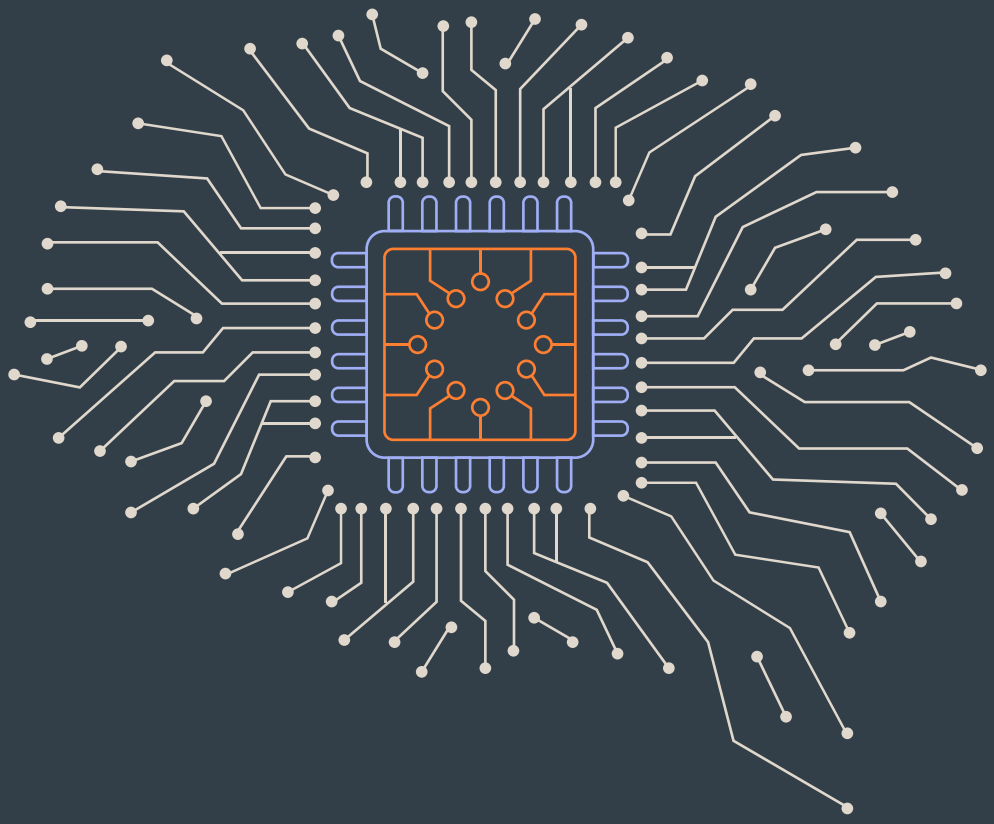
d-fine

ampega.

Anwendung von Agentic AI in der Modellvalidierung – Eine Case Study



Juni 2026

Über uns

ampega.

Die Ampega Gesellschaften sind der Vermögensverwalter der Talanx-Gruppe und einer der führenden bankenunabhängigen Asset Manager Deutschlands. Neben der Steuerung der Kapital- und Immobilienanlagen der Gruppe bietet Ampega institutionellen Investoren maßgeschneiderte Investment- und Servicelösungen entlang der gesamten Wertschöpfungskette. Insgesamt verwaltet Ampega Kapital- und Immobilienanlagen in Höhe von über 201 Mrd. Euro (Stand: 31.03.2026).

www.ampega.de

d-fine

d-fine ist ein europäisches Beratungsunternehmen mit Fokus auf analytische und quantitative Herausforderungen und die Entwicklung nachhaltiger technologischer Lösungen. Die Kombination aus 2.000 naturwissenschaftlich geprägten Mitarbeiterinnen und Mitarbeitern und langjähriger Praxiserfahrung ermöglicht passgenaue, effiziente und nachhaltige Umsetzungen für unsere mehr als fünfhundert Kunden aus allen Wirtschaftsbereichen und dem öffentlichen Sektor.

www.d-fine.com

Inhaltsverzeichnis

01. Einleitung und Problemstellung

Große Sprachmodelle (LLMs) und die darauf aufbauende agentische künstliche Intelligenz (KI) markieren derzeit einen technologischen Wendepunkt und werden intensiv diskutiert. Obwohl viele Financial-Services-Unternehmen bereits erste Schritte unternommen haben, die Potenziale dieser neuen Techniken zu heben, liegt der Fokus häufig immer noch darauf, einzelne bestehende Prozesse punktuell durch KI-basierte Werkzeuge zu ergänzen, etwa in Form der Nutzung von Chatbots zu Recherchezwecken oder zur Überarbeitung von Texten.[1] Hierbei bleibt jedoch ein Großteil des Potenzials zur Produktivitätssteigerung ungenutzt. Möglichkeiten zur signifikanteren Effizienzgewinn ergeben sich bei tiefer Verankerung von KI, und insbesondere deren Erweiterung zu agentischer KI, in die Betriebsprozesse selbst.

Um das Potenzial der Effizienzsteigerung durch Integration von agentischer KI in die Betriebsprozesse zu evaluieren, haben die Ampega Asset Management GmbH und die d-fine GmbH gemeinsam einen Proof-of-Concept im Kontext des fachlich und regulatorisch besonders anspruchsvollen Prozesses der Validierung von Risiko- und Bewertungsmodellen durchgeführt.

Die Modellvalidierung gilt als einer der anspruchsvollsten Prozesse des Risikomanagements: Sie geht weit über mechanistisches Testen hinaus und erfordert die systematische Prüfung der Angemessenheit getroffener Modellannahmen, verwendeter Daten, Methoden und Implementierungen. Hinzu kommen hohe regulatorische Anforderungen an Nachvollziehbarkeit, Dokumentation und Governance bei der Validierung von Risikomodellen. Ergeben sich Änderungen in den unterliegenden Modellen, Portfolien und Daten, so entstehen wechselnde Fokusthemen in der Validierung, welche über eine mechanische Automatisierung nur schwer abzubilden sind.

Hier versprechen agentische KI-Systeme Fortschritte: Werden sie direkt in den Prozess integriert, so kombinieren sie u. a. LLMs mit Datenzugriff („Talk to your data“) oder dem Einsatz von Tools und erlauben es, Routineaufgaben zu automatisieren, Analysen zu beschleunigen und menschliche Expertinnen und Experten zu entlasten. Dabei stellen sich jedoch auch grundsätzliche Fragen: Wie lässt sich ein solcher Einsatz von KI mit der Anforderung vereinbaren, dass die Validierungsfunktion selbst die Verantwortung für das Validierungsergebnis übernehmen muss? Wie hoch ist der Aufwand, alle relevanten Daten und Dokumente wie Validierungsskripte und Modelldokumentationen der Nutzung durch KI-Agenten zugänglich zu machen?

Im Rahmen eines Proof-of-Concept-Projekts (PoCs) mit Fokus auf den Bewertungs- und Marktrisikomodellen im Asset Management haben die Ampega und d-fine diese Fragestellungen gemeinsam untersucht. Das Ergebnis: Die Effizienz des Modellvalidierungsprozesses kann bei vertretbarem initialem Aufwand durch eine systematische Durchdringung mit agentischer KI signifikant gesteigert werden. KI kann menschliche Expertinnen und Experten bei verschiedensten wiederkehrenden Tasks entlasten, sie bei der Durchführung zusätzlicher Analysen unterstützen und dadurch Freiräume schaffen, sich auf ökonomisch-inhaltliche Kernfragestellungen sowie auf die Qualitätssicherung zu konzentrieren.

1 Vgl. „KI in Banken 2025“, cofinpro und VÖB-Service, https://www.voeb-service.de/aktuelles/detail/studie-vom-hype-zur-realitaet-78-prozent-der-banken-setzen-ki-produktiv-ein

02. Regulatorische Anforderungen

Bei der Verwendung von KI als Bestandteil eines zentralen, regulatorisch geforderten Prozesses ist die Berücksichtigung sämtlicher relevanter regulatorischer und interner Anforderungen von grundlegender Wichtigkeit.

In diesem Kontext ist zunächst anzumerken, dass es sich bei zur Modellvalidierung verwendeten agentischen KI-Verfahren typischerweise nicht selbst um Modelle im Sinne der einschlägigen Regulierungen (SR 26-2, SS 1/23, MaRisk AT 4.3.4, Derivateverordnung i. V. m. KAGB, …) handelt, insbesondere wenn sie keine eigenen quantitativen Schätzungen unter Verwendung eigener Annahmen durchführen.

Beim Einsatz von KI in der Modellvalidierung sind ferner interne Anforderungen zu beachten, insbesondere den Schutz auch nicht personenbezogener Daten (etwa Modelldokumentation oder Modellparameter) betreffend. Idealerweise werden solche Anforderungen anwendungsfallübergreifend festgesetzt, etwa in Form einer KI-Strategie. Die verbleibenden technologischen Risiken sind, unabhängig vom konkreten Anwendungsfall, zentral zu steuern.[2] Im Rahmen des PoCs wurden ausschließlich Cloud Services und LLMs einer firmenintern vorgegebenen Whitelist verwendet, und es wurde sichergestellt, dass keine Daten die EU verlassen.

Eine weitere wesentliche regulatorische Anforderung an die Modellvalidierung besteht darin, dass menschliche Expertinnen und Experten alle Validierungsschritte verstehen und nachvollziehen können müssen und dass sie letztlich die Verantwortung für das Ergebnis der Modellvalidierung zu übernehmen haben. Um dies zu unterstützen, wurde das im PoC verwendete agentische KI-System entsprechend ausgestaltet. Insbesondere wurden alle von der KI genutzten Zwischenergebnisse (z. B. Input- und Outputdaten ausgeführter Funktionen oder erzeugter Code), das sogenannte „Agentic Thinking“, gespeichert und für den Menschen einsehbar gemacht. Die Konsistenz und Übereinstimmung von KI-generierten Inhalten bei wiederholtem Aufruf mit gleichen Eingabedaten werden fortlaufend automatisch anhand spezifischer KPIs, z. B. anhand der Kosinus-Ähnlichkeit, gemessen und durch sorgfältiges Design von System-Prompts[3] sichergestellt. Der Mensch verbleibt stets „im Loop“, wird zum Review angehalten und kann auf einzelne Validierungsschritte explizit Einfluss nehmen, beispielsweise durch manuelles Editieren der generierten Inhalte oder durch Änderung der Prompts.

2 Vgl. hierzu auch die BaFin-Orientierungshilfe zu IKT-Risiken beim Einsatz von KI in Finanzunternehmen (https://www.bafin.de/SharedDocs/Downloads/DE/Anlage/dl_Anlage_orientierungshilfe_IKT_Risiken_bei_KI.html), die den Einsatz von KI-Systemen aus Perspektive des IKT-Risikomanagements nach DORA betrachtet und insbesondere deren Einbindung in den IKT-Risikomanagementrahmen, Governance- und Organisationsfragen sowie Cloud-, Drittparteien-, Cyber- und Datensicherheitsaspekte adressiert. Für den vorliegenden Kontext ist insbesondere relevant, dass KI-Systeme dort als Kombination von IKT-Assets und IKT-Infrastruktur verstanden werden und – analog zu anderen IKT-Assets – in den IKT-Risikomanagementrahmen zu integrieren sind.

3 System Prompts werden initial festgelegt und gestalten, wie das LLM grundsätzlich antwortet und in welcher Rolle es agiert (z.B. in der Rolle einer Marktrisiko- oder Bewertungsexpertin).

03. Technologischer Lösungsansatz

Ein agentisches System ist ein Netzwerk autonom, zugleich jedoch koordiniert agierender Agenten. Jeder Agent besteht aus einer LLM-Instanz, einem System-Prompt (der die Rolle des jeweiligen Agenten beschreibt) und einer Menge an Tools. Die Koordination wird in dem im PoC verwendeten Setting durch die Graphenstruktur des Netzwerks sichergestellt: Vom allgemeinen Supervisory Agent abgesehen hat dort jeder Agent genau einen „Vorgesetzten", der ihn mit Informationen und Aufgaben versorgt und an den er zurückmeldet. Im gewählten Aufbau werden Informationen zwischen Agenten ferner lediglich „vertikal", also nur zwischen einem Agenten und seinem unmittelbaren Vorgesetzten, ausgetauscht. Alle weiteren IT-relevanten Berechtigungen, z. B. für das Lesen von Dateien oder das Ausführen von Programmcode, sind explizit in der Konfiguration des agentischen Systems berücksichtigt und fixiert worden.

Der Supervisory Agent hat die Aufgabe, einen konkreten Lösungsplan zu erstellen, die zur Abarbeitung des Plans benötigten Agenten zu identifizieren und zu koordinieren sowie die Ergebnisse der einzelnen Agenten zu konsolidieren. Die dem Supervisory Agent untergeordneten spezialisierten Agenten verfügen hierbei gegebenenfalls über dedizierte Tools etwa zur Extraktion von Informationen aus Datenbanken, zur Ausführung von Code, zur Durchführung von Web-Suchen oder auch zur Qualitätssicherung von Zwischenergebnissen. Supervisor und untergeordnete Agenten nutzen dabei nicht zwingend Instanzen desselben LLMs. Über spezialisierte Systemprompts sind alle Agenten mit Denkmustern, Lösungsansätzen und Rollen konfigurierbar. Auch hierbei findet menschliches Wissen Eingang in die Struktur eines agentischen KI-Systems.

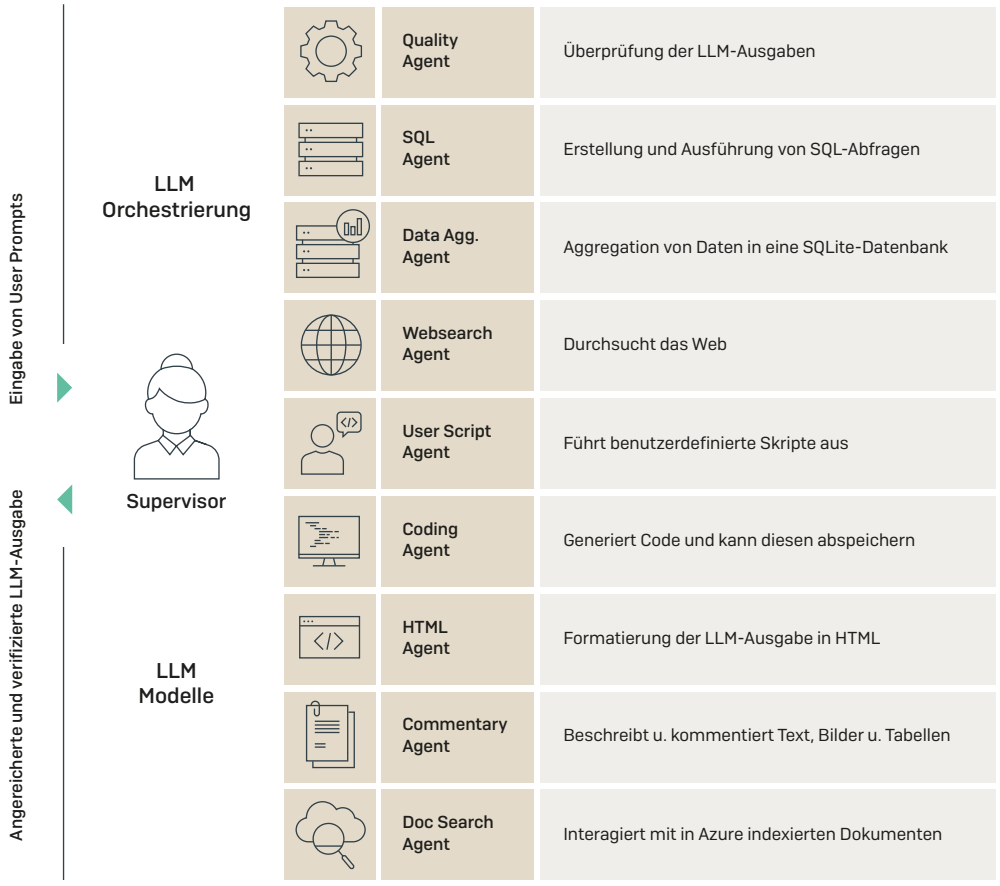


Abbildung 1: Übersicht über die im PoC eingesetzten Agenten

Hiermit sind nun klar vorstrukturierte Workflows (schematisch etwa: „Führe die Analyse X aus und bewerte das Ergebnis hinsichtlich der Schwellwerte aus dem Validierungskonzept."), aber auch strukturell variable Workflows oder ad hoc Analysen (schematisch etwa: „Erkläre die im vorherigen Validierungsschritt beobachteten Ausreißer.") abbildbar. Insbesondere in letzterem Fall spielt die Autonomie der Agenten ihre Stärken aus, da sie es erlaubt, Standard-Analysen und Tools situationsgerecht zu rekombinieren. Ein genauerer Einblick in die im Zuge des PoCs umgesetzten Analysen wird im nächsten Abschnitt gegeben.

Ein Spezialfall des agentischen Designs soll hier aufgrund seiner häufigen Anwendung in den PoCs hervorgehoben werden: Talk-to-Your-Data. Hierunter versteht man agentische Systeme, die auf die Erstellung von Datenbankabfragen sowie die Aggregation von Daten spezialisiert sind. Maßgeblich für die erzielbare Ergebnisqualität – insbesondere bei großen und komplexen Datenmengen – ist hierbei eine präzise Kontextualisierung der Agenten, etwa durch relevante Metadaten, Schemata und fachliche Logik.

Zur Unterstützung der Validierenden wurde ferner ein unabhängiger Quality Agent implementiert. Dieser prüft alle Ergebnisse auf Vollständigkeit und Konsistenz, sowohl in Bezug auf die in den Prompts gestellten Fragen als auch über verschiedene Abschnitte hinweg. Trotz dieses automatisierten ersten Quality-Gates verbleibt die Verantwortung für die Validierung, und die Qualität deren Ausführung bei den Validierenden.

Wichtig ist zu betonen, dass die Qualität aller Ergebnisse signifikant vom Prompt-Design abhängt. Kein noch so spezialisiertes agentisches Netzwerk kann unzureichende User Prompts kompensieren. Zur Entwicklung geeigneter Prompts für die Erzeugung stabiler und konsistenter Ergebnisse wurden im PoC neben der Orientierung an eigens entwickelten „Best Practice Prompting Guidelines"[4] auch geeignete KPIs definiert und automatisiert ausgewertet, um Verbesserungspotenziale in den Prompts zu identifizieren. Ein Beispiel für eine solche Metrik ist die sogenannte Kosinus-Ähnlichkeit, die darauf abzielt, die Ähnlichkeit von KI-Ausgaben über mathematische Verfahren in einer Kennzahl zu bündeln.

4 Diese wurden innerhalb des gemeinsamen Projekts entwickelt und sind nur intern zugänglich.

04. Mehrwert für die Modellvalidierung

04.01 Validierung von Bewertungsmodellen

Ausgangslage

Bei der Validierung von Bewertungsmodellen sind sowohl qualitative als auch quantitative Aspekte von Bedeutung (vergleiche z. B. §28 KARBV, AT 4.3.4 Tz. 5 MaRisk). Erstere umfassen beispielsweise die Prüfung der grundsätzlichen Eignung des Modells für die gestellte Aufgabe (etwa in Bezug auf die zu bewertenden Produkte und die angenommene Marktlage), während Letztere insbesondere die Prüfung der Korrektheit der Implementierung und der numerischen Stabilität des verwendeten Bewertungsmodells abdecken.

In der quantitativen Validierung wird der Aufwand wesentlich durch die Strukturierung der Analysen, die Sichtung relevanter Quellen sowie die Beurteilung der Angemessenheit des Modells für die konkreten Anwendungsfälle getrieben. In der quantitativen Validierung hingegen liegt ein wesentlicher Aufwandstreiber in der Entwicklung unabhängiger Challenger- und Benchmark-Modelle.

Integration agentischer KI in den Validierungsprozess

Agentische KI-Systeme unterstützen die Validierenden in beiden Bereichen. Im Rahmen der qualitativen Validierung werten sie verschiedene Wissensquellen, beispielsweise frühere Validierungsberichte, Modelldokumentationen und Fachliteratur[5] aus. Auf dieser Grundlage übernehmen sie auch eine (Vor-)Strukturierung des Validierungsvorgehens insgesamt.

Agentische KI beschleunigt die quantitative Validierung, indem sie beispielsweise auf Basis der bereitgestellten Modelldokumentation eine Referenzimplementierung des Bewertungsmodells unter Einbindung etablierter Bewertungsbibliotheken (z. B. QuantLib) erstellt.

Die durch die agentische KI erzeugte Referenzimplementierung in Form von Skripten wird vor ihrer Verwendung fachlich und technisch durch den verantwortlichen Validierenden geprüft bzw. abgenommen und im Anschluss durch die agentische KI eigenständig eingesetzt. Diese Skripte beinhalten auch Datenbankabfragen zur Extraktion bewertungsrelevanter Markt- und Produktdaten. Solche Abfragen werden, auf Basis einer fachlichen Beschreibung der Datenbankinhalte und der technischen Beschreibung des Datenmodells, durch agentische KI autonom erzeugt. Die Qualität der KI-generierten Datenbankabfragen hängt hierbei wesentlich von der bereitgestellten Beschreibung der Daten ab (Metadaten-Management).

Frühere Versionen großer Sprachmodelle (z. B. ChatGPT 4.1) zeigten noch vereinzelt Schwächen in der Beurteilung fachlicher Detailfragen, insbesondere im Rahmen der qualitativen Validierung. Doch selbst mit den Qualitätssteigerungen aktueller LLMs bleibt der verantwortliche Validierende als aufmerksamer Prüfer, Qualitätssicherer und finaler Abnehmer der Ergebnisse quantitativer und qualitativer Analysen unverzichtbar.

5 Aktualisierungen der regulatorischen Anforderungen beziehungsweise Fortschritte in der finanzmathematischen Modellierung können durch das agentische System grundsätzlich per Web-Suchen identifiziert werden.

Im PoC konkret umgesetzt:

- Im Rahmen des PoCs wurde das von der Ampega verwendete Bewertungsverfahren für kündbare Floater betrachtet. Dieses wird in Abwesenheit eines liquiden Marktpreises zur Bestimmung von Fair Values sowie für alle kündbaren Floater in der Marktpreisrisikorechnung eingesetzt. Da die Credit-Spread-Sensitivität der Wertpapiere von der erwarteten Restlaufzeit abhängt, spielt dieses Bewertungsmodell in der Marktpreisrisikorechnung eine wichtige Rolle
- Qualitative Validierung des Bewertungsverfahrens mittels RAG-basiertem Zugriff auf die Modelldokumentation und das modellübergreifende Validierungskonzept
- KI-gestützte Implementierung einer Challenger-Implementierung in Python, basierend auf der fachlichen Modelldokumentation in Verbindung mit einem Metadaten-Management für SQL-Datenbankabfragen
- Durchführung der quantitativen Validierung durch die agentische KI basierend auf dem als Tool zur Verfügung gestellten Challenger-Modell

04.02 Validierung von Marktrisikomodellen

Ausgangslage

In Analogie zur oben beschriebenen Validierung der Bewertungsmodelle kann auch bei der Validierung von Marktpreismodellen eine Unterscheidung zwischen qualitativen und und quantitativen Validierungshandlungen getroffen werden. Im Rahmen des PoCs wurden qualitative und quantitative Aspekte bei der Validierung von Marktrisikomodellen allerdings zumeist zeitgleich betrachtet.

Die quantitativen Aspekte umfassen hierbei vor allem die Überprüfung statistischer Voraussetzungen der gewählten Modelle, deren Kalibrierungsroutinen sowie die Analyse der Angemessenheit und Granularität ausgewählter Risikofaktorgruppen.

In der qualitativen Validierung des Marktpreisrisikos werden die Fragen der Marktüblichkeit und der regulatorischen Compliance angesiedelt. Weiterhin stellen sich auch qualitative Fragen zur Aussagekraft der durchgeführten statistischen Tests und deren Abhängigkeiten; so führt zum Beispiel eine Abweichung in den Autokorrelationsannahmen zu einer eingeschränkten Nutzbarkeit anderer Tests.

Integration agentischer KI in den Validierungsprozess

Über den PoC wurde untersucht, inwieweit agentische KI die Validierung unterstützen beziehungsweise eigenständig durchführen kann. Der Schwerpunkt lag auf den quantitativen Aspekten, aber auch der qualitativen Einordnung von Ergebnissen. Insbesondere wurde der Arbeitseinsatz in den folgenden Gebieten erprobt:

- Überprüfung der statistischen Modellparameter
- Analyse der Angemessenheit und Granularität einzelner Risikofaktorgruppen
- Auswertung der beobachteten Calibration Spreads

Im Rahmen dieser Tätigkeiten konnte agentische KI insbesondere bei klar strukturierten Arbeitsschritten wirksam unterstützen. So wurden Analyse-Skripte automatisiert ausgeführt, Datensätze strukturiert ausgelesen sowie Tabellen, Grafiken und deren Interpretationen konsistent in Berichtsstrukturen eingebunden.

Auch beim Erstellen komplexerer, erstmalig auftretender Analysen erwies sich der Einsatz agentischer KI als hilfreich. Bestehende Datenextraktionen und Analyseskripte konnten flexibel durch die KI rekombiniert und für neue Fragestellungen genutzt werden, um somit Ursachen für statistische Ausreißer zu identifizieren und einzuordnen.

Im PoC konkret umgesetzt:

- Im Rahmen der Erprobung wurde untersucht, inwieweit agentische KI auf Basis bestehender Analysen neue Analyseansätze ableiten kann. Hierfür standen verschiedene R-basierte Werkzeuge zur Durchführung statistischer Tests sowie eine Datenbasis von Risikofaktorzeitreihen zur Verfügung. Die KI konnte die relevanten Datengrundlagen gezielt adressieren und Analysen auf dieser Basis durchführen.
- Ein häufig verwendeter Ablauf im PoC war das Durchführen von Standardanalysen sowie das Interpretieren von dabei erstellten Plots, Tabellen und Kennzahlen. Insbesondere konnten über den RAG-Agenten auch Bezüge zu schon bekannten Mustern aus vergangenen Validierungen hergestellt werden.
- Im Bereich der qualitativen Validierung konnte die KI genutzt werden, um als fachlicher Sparringspartner, basierend auf dem existierenden Modell und ausgewerteten statistischen Eigenschaften bezüglich der Verteilung der Residuen, mögliche Modellverbesserungen vorzuschlagen.

Die Erprobung im Rahmen der Marktpreisrisikovalidierung hat gezeigt, dass mithilfe agentischer KI zahlreiche Standardanalysen und Interpretationen KI-basiert durchgeführt werden können. Durch den Einsatz von persistierten Prompts stellen sich hier Zeitersparnisse ein, da Analyse und Interpretation auf „Knopfdruck“ entstehen und in der nächsten Validierung wiederverwendet werden können. Besonders hervorzuheben sind hierbei statistische Tests. In diesem Bereich zeigt agentische KI großes Potenzial und eine hohe Ausführungssicherheit, auch bei der Durchführung und Interpretation mehrstufiger Analysen.

05. Zusammenfassung und Ausblick

Die wesentlichen Arbeitsschritte im Rahmen der Modellvalidierung lassen sich anhand weniger zentraler Tätigkeiten beschreiben. Dazu gehören insbesondere die Extraktion von Daten aus Datenbanken und Fachdokumenten, die Durchführung qualitativer und quantitativer Analysen (gegebenenfalls unter Einsatz von Challenger-Modellen), die Dokumentation der Ergebnisse sowie deren fachliche Einordnung im hausinternen und regulatorischen Kontext. Sofern erforderlich werden hieraus Feststellungen und Handlungsempfehlungen zur Behebung identifizierter Modellschwächen abgeleitet.

Bereits vor Durchführung des PoCs gingen die Ampega und d-fine gemeinsam davon aus, dass sich bei quantitativen Analysen durch Workflow-Automatisierung und den Einsatz agentischer KI zur Ergebnisbeschreibung erhebliche Effizienzpotenziale realisieren lassen. Diese Erwartung konnte im Rahmen des PoCs klar bestätigt werden.

Positiv überraschend war darüber hinaus, wie gut sich der agentische Ansatz in die Entwicklung neuer Analysen, Skripte und Datenauswertungen innerhalb des bestehenden Validierungsprozesses integrieren ließ. Der Mehrwert agentischer KI zeigte sich damit nicht allein in der Automatisierung bestehender Analysen. Vielmehr ermöglichte ihr Einsatz inhaltlich wertvolle zusätzliche Analysen und Deep Dives.

Abbildung 2 fasst zusammen, in welchem Umfang die genannten Tätigkeiten im PoC durch agentische KI übernommen werden konnten und skizziert zugleich die Bereiche, in denen weiteres Verbesserungspotenzial identifiziert wurde.

Die größten Stärken liegen in Talk to your Data Anwendungen, der auf Fachkonzepten basierenden Erstellung und Anbindung von Analyse-Skripten sowie der Nutzung verschiedener Informationsquellen und Werkzeuge in einem Workflow.

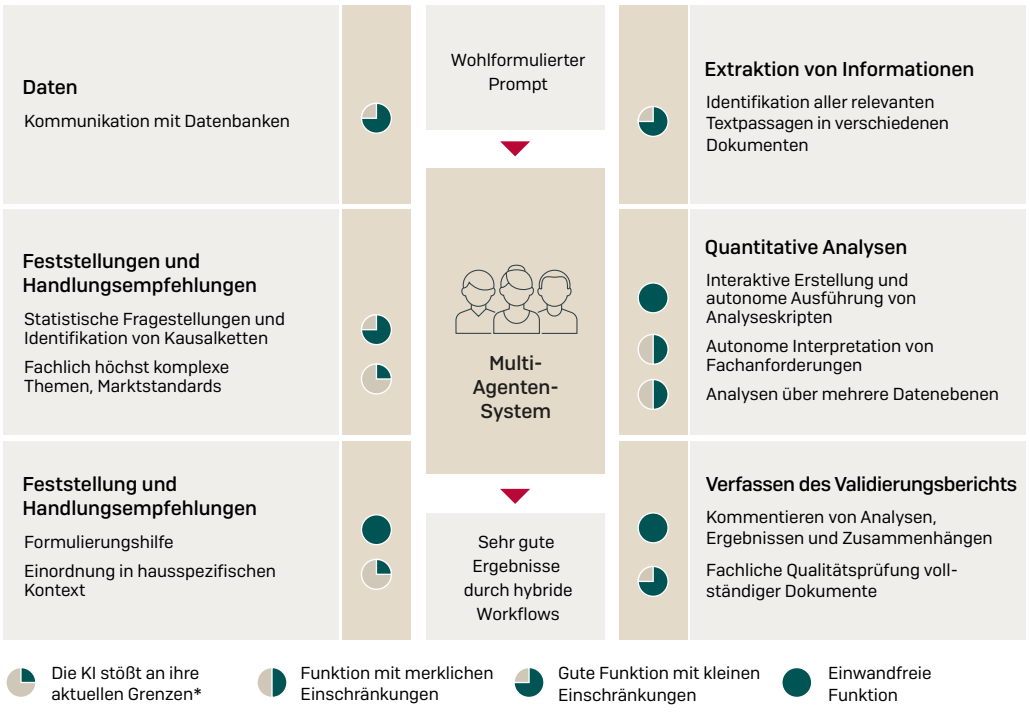


Abbildung 2: Übernahme von Tätigkeiten durch agentische KI in der Modellvalidierung

Aus unserer Sicht konnten die folgenden Validierungshandlungen im Rahmen des PoCs sehr gut automatisiert werden:

- die Analyse von Daten via „Talk To Your Data“,
- die Beantwortung spezifischer und statistischer Fragestellungen,
- die Durchführung quantitativer Analysen von (Zwischen-)Ergebnissen
- und die qualitative Einwertung und Dokumentation der Validierungsergebnisse.

Daher gelangen die Ampega und d-fine gemeinsam zu der Einschätzung, dass bei entsprechender Verzahnung der agentischen KI mit den technischen Systemen, welche die zu validierenden Modelle beziehungsweise Challenger-Modelle vorhalten, sowie der Datenhaltung, von erheblichen Potenzialen zur Effizienzsteigerung und Skalierung sowie damit einhergehend auch von wesentlichen Qualitätsgewinnen im Prozess der Modellvalidierung auszugehen ist.

Selbst im Fall lediglich unstrukturierter Daten kann agentische KI helfen, bei Bedarf automatisiert temporäre Datenbanken zu erstellen, um so schließlich über „Talk To Your Data“ in Verbindung mit einem schlanken Metadaten-Management ad hoc Analysen zu ermöglichen. Damit sind grundsätzlich auch Automatisierungen auf Basis unstrukturierter Informationen möglich, die ohne KI-Einsatz nur mit erheblichem manuellem Aufwand realisierbar wären.

In anderen Teilen des Modellvalidierungsprozesses haben d-fine und die Ampega im Rahmen des PoCs größere Herausforderungen für eine aufwandsarme Automatisierung durch agentische KI identifiziert. Dies betrifft insbesondere den Umgang mit hausspezifischen Inhalten, die autonome Interpretation komplexer fachlicher Anforderungen sowie sehr anspruchsvolle Fragestellungen in fachlichen Spezialbereichen.

Gerade bei Letzteren zeichnete sich jedoch bereits während des PoCs sowie in dessen Nachgang ein deutlicher Fortschritt ab, der auf den Einsatz zunehmend leistungsfähiger Sprachmodelle zurückzuführen ist — insbesondere im Vergleich zu GPT-4.1, das im PoC primär verwendet wurde. Darüber hinaus ermöglichen neuere Sprachmodelle durch größere Kontextfenster zunehmend, umfangreichere hausspezifische Informationen innerhalb eines Analyse- oder Validierungskontexts zu berücksichtigen.

Ein umfangreicher, produktiver Einsatz von KI im Rahmen der Modellvalidierung wird in der Praxis mit Anforderungen an die hausintern eingesetzte IT, an die verantwortlich Validierenden sowie, wie oben beschrieben, auch mit Fragen im Bereich Regulatorik und Governance einhergehen. In der Praxis ist also eine Abwägung zwischen den oben beschriebenen erheblichen Potenzialen und den mit der Hebung dieser Potenziale einhergehenden Aufwänden erforderlich.

Vor dem Hintergrund der Aufwände ist zu bemerken, dass die im PoC betrachteten KI-Tools breite Anwendungsmöglichkeiten über den eigentlichen Modellvalidierungsprozess hinaus haben und sich dazu eignen, Prozesse entlang der gesamten Wertschöpfungskette im Unternehmen signifikant zu verbessern. In diesem Zusammenhang bergen insbesondere KI-gestützte fachliche Datenqualitätsanalysen sowie die „natürlichsprachliche“ Kommunikation mit komplexen Asset-Beständen und zugehörigen Metadaten in der Praxis großes Potenzial.

Autoren/Autorinnen

ampega

Jakob Schwibbert
Quantitative Modelling Specialist

Svetlana Kell
Senior Market Risk Specialist

Dr. Gregor Wergen
Head of Pricing Methodology & ESG Controlling

d-fine

Tim Kopischke
Manager, Agentic AI / GenAI Expert

Dr. Frank Könnig
Manager, Product Manager LARA

Thorben Peters
Manager, Market Risk Modelling Expert

Ferdinand Reuther
Manager, Quantitative Modelling Expert

Dr. Christian Kappen
Senior Manager, Quantitative Modelling Expert

d-fine

ampega.