosten.

Die Entscheidungsregel (etwa Mindestwerte je Messgröße) sollte vor Beginn festgelegt werden, damit das Ergebnis nicht nachträglich passend gemacht wird.

FAZIT

Native KI-Umsetzung ist günstig im Start und riskant im Kern: Sie spart Aufwand dort, wo er sichtbar ist, und verlagert ihn nach Test, Abnahme und Folgeprojekte. Das KI-gestützte Fachmodell mit anschließender KI-gestützter Umsetzung kostet früher, macht den Aufwand planbar und sorgt dafür, dass die Fachseite versteht, was geschieht. Es liefert belegtes Know-how, eine bessere Anwendungsqualität und aktuelle fachliche wie technische Dokumentation. Welcher Weg wirtschaftlicher ist, entscheiden Größe, Alter, Regulierung und Lebensdauer des Systems – am besten gemessen am eigenen Fall.

Hinweis: Aussagen zu FIN-AI sind Herstellerangaben der FINIUS GmbH und nicht extern validiert. Die Bewertungen in den Tabellen sind Einschätzungen, keine Messwerte. Genannte Studien: IBM Institute for Business Value mit BIAN (2025), McKinsey (2024).

FINIUS GmbH

Beratung seit 2000 · BSM-Methode · FIN-AI

finius.de
fin-ai@finius.de