

Ergebnisbericht des Ausschusses Enterprise Risk Management

Einsatz von Künstlicher Intelligenz in der Risikomodellierung der Lebens- versicherung

Köln, 23. Februar 2026

Präambel

Die Unterarbeitsgruppe *KI in Risikomodelle Leben* der Arbeitsgruppe *Risikomodelle Lebensversicherung* des Ausschusses Enterprise Risk Management der Deutschen Aktuarvereinigung e. V. (DAV) hat den vorliegenden Ergebnisbericht erstellt.¹

Anwendungsbereich

Der Ergebnisbericht betrifft Aktuarinnen und Aktuare² welche im Bereich der Risikomodellierung in der Lebensversicherung tätig sind.

Viele der erörterten Themen lassen sich jedoch auch in andere Sparten und Tätigkeitsbereiche übertragen. Somit bietet der Bericht potenziell allen Aktuarinnen und Aktuaren einen wertvollen Erkenntnisgewinn.

Der Ergebnisbericht ist an die Mitglieder und Gremien der DAV zur Information über den Stand der Diskussion und die erzielten Erkenntnisse gerichtet und stellt keine berufsständisch legitimierte Position der DAV dar.³

Inhalt

Der Ergebnisbericht richtet sich an Aktuarinnen und Aktuare in der Risikomodellierung von Lebensversicherern und zeigt auf, in welchen Prozessen der Einsatz von KI-Technologie einen besonderen Mehrwert bieten kann. Zu Beginn wird ein Überblick über die relevanten Prozesse in der Risikomodellierung der Lebensversicherung gegeben. Anschließend erfolgt eine Einordnung des Begriffs der Künstlichen Intelligenz auf Basis des EU AI Act sowie der damit verbundenen regulatorischen Anforderungen und verschiedener Arten von KI-Systemen. Im weiteren Verlauf werden die einzelnen Prozesse detaillierter betrachtet. Dabei werden exemplarische KI-Anwendungen vorgestellt und die daraus resultierenden Chancen und Herausforderungen erläutert. Ziel des Ergebnisberichts ist es, Aktuarinnen und Aktuare anzuregen, die bestehenden Prozesse in ihren Unternehmen kritisch zu hinterfragen und Potenziale für den sinnvollen Einsatz von KI-Technologien zu identifizieren.

Schlagworte

Lebensversicherung, Risikomanagement, Risikomodelle, Künstliche Intelligenz, Machine Learning, Deep Learning

Verabschiedung

Dieser Ergebnisbericht ist durch den Ausschuss Enterprise Risk Management am 23. Februar 2026 verabschiedet worden.

¹ Der Ausschuss dankt der Unterarbeitsgruppe *KI in Risikomodelle Leben* der Arbeitsgruppe *Risikomodelle Lebensversicherung* ausdrücklich für die geleistete Arbeit, namentlich: Andreas Trusow, Dr. Aleksander Rejman, Dr. Mischa Pupashenko (Leitung), Jessica Wenzler (Gast), Matthias Sohn und Miglena Gavrilova.

² Auch wenn hier und im Folgenden die Aktuarinnen und Aktuare explizit genannt werden, spricht die DAV alle Geschlechter und Identitäten gleichermaßen an. Dies gilt auch für alle anderen hier genannten Personengruppen.

³ Die sachgemäße Anwendung des Ergebnisberichts erfordert aktuarielle Fachkenntnisse. Dieser Ergebnisbericht stellt deshalb keinen Ersatz für entsprechende professionelle aktuarielle Dienstleistungen dar. Aktuarielle Entscheidungen mit Auswirkungen auf persönliche Vorsorge und Absicherung, Kapitalanlage oder geschäftliche Aktivitäten sollten ausschließlich auf Basis der Beurteilung durch eine(n) qualifizierte(n) Aktuar DAV/Aktuarin DAV getroffen werden.

This abstract summarises the report on findings „Einsatz von Künstlicher Intelligenz in der Risikomodellierung der Lebensversicherung“ which was approved by the DAV committee Enterprise Risk Management on 23.02.2026.

Abstract

The report is aimed at actuaries working in risk modelling for life insurers and highlights the processes in which the use of AI technology can provide added value. It begins with an overview of the key processes involved in risk modelling within life insurance. This is followed by a discussion of the concept of artificial intelligence based on the EU AI Act, including the related regulatory requirements and different types of AI systems. Subsequently, the individual processes are examined in greater detail. Exemplary AI applications are presented, and the resulting opportunities and challenges are discussed. The objective of the report is to encourage actuaries to critically review existing processes within their organizations and to identify potential areas for the meaningful application of AI technologies.

1. Einleitung	6
1.1. Übersicht der Prozesse in der Risikomodellierung in der Lebensversicherung	6
1.2. Definition von Künstlicher Intelligenz	7
1.3. Arten von Data-Science-Techniken und KI	7
1.4. Regulatorische Rahmenbedingungen	10
1.5. Allgemein Nutzen und Herausforderungen	10
1.5.1. Nutzen	10
1.5.2. Herausforderungen	11
2. Datenaufbereitung und Qualitätssicherung	12
2.1. Einleitung	12
2.2. Qualitätssicherung der Eingabedaten – bisherige Erfahrungen	12
2.3. Qualitätssicherung der Eingabedaten mit KI-Systemen	13
3. Herleitung der biometrischen Annahmen	16
3.1. Einleitung	16
3.2. Festlegung der biometrischen Annahmen – bisherige Erfahrungen	16
3.3. (Automatisierte) Herleitung der Annahmen durch KI	17
4. Modellierung	18
4.1. Einleitung	18
4.2. Modellierung – bisherige Erfahrungen	18
4.3. Entwicklung von aktuarieller Software mit KI	19
4.4. Beispiel Show-Case – Modellierung / Testen eines neuen Produkts	21
5. Berechnungen	22
5.1. Einleitung	22
5.2. Berechnungen – bisherige Erfahrungen	22
5.3. Automatisierung durch KI	23
5.4. KI-proxy bzw. Ersatz inkl. Fortschreibung / Prognose von SCR	23
6. Validierung	27
6.1. Einleitung	27
6.2. Validierung – bisherige Erfahrungen	27
6.3. Validierung mit KI	29
7. Dokumentation sowie Berichtswesen	32
7.1. Einleitung	32
7.2. Dokumentation und Berichtswesen – bisherige Erfahrungen	32
7.3. Erstellung von Dokumentationsentwürfen durch KI	32

7.4.	Automatisiertes Aktualisieren von Dokumentationen und Berichten	32
7.5.	Automatisierte und KI-gestützte Berichtserstattung unter Solvency II.....	33
8.	Fazit und Ausblick.....	34
Anhang	36
8.1.	Abkürzungen	36
8.2.	Literaturverzeichnis	39

1. Einleitung

Die Risikomodellierung in der Lebensversicherung ist geprägt durch eine hohe fachliche, methodische und technische Komplexität. Insbesondere im Rahmen von Solvency II erfordert sie eine enge Verzahnung stochastischer Cashflow Modelle, granularer versicherungstechnischer Annahmen sowie umfangreicher Datenaufbereitungs- und Validierungsprozesse. Die Erstellung von QRTs, internen und aufsichtsrechtlichen Berichten, sowie die Berechnung des SCR unter Einbezug verschiedener Risikomodule erfordern signifikante Rechnerkapazität sowie Ressourcen zur Sicherstellung von Modellgovernance und Nachvollziehbarkeit. Viele dieser Prozesse, von der Datenaufbereitung bis zur Ergebnisaggregation, finden auch im Kontext von IFRS 17 Anwendung.

Angesichts steigender regulatorischer Anforderungen, wachsender Datenmengen und zunehmender Rechenintensität wächst der Bedarf an Optimierung, Automatisierung und Beschleunigung innerhalb dieser Wertschöpfungskette. Gleichzeitig steigen die Ansprüche an Transparenz, Konsistenz und Qualität der Ergebnisse. Dies öffnet den Raum für den Einsatz neuartiger Technologien, insbesondere von Künstlicher Intelligenz (KI).

Die KI gehört zu den einflussreichsten Technologien unserer Zeit und beeinflusst zunehmend zahlreiche Branchen, darunter auch die Versicherungswirtschaft. Der Einsatz von KI eröffnet neue Möglichkeiten zur Effizienzsteigerung, Verbesserung der Datenqualität und Optimierung komplexer Prozesse. Insbesondere in datengetriebenen Bereichen können KI-Technologien vielfältige Potenziale entfalten.

1.1. Übersicht der Prozesse in der Risikomodellierung in der Lebensversicherung

Dieser Abschnitt gibt einen strukturierten Überblick über die wesentlichen Prozessschritte der Risikomodellierung in der Lebensversicherung – von der Datenaufbereitung über die aktuariellen Bewertungsmodelle bis zur Ableitung aufsichtsrechtlicher Kennzahlen und Berichte. Die Darstellung erhebt keinen Anspruch auf Vollständigkeit, sondern soll einen generischen, und strukturierten Überblick über typische Prozessschritte in der Risikomodellierung geben.

1. Datenaufbereitung und Qualitätssicherung
 - (Rück-)Versicherungsbestand inkl. Verdichtung für stochastische Berechnungen
 - Kapitalanlagen, Marktdaten und Zinsszenarien
 - Rechnungsgrundlagen (z. B. Sterblichkeit, Invalidität, Langlebigkeit)
 - Kosten inkl. Inflation
 - Weitere Eingabedaten (z. B. Eröffnungsbilanz, Parametrisierung der Management Regeln)
2. Herleitung der Annahmen
 - Versicherungstechnische Annahmen
 - Rechnungsgrundlagen (z. B. Sterblichkeit, Invalidität, Langlebigkeit)
 - Versicherungsnehmerverhalten (z. B. Storno, Betragsfreistellung, Kapitalwahl, Dynamisches VN-Verhalten)
 - Kosten inkl. Inflation
 - Ökonomische Annahmen (Zinskurve, Kapitalmarktszenarien, Inflation)
 - Management Regeln inkl. Kapitalanlagestrategie und Gewinnverteilung
3. Modellierung des Asset-Liability Managements (ALM)
 - Entwicklung / Nutzung aktuarieller Software (z. B. Prophet, RiskAgility FM, PLA.NET)
 - Abbildung der Produkte (inkl. VN-Verhalten, Optionen und Garantien (O&G)) und Kapitalanlagen (inkl. Kapitalanlagestrategie)
 - Cashflow (CF) Projektionen inkl. Methoden und regulatorischer Anforderungen (z. B. Mindestzuführungsverordnung (MindZV)), Rückstellung für Beitragsrückerstattung (RfB))
 - Stochastische Modellierung (z. B. O&G-Bewertung)
 - Risikomodellierung (z. B. Standardstresse, SCR-Aggregation, Replikation und LSMC)

- Datenmodell- und Modell-Governance (Change-Management, Interne Modelle, Dokumentation)
- 4. Berechnungen (z. B. Quartal- und Jahresabschlüsse)
 - Konfiguration der Berechnungsläufe inkl. Sensitivitäten und ad-hoc Berechnungen
 - Ergebnisaufbereitung (z. B. QRTs, internes Reporting, KPIs und KRIs)
 - Analysen und Erklärung der Ergebnisse (z. B. Veränderungsanalyse)
 - Proxy-Modelle, Forecasts und Planungsrechnungen
- 5. (Unabhängige) Validierung
 - Datenqualität (siehe Punkt 1) und Abdeckungsgrad
 - Angemessenheit der Annahmen (siehe Punkt 2)
 - Methoden inkl. Model Changes (siehe Punkt 3)
 - Vereinfachungen
 - Ergebnisse (siehe Punkt 4)
 - Stresstests und Szenariorechnungen
 - Dokumentations- sowie Kontrollstandards
 - Erstellen des Validierung-Reports (siehe Punkt 6)
- 6. Dokumentation sowie Berichtswesen
 - Solvency II Berichte
 - Internes und aufsichtliches Reporting
 - Dokumentation (z. B. Herleitung der Annahmen, Modellierung)

1.2. Definition von Künstlicher Intelligenz

Gemäß dem EU AI Act wird ein KI-System wie folgt definiert:

"Ein KI-System ist ein maschinengestütztes System, das für einen in unterschiedlichem Grade autonomen Betrieb ausgelegt ist und das nach seiner Betriebsaufnahme anpassungsfähig sein kann und aus den erhaltenen Eingaben für explizite oder implizite Ziele ableitet, wie Ausgaben wie etwa Vorhersagen, Inhalte, Empfehlungen oder Entscheidungen erstellt werden, die physische oder virtuelle Umgebungen beeinflussen können.

(Rechtsgrundlage: Art. 3(1), Annex I AI Act)⁴

Allgemein versteht man unter Künstlicher Intelligenz Technologien, die in der Lage sind, Aufgaben zu erfüllen, die typischerweise menschliche Intelligenz erfordern. Dazu gehören das Lernen aus Daten, das Erkennen von Mustern, das Treffen von Entscheidungen sowie die Anpassung an neue Informationen.

1.3. Arten von Data-Science-Techniken und KI

Künstliche Intelligenz (KI) umfasst Verfahren, die Aufgaben ausführen, die typischerweise menschliche kognitive Fähigkeiten erfordern (z. B. Erkennen von Mustern, Lernen aus Daten, Ableiten von Entscheidungen). In der Praxis basiert ein Großteil aktueller KI-Anwendungen auf maschinellem Lernen. Im Folgenden werden unter „klassischen Machine-Learning-Verfahren“ vor allem etablierte, überwiegend nicht-generative Methoden verstanden, die auf strukturierten Daten Prognosen, Klassifikationen oder Mustererkennung ermöglichen (z. B. lineare Modelle, Entscheidungsbaum- und Ensembleverfahren). Deep-Learning-Ansätze werden davon abgegrenzt und separat behandelt.

⁴ Verordnung (EU) 2024/1689 des Europäischen Parlaments und des Rates vom 13. Juni 2024

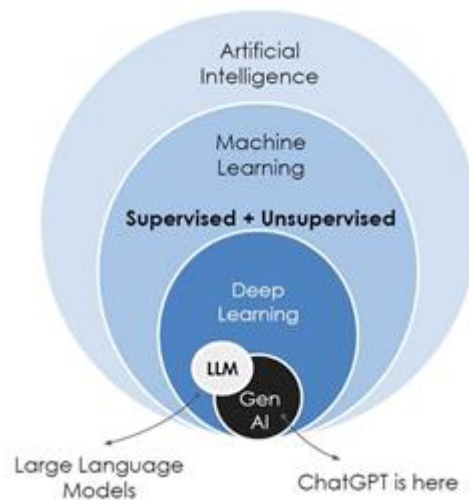


Abbildung 1: Schematische Einordnung von AI, Machine Learning, Deep Learning und Generativer KI.

Maschinelles Lernen (ML)

Maschinelles Lernen (ML) bezeichnet Verfahren, bei denen Modelle aus Daten Zusammenhänge ableiten, um Vorhersagen zu treffen oder Strukturen in Daten zu erkennen. Üblicherweise wird ML in überwachtes Lernen und unüberwachtes Lernen unterteilt. Während beim überwachten Lernen ein Zielwert (Label) vorgegeben ist, werden beim unüberwachten Lernen Strukturen und Muster ohne vorgegebene Zielvariable identifiziert.

- **Überwachtes Lernen (Regression/Klassifikation)**

- Random Forest (RF): ein Verfahren, das viele Entscheidungsbäume kombiniert, um robuste Vorhersagen zu treffen. Es eignet sich besonders gut, wenn Zusammenhänge in den Daten komplex oder nicht linear sind, z. B. zur Schadenprognose.
- Isolation Forest (IF): ein speziell für die Anomalie Erkennung entwickelter Machine-Learning-Algorithmus. Das Verfahren detektiert Anomalien in einem Datensatz durch die sukzessive Isolierung mithilfe einer Vielzahl von Isolation Trees. Anomalien werden anhand ihrer kürzeren durchschnittlichen Pfadlängen innerhalb dieser Bäume identifiziert, was darauf hinweist, dass sie leichter vom restlichen Datenbestand zu trennen sind.
- Gradient Boosting Machines (GBM): eine Modellfamilie, bei der Entscheidungsbäume schrittweise (sequenziell) aufgebaut werden. Jeder nachfolgende Baum dient dazu, die Fehler der vorherigen Bäume systematisch zu korrigieren. Dadurch entstehen besonders leistungsfähige Vorhersagemodelle, z. B. für Storno- oder Mortalitätsrisiken.
- Support Vector Machines (SVM): ein klassisches Verfahren zur Klassifikation, das besonders effektiv ist, wenn die Gruppen in den Daten durch klare Trennlinien voneinander unterscheidbar sind. SVM-Modelle eignen sich z. B. gut zur Einteilung von Kunden oder Risiken in unterschiedliche Kategorien, etwa zur Abgrenzung zwischen hohen und niedrigen Stornorisiken.

- **Unüberwachtes Lernen (Clustering & Dimensionsreduktion)**

- K-Means Clustering: ein Verfahren, das Datensätze automatisch in Gruppen (Cluster) unterteilt, basierend auf der Ähnlichkeit bestimmter Merkmale. Es kann beispielsweise dazu verwendet werden, Kundensegmente mit ähnlichem Verhalten zu identifizieren – etwa zur zielgerichteten Tarifgestaltung oder zur Erkennung auffälliger Muster bei Betrugsprüfungen.
- Density-Based Spatial Clustering of Applications with Noise (DBSCAN): ein nicht-parametrischer, dichtebasierter Clustering-Algorithmus bei dem gegeben eine Menge von Punkten in einem Raum, diejenigen Punkte gruppiert werden, die dicht beieinander liegen (also viele nahe Nachbarn haben), jene Punkte, die isoliert in

- Regionen mit geringer Dichte liegen (deren nächste Nachbarn also zu weit entfernt sind), werden als Ausreißer markiert.
- Principal Component Analysis (PCA): ein Verfahren zur Reduktion der Anzahl von Merkmalen in einem Datensatz. Dabei werden stark korrelierte Variablen zusammengefasst und durch sogenannte Hauptkomponenten ersetzt. Dies erleichtert die Modellierung, verringert Rechenaufwand und kann helfen, Overfitting zu vermeiden.

Neuronale Netze (NN) und Deep Learning (DL)

Neuronale Netze und Deep Learning sind Teilbereiche des maschinellen Lernens. Deep Learning bezeichnet besonders tiefe neuronale Netze.

- Neuronale Netze (NN): Modelle, die aus mehreren Schichten (Layern) bestehen und darauf ausgelegt sind, komplexe Muster und Zusammenhänge in großen Datensätzen zu erkennen. Sie eignen sich gut für strukturierte, aber vielschichtige Daten, z. B. zur Analyse von Leistungsfällen (Claims) oder Gesundheitsangaben bei der Risikoprüfung.
- Deep Learning (DL): eine spezielle Form neuronaler Netze mit besonders vielen Schichten. Sie ermöglichen es, relevante Merkmale automatisch aus den Daten zu extrahieren – ohne dass vorher manuell Features definiert werden müssen. DL ist vor allem bei unstrukturierten Daten wie Texten, Bildern oder Audiodaten nützlich, etwa bei der Verarbeitung von Freitexten in Leistungsanträgen.
- Rekurrente Neuronale Netze (RNN) / Long Short-Term Memory (LSTM): Varianten neuronaler Netze, die speziell für sequenzielle Daten entwickelt wurden. Sie sind in der Lage, Informationen über vergangene Werte zu berücksichtigen und eignen sich daher besonders für die Analyse von Zeitreihen, z. B. zur Vorhersage von Beitragsentwicklungen, Stornoverhalten oder medizinischen Parametern über die Zeit.
- Autoencoder: ein Autoencoder ist eine Art künstliches neuronales Netzwerk, das zum Erlernen effizienter Datenkodierungen auf unbeaufsichtigte Weise verwendet wird. Das Modell optimiert zwei Funktionen: eine Kodierungsfunktion zur Kompression der Eingabedaten und eine Dekodierungsfunktion zur Rekonstruktion der ursprünglichen Eingaben aus der komprimierten Repräsentation. Autoencoder werden hauptsächlich für Aufgaben wie Dimensionsreduktion, Datenkompression, Anomalie Erkennung und Merkmalslernen eingesetzt.

Generative KI (GenAI)

Im Folgenden wird ein Überblick über die Klasse der Generative KI (GenAI) Systeme gegeben. Unter diesem Oberbegriff werden KI-Verfahren verstanden, die neue Inhalte erzeugen. Eine Teilmenge davon sind Vortrainierten Generative AI (GenAI) Modell. Darunter werden KI-Systeme verstanden, die eigenständig neue Inhalte erzeugen, z. B. Texte, Bilder, Code oder Simulationen. Grundlage dieser Systeme sind komplexe neuronale Netze. In der Versicherungswirtschaft kann GenAI z. B. zur automatisierten Erstellung von Berichten eingesetzt werden. Große Sprachmodelle (LLMs) sind die bekannteste Form generativer KI im Textbereich.

- Natural Language Processing (NLP): ein Teilgebiet der KI, das sich mit der Verarbeitung, Analyse und Interpretation natürlicher Sprache befasst. Es bildet die methodische Grundlage vieler textbasierter Automatisierungen. In der Versicherungswirtschaft kann NLP dazu eingesetzt werden, Verträge automatisch zu prüfen.
- Large Language Models (LLMs): eine moderne Klasse von NLP-Modellen, die auf der sogenannten Transformer-Architektur basieren, z. B. GPT, BERT oder LLaMA. Diese Modelle sind in der Lage, umfangreiche Texte zu verstehen, sinnvoll zu vervollständigen oder eigenständig zu generieren. Sie kommen etwa bei Chatbots oder FAQ-Antwortsystemen zum Einsatz.
- KI-Agenten: aufbauend auf LLMs ermöglichen KI-Agenten die autonome Ausführung komplexer Aufgaben. Im Unterschied zu reinen LLMs kombinieren Agenten Sprachverarbeitung mit Zugriff auf externe Tools, Datenbanken oder Regeln. Sie können dadurch nicht nur

Texte generieren, sondern auch aktiv Informationen abrufen, Berechnungen durchführen oder Prozesse anstoßen.

- Retrieval-Augmented Generation (RAG): ein Ansatz, bei dem ein LLM mit einer externen Wissensquelle kombiniert wird. Während das LLM für die Sprachverarbeitung und Antwortgenerierung zuständig ist, liefert der vorgeschaltete „Retriever“ relevante Informationen aus einer strukturierten oder unstrukturierten Datenbasis. So kann ein RAG-Modell fundierte, aktuelle und kontextspezifische Antworten geben.
- Retrieval-Augmented Generation with Graphs (Graph-RAG): eine Weiterentwicklung des klassischen RAG-Konzepts, bei der das externe Wissen zusätzlich in Form eines semantischen Graphen organisiert wird. Dieser Graph bildet Beziehungen explizit ab. Dadurch lassen sich komplexe Zusammenhänge gezielter analysieren und erklären. Graph-RAG-Modelle verbessern auch die Nachvollziehbarkeit und Konsistenz der Antworten durch eine visuell und logisch nachvollziehbare Struktur.
- Reasoning-Modelle (RM): ein Ansatz, bei dem das LLM gezielt dazu gebracht wird, tatsächlich „nachzudenken“ und sich selbst zu hinterfragen. Aufgaben werden in Teilschritte zerlegt, Hypothesen/Alternativen gegengerechnet und Zwischenergebnisse überprüft. Bei Bedarf nutzt das Modell Tools und Regeln, um Aussagen zu verifizieren. Ergebnis sind begründete, auditierbare Antworten mit geringerer Halluzinationsrate – hilfreich für Datenqualitätsprüfungen, Annahmenherleitung und Validierung.

Visualisierung

- t-Distributed Stochastic Neighbour Embedding (t-SNE): wird zur Dimensionsreduktion eingesetzt, um hochdimensionale Daten in zwei- oder dreidimensionalen Raum zu projizieren. Dabei bleibt die lokale Nachbarschaftsstruktur weitgehend erhalten, sodass Muster, Cluster und Ähnlichkeiten in komplexen Datensätzen visuell erkennbar werden. Besonders häufig findet es Anwendung bei genetischen Daten, Bildern oder Textrepräsentationen.
- Uniform Manifold Approximation and Projection (UMAP): Technik zur Dimensionsreduktion, die ähnlich wie t-SNE für Visualisierungen verwendet werden kann, aber auch für allgemeine nichtlineare Dimensionsreduktion geeignet ist.

1.4. Regulatorische Rahmenbedingungen

Neben den technologischen Entwicklungen spielt die Einhaltung regulatorischer Vorgaben eine zentrale Rolle beim Einsatz von KI. Der EU AI Act definiert klare rechtliche Rahmenbedingungen, die den sicheren und verantwortungsvollen Einsatz von KI gewährleisten sollen. Auch wenn detaillierte Regularien nicht im Fokus dieses Papiers stehen, ist deren Berücksichtigung unerlässlich.

Diese Entwicklungen verdeutlichen, dass KI nicht nur ein technischer Fortschritt, sondern ein strategischer Erfolgsfaktor für die Versicherungswirtschaft ist. Im weiteren Verlauf dieses Papiers wird vertiefend beschrieben, wie KI gezielt eingesetzt werden kann, um Prozesse zu verbessern und Wettbewerbsvorteile zu erzielen.

Eine Einschätzung der Auswirkungen des EU AI Act auf die Tätigkeit von Aktuarinnen und Aktuaren findet sich im Ergebnisbericht der DAV „Der Artificial Intelligence Act im aktuariellen Kontext“⁵.

1.5. Allgemein Nutzen und Herausforderungen

1.5.1. Nutzen

- Punktuelle Effizienzsteigerung: KI-basierte Tools ermöglichen die Automatisierung auch kleiner, aber repetitiver Aufgaben entlang der Wertschöpfungsketten. Dadurch wird aktuarielle Kapazität für höherwertige Analysen frei und die Durchlaufzeit komplexer Prozesse reduziert.

⁵ DAV, Ergebnisbericht des Ausschusses Actuarial Data Science „Der Artificial Intelligence Act im aktuariellen Kontext“ vom 11. März 2025

- Unterstützung und Beschleunigung von Analysen: Generative KI kann bei der Interpretation von Ergebnissen, der Herleitung von Annahmen oder dem Vergleich historischer Trends unterstützen.
- Steigerung der Qualität: Durch den Einsatz von KI-basierten Kontrollen lassen sich zusätzliche Validierungsschritte automatisieren, Inkonsistenzen frühzeitig erkennen und die Nachvollziehbarkeit erhöhen.
- Bessere Nutzung von unstrukturierten Daten: KI kann auch großen Mengen an Informationen aus Berichten, regulatorischen Texten oder interner Dokumentation effizient verarbeiten und gezielt nutzbar machen.
- Skalierbarkeit und Wiederverwendbarkeit: Einmal etablierte KI-Lösungen lassen sich mit geringem Zusatzaufwand auf weitere ähnliche Prozesse in den anderen Wertschöpfungsketten übertragen.

1.5.2. Herausforderungen

- Datenschutz und Vertraulichkeit bei unternehmensinternen LLMs: Während öffentlich verfügbarer Modelle wie z. B. ChatGPT große Funktionalität bieten, sind unternehmensinterne Anwendungen oft durch Einschränkungen in Datenverfügbarkeit, Infrastruktur und Governance limitiert. Der Schutz sensibler Kunden- und Vertragsdaten sowie von Unternehmensgeheimnissen (Intellectual Property) muss gewährleistet sein, was die Anwendung zum Teil einschränkt.
- Governance: Die Integration von KI in aktuarielle Prozesse erfordert ein klar definiertes Governance-Rahmenwerk. In einer sich rasant entwickelnden Welt ist davon auszugehen, dass auch die regulatorischen Anforderungen weiter steigen. Es ist daher mit einem erhöhten Umsetzungsaufwand und möglicherweise wiederholten Anpassungen zu rechnen.
- Wartung und Pflege: KI-gestützte Systeme erfordern eine kontinuierliche Wartung und Anpassung, etwa bei Änderungen fachlicher Prozesse. Ohne dedizierte Ressourcen und ausreichende Expertise besteht das Risiko von Veralterung oder Fehlanwendungen.
- Fehlendes Know-how und Kompetenzaufbau: Die Verknüpfung aktuariellen Fachwissens mit KI-Methodik erfordert interdisziplinäre Kompetenzprofile, die in vielen Unternehmen noch im Aufbau sind.
- Transparenz und Erklärbarkeit⁶: Da viele KI-Modelle als sogenannte Black-Box-Verfahren agieren, stellt sich die Frage der Nachvollziehbarkeit und Erklärbarkeit der Ergebnisse. Erklärbarkeit bezeichnet hierbei kein einheitliches Verfahren, sondern eine Methodensammlung zur (lokalen oder globalen) Interpretation von Modellverhalten und Einzelfallentscheidungen. Das ist für Versicherungsunternehmen besonders herausfordernd, da die strengen aufsichtsrechtlichen Vorgaben (z. B. durch BaFin, EIOPA oder Solvency II) erfordern, dass eine transparente, nachvollziehbare und prüfbare Entscheidungsfindung vorliegt.
- Vollständigkeit und Überwachung der Ergebnisse: KI-Modelle benötigen laufend Kontrollen hinsichtlich ihrer Prognosequalität und möglicher Fehlklassifikationen⁷ (z. B. infolge von Bias in den Trainingsdaten). Insbesondere im regulatorischen Umfeld dürfen Entscheidungen nicht allein auf schwer überprüfbaren KI-Ergebnissen beruhen.

⁶ DAV, Ergebnisbericht des Ausschusses Actuarial Data Science „Explainable Artificial Intelligence: Ein aktueller Überblick für Aktuarinnen und Aktuare“ vom 27. Mai 2024.

⁷ DAV, Ergebnisbericht des Ausschusses Actuarial Data Science „Vermeidung von Bias und Diskriminierung im Bereich Data Science“ vom 11. März 2025.

2. Datenaufbereitung und Qualitätssicherung

2.1. Einleitung

Die Datenqualität ist ein entscheidender Faktor für die Genauigkeit und Zuverlässigkeit von Modellen. Aufgrund der hohen Komplexität und Menge der Daten wird bereits heute ein erheblicher Aufwand betrieben, um die Datenqualität sicherzustellen. Dies umfasst eine Vielzahl manueller und (teil-)automatisierter Verfahren sowie interne Kontrollsysteme (IKS), die darauf abzielen Daten zu generieren und fehlerhafte oder inkonsistente Daten zu erkennen und zu korrigieren. Dieses Kapitel konzentriert sich auf die gezielte Überprüfung der formalen und inhaltlichen Plausibilität von Eingabedaten im Hinblick auf deren Eignung und Richtigkeit für den jeweiligen Anwendungszweck vor deren Nutzung. Eine unabhängige oder nachgelagerte Validierung ist Gegenstand von Kapitel 6.

2.2. Qualitätssicherung der Eingabedaten – bisherige Erfahrungen

Nachfolgend wird ein Überblick über die Eingabedaten in der Risikomodellierung der Lebensversicherung gegeben. Für eine Qualitätssicherung sind gezielte Prüfungen notwendig, welche wir im Folgenden kurz darstellen.

(Rück-)Versicherungsbestand – Aufbereitung

- Vollständigkeit der Eingangsdaten: Sicherstellung, dass alle relevanten Attribute (z. B. Vertragsstatus, Altersgruppe, Tarifnummer, Rückstellungskennzeichen) vorhanden sind.
- Konsistenzprüfung: Vermeidung widersprüchlicher Werte (z. B. negative Versicherungssummen, ungültige Vertragsnummern).
- Prüfung der Filterkriterien: Sicherstellung, dass relevante Verträge eingeschlossen und irrelevante ausgeschlossen sind.
- Analyse von Ausreißern: Identifikation und Plausibilitätsprüfung extremer Werte (z. B. ungewöhnlich hohe Rückstellungen).

(Rück-)Versicherungsbestand – Verdichtung

- Konsistenz der Gruppierungskriterien: Angemessenheit der angewendeten Kriterien (z. B. Tarifstruktur, Altersklassen, Rückstellungskennzeichen).
- Skalierung der Werte: adäquate Abbildung der Originaldaten durch verdichtete Werte (z. B. Beiträge, Rückstellungen).
- Konsistenzprüfung der Daten: Abwesenheit logischen Widersprüche innerhalb verdichteter Daten.
- Berücksichtigung von Spezialfällen: separate Prüfung seltener Tarife oder Sonderverträge, um Verzerrungen zu vermeiden.

Kapitalanlagen

- Klassifizierung von Kapitalanlagen: Sicherstellung, dass Assets korrekt den Anlageklassen (z. B. Anleihen, Aktien, Immobilien, Derivate) zugeordnet sind.
- Bewertung der Vermögenswerte: Prüfung der Marktpreise auf Aktualität und marktkonforme Kalibrierung.
- Ertrags- und Risikomodellierung: Annahmen zu Renditen, Volatilitäten und Risikoaufschläge sind mit historischen Daten und Benchmarks abzugleichen.
- Liquiditätsannahmen: Prüfung der Verfügbarkeit und Veräußerbarkeit von Assets im Modell, insbesondere in Stresssituationen.
- Aggregation zu Portfolios: Kontrolle der korrekten Aggregation von Einzelwerten, inkl. Gewichtungen, Diversifikationseffekten und Risikoverteilungen.

Zinsszenarien

- Überprüfung der Kalibrierung: Die Kalibrierung des Ökonomischen Szenariengenerators (ESG) muss auf aktuellen und belastbaren Marktdaten basieren.

- Überprüfung der Modellannahmen: Die Annahmen zu wirtschaftlichen Zusammenhängen (z. B. Zinsentwicklung, Inflationsraten) müssen realistisch und konsistent mit historischen Daten und Marktprognosen sein.
- Backtesting und Benchmarking: Durch Backtesting kann überprüft werden, ob das ESG-Modell vergangene Marktbewegungen korrekt abbilden konnte.
- Sicherstellung der Arbitragefreiheit: Das Modell muss arbitragefrei sein, um eine wirtschaftlich konsistente Bewertung zu gewährleisten.
- Prüfung der Abhängigkeitsstrukturen: Die Modellierung von Abhängigkeiten zwischen wirtschaftlichen Variablen sollte logisch und nachvollziehbar sein, insbesondere in Stresssituationen.
- Sensitivitäts- und Stresstests: Sensitivitätsanalysen prüfen, wie empfindlich das Modell auf Änderungen einzelner Parameter reagiert.

Weitere Eingabedaten

- Parametrisierung: Sicherstellung der Konsistenz und Richtigkeit der Parametrisierungen (z. B. Management-Regeln) sowohl innerhalb der Inputs als auch im Vergleich zu historischen Daten und bisherigen Einstellungen.
- Startwerte: Plausibilisierung der Startwerte im Vergleich zu der Vorperiode. Abweichungen sind unter Berücksichtigung definierter Toleranzschwellen zu prüfen, um unerwartete Veränderungen frühzeitig zu erkennen.
- Rechnungsgrundlagen: Überprüfung der Verwendung korrekter Rechnungsgrundlagen (z. B. Sterblichkeit, Invalidität, Kostenansätze). Sicherstellung der Konsistenz der Kombination verschiedener Rechnungsgrundlagen untereinander sowie in Bezug auf den Bestand (z. B. identische Stichtagsbezüge).

2.3. Qualitätssicherung der Eingabedaten mit KI-Systemen

Die DAV hat sich bereits intensiv mit Fragen der Regulierung und Validierung von KI-Modellen auseinandergesetzt. Einen vertiefenden Überblick hierzu bietet der Ergebnisbericht „Regulierung und Validierung von KI-Modellen“⁸, in welchem zentrale Fragestellungen und Lösungsansätze umfassend dargestellt werden.

Der Einsatz von KI-gestützten Methoden wie RAG, Graph-RAG und KI-Agenten kann die Qualität und Konsistenz von Inputs erheblich verbessern. Diese Technologien bieten spezifische Vorteile für unterschiedliche Eingabe-Kategorien, indem sie komplexe Datenbeziehungen analysieren, inkonsistente oder fehlerhafte Daten identifizieren und dynamisch auf Abweichungen reagieren. Im Folgenden werden die vielseitigen Einsatzmöglichkeiten dieser Ansätze dargestellt.

Kapitalanlagen

- Automatisierte Klassifizierung und Kategorisierung von Assets: Ein KI-Agent extrahiert Asset-Daten und stellt sie standardisiert bereit. Graph-RAG überprüft die Klassifizierung anhand interner Vorgaben, analysiert Asset-Beziehungen und erkennt Fehlklassifikationen wie beispielsweise fälschliche Einstufungen von Anleihen als Aktien.
- Überprüfung der Bewertungsmethoden und Marktpreise: Ein KI-Agent ruft automatisch aktuelle Marktdaten aus internen und externen Quellen ab und stellt sie für die Bewertung bereit. RAG durchsucht Dokumentationen und Benchmark-Datenbanken, um sicherzustellen, dass Bewertungsmethoden wie Marktwert oder Buchwert korrekt und konsistent angewendet wurden. Zudem gleicht es die abgerufenen Marktpreise mit den im Asset Model verwendeten Werten ab, um veraltete oder unplausible Bewertungen zu identifizieren. Der KI-Agent markiert fehlerhafte Werte, schlägt Anpassungen vor und aktualisiert die Bewertungslogik in den Systemen.

⁸ DAV, Ergebnisbericht des Ausschusses Enterprise Risk Management „Regulierung und Validierung von KI-Modellen“ vom 16. November 2018

- Prüfung der Ertrags- und Risikomodellierung: Graph-RAG visualisiert die Beziehungen zwischen Vermögenswerten, Marktrisiken und Erträgen, um un plausible Risikozuweisungen oder Annahmen zu erkennen.
- Plausibilisierung von Liquiditätsannahmen: Graph-RAG modelliert die Liquiditätsstrukturen im Portfolio, analysiert Abhängigkeiten und identifiziert Schwachstellen bei der Verfügbarkeit von Assets in Stresssituationen. Zudem erkennt es Unstimmigkeiten in den Liquiditätsannahmen einzelner Vermögenswerte.
- Überprüfung der Aggregation von Einzelassets zu Portfolios: Graph-RAG analysiert die Abhängigkeiten zwischen verschiedenen Asset-Klassen, visualisiert deren Beziehungen und überprüft, ob Aggregationsregeln korrekt angewendet wurden. So werden Unstimmigkeiten in Gewichtungen, Diversifikationseffekten und Risikoverteilungen erkannt und fehlerhafte Portfoliozusammensetzungen aufgedeckt.

(Rück-)Versicherungsbestand – Aufbereitung

- Automatisierte Prüfung der Filterkriterien und Datenvollständigkeit: Ein KI-Agent extrahiert und verarbeitet die relevanten Modelpoint-Daten aus Unternehmenssystemen und stellt sie standardisiert zur Verfügung. RAG gleicht die angewendeten Filterkriterien mit hinterlegten Vorgaben ab und identifiziert fehlende oder unvollständige Attribute wie Tarifnummern oder Altersangaben. Der Agent übernimmt zudem die automatische Korrektur fehlerhafter Filterlogiken und stößt bei kritischen Abweichungen eine manuelle Überprüfung an.
- Erkennung von Ausreißern und Inkonsistenzen: Graph-RAG erstellt einen Graphen aus den Daten, in dem Verträge, Tarife und Rückstellungen als Knoten und deren Beziehungen als Kanten dargestellt werden. Durch Analyse dieser Graph-Struktur erkennt das System un plausible Verbindungen oder fehlende Verknüpfungen.
- Dynamische Anpassung an neue Anforderungen: RAG überwacht interne und externe Quellen auf neue regulatorische oder prozessuale Änderungen und aktualisiert Prüfregeln automatisch. Ein KI-Agent implementiert die Anpassungen in den relevanten Systemen, validiert deren Umsetzung und dokumentiert die Änderungen für eine transparente Nachverfolgbarkeit.
- Automatisierte Verarbeitung von Fachkonzepten: RAG liest Produktbeschreibungen und Fachkonzepte aus, extrahiert relevante Parameter (z. B. Altersgruppen, Tarifmerkmale) und wandelt diese in Modelpoints um.
- Dynamische Anpassung: Änderungen in den Fachkonzepten werden durch RAG automatisch erkannt und die Modelpoints entsprechend aktualisiert. Ein KI-Agent identifiziert diese Anpassungen, führt gezielte SQL-Abfragen in den relevanten Datenbanken aus und aktualisiert die betroffenen Modelpoints.

(Rück-)Versicherungsbestand – Verdichtung

- Clusterbasierte Bestandsverdichtung (Model-Point-Compression): Verfahren des unüberwachten Lernens wie k-means können genutzt werden, um Modelpoints anhand relevanter Merkmale (z. B. Alter, Laufzeit, Produkt-/Garantieparameter) in Cluster zu gruppieren und pro Cluster repräsentative Modellpunkte (z. B. Clusterzentren) mit Gewichten abzuleiten. Dadurch lässt sich die Laufzeit von Projektions- und Sensitivitätsrechnungen bei kontrolliertem Genauigkeitsverlust deutlich reduzieren.
- Überprüfung der Gruppierungskriterien: Graph-RAG analysiert die Beziehungen zwischen Verträgen und erkennt Unstimmigkeiten in der Gruppierung, etwa wenn Verträge mit unterschiedlichen Merkmalen fälschlicherweise zusammengeführt wurden. Ein KI-Agent validiert die Ergebnisse, passt fehlerhafte Gruppierungen an und dokumentiert die Änderungen automatisch.
- Kontrolle der Skalierung von Werten: RAG vergleicht die verdichteten Werte, wie z. B. Beiträge und Rückstellungen, mit den bilanziellen Zielwerten, erkennt Abweichungen und überprüft, ob die Skalierung korrekt angewendet wurde. Dabei identifiziert es fehlerhafte Anpassungen, die zu Inkonsistenzen in der Verdichtung führen könnten.

- Konsistenzprüfung der verdichteten Daten: Ein Reasoning-Modell analysiert die Struktur der verdichteten Modelpoints, erkennt logische Widersprüche und deckt Inkonsistenzen in den Gruppierungen auf. Unstimmigkeiten, wie widersprüchliche Risikoklassen innerhalb einer Verdichtungsgruppe, werden identifiziert und für die Korrektur markiert.
- Berücksichtigung von Spezialfällen: RAG durchsucht Dokumentationen nach definierten Regeln für Sonderfälle, wie z. B. seltene Tarife oder Sonderverträge, und überprüft deren korrekte Behandlung in der Verdichtung. Dabei erkennt es, ob bestimmte Tarifgruppen fälschlicherweise zusammengeführt oder unberücksichtigt geblieben sind.

Zinsszenarien

- Prüfung der Kalibrierung: Ein KI-Agent ruft automatisch aktuelle Markt- und Wirtschaftsdaten aus internen und externen Datenbanken ab und stellt sie standardisiert zur Verfügung. RAG überprüft diese Daten auf Konsistenz mit den Kalibrierungsparametern des ESG, erkennt veraltete oder nicht marktgerechte Zinssätze und schlägt aktuelle Referenzwerte für eine präzise Kalibrierung vor.
- Überprüfung der Modellannahmen: Ein Reasoning-Modell vergleicht die Modellannahmen zu Zinsentwicklungen und Inflationsraten mit historischen Daten und aktuellen Marktprognosen. Es identifiziert unrealistische Annahmen, die nicht mit den erwarteten wirtschaftlichen Entwicklungen übereinstimmen, und markiert diese zur Anpassung.
- Backtesting und Benchmarking: Graph-RAG verknüpft ESG-Simulationen mit historischen Marktdaten und alternativen Modellen, um Abweichungen und Muster in der Modellgenauigkeit zu identifizieren. Es erkennt systematische Modellschwächen und hebt kritische Differenzen zu Benchmarks hervor, die auf fehlerhafte Annahmen oder Kalibrierungen hinweisen.
- Prüfung der Abhängigkeitsstrukturen: Graph-RAG modelliert die Abhängigkeiten zwischen wirtschaftlichen Variablen und überprüft deren Logik, insbesondere in Stressszenarien. Es erkennt unplausible Korrelationen, etwa zwischen Aktien- und Zinsmärkten in Krisenzeiten, und deckt Inkonsistenzen in den Modellannahmen auf.

Weitere Eingabedaten

- Parametrisierung: RAG sucht automatisch nach relevanten Parametern (z. B. Management-Regeln) in internen und externen Dokumenten und vergleicht sie mit historischen Daten. Ein KI-Agent passt bei Bedarf Parameter eigenständig an und dokumentiert die Änderungen.
- Startwerte: RAG liest interne Berichte oder Geschäftspläne aus, extrahiert Startwerte und plausibilisiert sie anhand von Vergleichs- und Vorjahresdaten. Ein KI-Agent schlägt Korrekturen vor und führt sie nach Freigabe direkt in den Systemen durch.
- Rechnungsgrundlagen: Graph-RAG kann durch die graphische Darstellung der Beziehungen zwischen verschiedenen Rechnungsgrundlagen (z. B. Sterblichkeit und Kosten) Widersprüche im Gesamtmodell schneller aufdecken. Änderungen in einzelnen Grundlagen können im Kontext weiterer Abhängigkeiten direkt visualisiert werden.

Vorteile der KI-Verfahren

- Strukturierung unstrukturierter Daten: LLMs ermöglichen die semantische Analyse und Umwandlung unstrukturierter Datenquellen (z. B. Dokumente) in strukturierte Formate. Dadurch werden bislang nicht systematisch nutzbare Informationen für die weitere Datenverarbeitung und Qualitätssicherung erschlossen.
- Skalierbare Datenverarbeitung: KI-Modelle sind in der Lage, große Datenmengen effizient, konsistent und automatisiert zu analysieren.

Nachteile der KI-Verfahren

- Hoher Rechenaufwand und Ressourceneinsatz: Insbesondere LLMs sind in der Verarbeitung großer Datenmengen rechenintensiv und erfordern spezialisierte Infrastruktur. Dies kann je nach Anwendungsfall und Datenvolumen zu Performancegrenzen führen.

- Informationsverluste durch Modellabstraktion: Sprachmodelle erzeugen strukturierte Ausgaben auf Basis statistischer Wahrscheinlichkeiten. Dabei kann es zu Detailverlusten, Missverständnissen oder unvollständigen Extraktionen kommen, was bei sensiblen Daten problematisch ist.
- Anpassungs- und Trainingsaufwand: KI-Modelle müssen gezielt auf die jeweilige Problemstellung angepasst werden, um zuverlässig zu funktionieren. Dieser Feintuning-Aufwand reduziert kurzfristig die Skalierbarkeit.

3. Herleitung der biometrischen Annahmen

3.1. Einleitung

Unter Solvency II sind die biometrischen Annahmen ein wesentlicher Teil der Berechnung des besten Schätzwerts der versicherungstechnischen Rückstellungen (BEL). Der BEL wird durch Projektion aller Einzahlungen (z. B. Prämien) und Auszahlungen (z. B. Todesfall- oder Stornoleistungen, Kosten) über die Laufzeit des Versicherungsportfolios ermittelt. Die biometrischen Annahmen legen die Parameter (wie z. B. Sterblichkeit, Stornoraten, Morbidität) dieser Projektion fest, um den Zeitpunkt und die Höhe der Schäden zu schätzen.

Deutsche Lebensversicherer modellieren diese Zahlungsströme auf granularer Ebene unter Verwendung aktuarieller Software. Die biometrischen Annahmen dienen als Eingabeparameter für die Software.

3.2. Festlegung der biometrischen Annahmen – bisherige Erfahrungen

Solvency II stellt Anforderungen an die Annahmen (und die zugrunde liegenden Daten), die zur Berechnung der versicherungstechnischen Rückstellungen verwendet werden.

Um diese Anforderungen zu erfüllen, leiten die deutschen Lebensversicherer ihre biometrischen Annahmen in der Regel auf Grundlage historischer Daten (z. B. als durchschnittliche Erfahrung über drei, vier oder fünf vergangene Beobachtungen) ab, die bei Bedarf durch Expertenurteil erweitert werden. Im Allgemeinen kann der derzeit auf dem Markt verwendete Ansatz zur Bestimmung der biometrischen Annahmen als GLM-Modell oder als vereinfachtes GLM-Modell angesehen werden. Insbesondere trifft dies auf Folgendes zu⁹:

- Stornoannahmen: in der Regel pro Versicherungsjahr und möglicherweise mit einer Aufteilung pro Vertriebsweg.
- Sterblichkeitsannahmen: als Prozentsatz der DAV-Sterblichkeitstabellen mit einer Aufteilung nach Alter und Geschlecht.
- Morbiditätsannahmen bei Berufsunfähigkeit: als Prozentsatz der DAV-Tabellen mit einer Aufteilung nach Alter, Geschlecht und Berufsart.

Diese herkömmlichen Verfahren bieten insbesondere in Bezug auf Transparenz und Modellumsetzung klare **Vorteile**: Die getroffenen Annahmen sind nachvollziehbar dokumentiert, und bei unvollständiger Datenlage erfolgt eine bewusste, systematische Extrapolation oder Übertragung aus vergleichbaren Portfolios. Zudem lässt sich die Implementierung in Bewertungsmodelle durch tabellarische Aufbereitung und standardisierte Schnittstellen vergleichsweise einfach realisieren.

Den genannten Vorteilen stehen jedoch auch **Einschränkungen** gegenüber. Die Modellansätze sind häufig stark vereinfacht, was sich insbesondere in einer begrenzten Anzahl berücksichtigter Parameter niederschlägt. Darüber hinaus bleibt potenziell eine Vielzahl relevanter Einflussfaktoren unbeachtet. Ohne eine multidimensionale Analyse der zugrunde liegenden Datenbasis ist eine fundierte Beurteilung der Eignung und Angemessenheit typischer Annahmefaktoren – wie etwa des Versicherungsjahres bei Stornoannahmen – nur eingeschränkt möglich.

⁹ DAV, Fachgrundsatz „Best Estimate in der Lebensversicherung“ vom 27. Juni 2019

3.3. (Automatisierte) Herleitung der Annahmen durch KI

Der Einsatz von KI zur Herleitung von Annahmen basierend auf historischen Daten

KI-Modelle könnten historische Daten analysieren (z. B. Erfahrungen der letzten Jahre) und Vorhersagen treffen, die dann zur Festlegung von Annahmen verwendet werden.

Diese fortgeschrittenen Modelle (wie beispielweise: NN, GBM oder die RF) sind in der Lage, mehr Wechselwirkungen (Zusammenhänge) in den Daten zu erkennen und ihre Vorhersagekraft erheblich zu steigern. Schwachstellen in der Anwendung der derzeitigen, vereinfachten GLM-Ansätze können erkannt und korrigiert werden, z. B.:

- Rückschluss aus Vergangenheit auf Zukunft: Einschränkungen bei der Extrapolation von Erfahrungen aus älteren Produkten (mit umfangreicherer Erfahrung) auf jüngere Produkte (mit begrenzter Erfahrung).
- Erkennung neuer Risikofaktoren, die über den typischen Ansatz hinausgehen, z. B.:
 - Storno: Das KI-Modell kann potenziell Stornomuster erkennen, die gleichzeitig vom Versicherungsjahr, dem Vermittler, dem Alter des Versicherten, der Art der Beschäftigung oder der Gewinnbeteiligungspolitik des Unternehmens abhängen. Ein interessantes Beispiel wären Policen mit einer Sparkomponente (z. B. aufgeschobene Rente), bei denen die Prämien bis zum offiziellen Rentenalter gezahlt werden. Es ist jedoch bekannt, dass dieses Rentenalter in bestimmten Berufen aus gesundheitlichen Gründen schwer zu erreichen ist. Daher ist es wahrscheinlich, dass ein Kunde mit einer solchen Police im Krankheitsfall entweder die Police vorzeitig kündigt oder in eine prämienfreie Police wechselt. Die Einbeziehung von Informationen über den Beruf des Versicherten in die Stornoanalyse könnte daher helfen, die Wahrscheinlichkeit eines vorzeitigen Rückkaufs im Vorruhestand abzuschätzen.
 - Sterblichkeit: Beim Abschluss einer Versicherung, die im Todesfall des Versicherten eine Leistung erbringt, wird in der Regel eine Gesundheitsprüfung durchgeführt. Dabei übermittelt der Versicherte dem Versicherer Daten über seinen Gesundheitszustand (z. B. Body-Mass-Index (BMI), Blutdruck, Vorerkrankungen). Diese Daten auf Vertragsebene werden meistens nicht direkt für die Bewertung des BEL verwendet. Natürlich wirkt sich die Zeichnungspolitik auf die Realisierung der Sterblichkeit im gesamten Portfolio aus und beeinflusst daher die Art und Weise, wie die Sterblichkeitsannahmen für das gesamte Portfolio festgelegt werden. Durch die Einbeziehung von Gesundheitsinformationen auf der Ebene der einzelnen Verträge können detailliertere mehrdimensionale Annahmen getroffen werden. In diesem Fall hängt die Sterblichkeit nicht nur von der in der Vergangenheit beobachteten Sterblichkeit für das gesamte Portfolio ab, sondern auch von der zukünftigen Struktur des Portfolios (dem Gesundheitszustand der Versicherten im prognostizierten Portfolio).

Vorteile der KI-Verfahren

- Mehr Informationen über die Treiber der Annahmen (Wechselwirkungen / Zusammenhänge):
 - Möglicherweise werden neue wichtige Faktoren entdeckt, die Annahmen beeinflussen.
 - Insbesondere wenn bisherige Untersuchungen nur eindimensional durchgeführt wurden, können Wechselwirkungen zwischen bereits bekannten Faktoren aufgedeckt werden. Es kann sich herausstellen, dass das gemeinsame Auftreten von zwei oder drei Faktoren einen deutlich nichtlinearen Effekt hat.
- Verbesserung der Vorhersagekraft der Annahmen: Durch die systematische und mehrdimensionale maschinelle Analyse großer Datenmengen kann die Qualität des Modells deutlich verbessert werden. Es erfolgt eine Selektion der Faktoren, wobei ausschließlich jene berücksichtigt werden, die zur Optimierung der Vorhersagekraft des Modells beitragen.

Nachteile der KI-Verfahren

- Datenverfügbarkeit: Voraussichtlich werden umfangreichere Daten benötigt als derzeit verfügbar. Dies würde die Einführung neuer Datenverarbeitungsprozesse erfordern, um Annahmen zu treffen. Dies wäre wahrscheinlich ein einmaliger Aufwand.
- (Mögliche) Mangelnde Transparenz: Modelle wie neuronale Netze können sehr komplex und ihre Parametrisierung nicht intuitiv sein. Daher müssen die Analyseergebnisse vom Benutzer validiert werden.
- Veränderungsanalyse: Es schwierig sein, Änderungen der Annahmen in nachfolgenden Berichtszeiträumen zu erklären, wenn diese mit komplexen KI-Modellen berechnet werden. Dies könnte insbesondere dann wichtig sein, wenn diese Änderung wesentliche Auswirkungen auf die Bewertung vom BEL hätte
- Human-in-the-Loop: Das Einholen eines Expertenurteils kann weiterhin erforderlich sein, um sicherzustellen, dass die von der KI getroffenen Annahmen plausibel bleiben. Dies würde wiederum die Einführung von Prozessen zur Überprüfung der Annahmen und möglicherweise einen manuellen Prozess zur Korrektur der Annahmen erfordern.
- Anpassung der Struktur von Bewertungsmodellen: Derzeit werden in Bewertungsmodellen typischerweise Annahmen in Tabellenform verwendet. Es ist sehr wahrscheinlich, dass die Annahmen, die mit Hilfe von KI-Modellen entwickelt werden, komplexer werden und dass eine Anpassung der Bewertungsmodelle an diese komplexen Annahmen erforderlich sein wird.
- Mehrwert aus BEL-Sicht: Obwohl mehrdimensionale Analysen eine Fülle von Informationen über ein Portfolio liefern können, bleibt die Frage, ob all diese Informationen im standardmäßigen vierteljährlichen BEL-Reporting verwendet werden können. Diese Informationen sind für die Preisgestaltung, das Underwriting oder das allgemeine Portfoliomanagement nützlicher.

4. Modellierung

4.1. Einleitung

Die aktuarielle Modellierung bildet das Herzstück der Risikobewertung in der Lebensversicherung. Sie dient der Abbildung komplexer Vertragsbestandteile, Kapitalanlagestrategien sowie des versicherten Kollektivverhaltens über lange Projektionshorizonte hinweg. Diese Modelle liefern die Grundlage für zentrale Unternehmensprozesse wie z. B. Tarifikalkulation, Solvency II Berechnungen, strategische ALM-Steuerung. Der Aufwand zur Erstellung, Pflege und Weiterentwicklung dieser Modelle ist hoch und unterliegt strengen Anforderungen an Nachvollziehbarkeit, Dokumentation und Governance. Durch KI ergeben sich zusätzliche Potenziale, insbesondere im Umfeld von Effizienz, Qualitätssicherung und Automatisierung.

4.2. Modellierung – bisherige Erfahrungen

In der Praxis erfolgt die Modellierung in der Lebensversicherung typischerweise innerhalb spezialisierter Softwareumgebungen (z. B. Prophet oder RiskAgility FM), die sowohl deterministische als auch stochastische Berechnungen unterstützen. Die Modellierung folgt dabei einem strukturierten Prozess:

- Produktabbildung und Policy-Modellierung: die Verträge werden in der Modellplattform möglichst detailliert nachgebildet. Hierzu zählen Tarifelemente, beitragsabhängige Mechanismen, Kundenoptionen (wie Beitragsfreistellung oder Kapitalwahlrechte) sowie Garantien und Überschussmechanismen. Auch versichertenverhaltensabhängige Modellierungen (z. B. Stornoannahmen) fließen ein.
- Kapitalanlage- und Bilanzmodellierung: Parallel zur Produktseite werden Kapitalanlagen modelliert. Dies betrifft sowohl Portfoliostrukturen als auch Bewertungsregeln und Wiederanlagestrategien.

- Cashflow-Projektionen: Die zentrale Modellierungseinheit besteht aus der Projektion von Zahlungsströmen unter Interaktion der Aktiv- und Passiv-Seite des Modells. Diese erfolgt unter Berücksichtigung regulatorischer Vorgaben wie z. B. der Mindestzuführungsverordnung (MindZV).
- Stochastische Bewertung: Um Optionen und Garantien (O&G) adäquat zu messen, werden oft stochastische Simulationen pfadweise durchgeführt.
- Risikomodellierung und Aggregation: die Berechnung der Kapitalanforderung nach Solvency II erfolgt entweder durch Kombination von Stress- und Szenario-Rechnungen für die Standardformel oder mit Hilfe von Replikationsportfolios oder Least-Squares Monte Carlo (LSMC) für das Interne Modell.
- Modellgovernance: Alle Änderungen am Modell sind reversionssicher zu dokumentieren, versioniert zu speichern und nachvollziehbar zu begründen. Dieser Prozess, einschließlich Change-Management, Auswirkungsanalysen, Tests und Freigaben, unterliegt regulatorischen Anforderungen (z. B. MaGo, Solvency II).

In der Praxis haben sich bereits zahlreiche Effizienzmaßnahmen etabliert, um den hohen Komplexitäts- und Dokumentationsanforderungen zu begegnen. Ein wesentlicher Baustein ist die automatisierte individuelle Datenverarbeitung (IDV). Wiederkehrende Schritte, z. B. das Einlesen und Aufbereiten von Ergebnissen oder das Erzeugen von Inputdateien, werden in vielen Unternehmen durch Skripte in VBA, Python oder R automatisiert. Diese Robotic-Process-Automation (RPA)-Ansätze reduzieren manuelle Eingriffe, beschleunigen Abläufe und erhöhen die Reproduzierbarkeit.

Auch bei der Analyse und Auswertung der Modellläufe kommen zunehmend automatisierte Werkzeuge zum Einsatz. Oftmals werden Dashboards genutzt, die direkt an Modelloutputs angebunden sind. Sie ermöglichen eine strukturierte Visualisierung zentraler Ergebniskennzahlen, z. B. ausgewählte Barwertverläufe.

Zur Qualitätssicherung und Unterstützung der Modellentwicklung haben sich standardisierte Testverfahren etabliert. Dazu zählen Regressionstests, die Veränderungsanalysen zwischen Modellversionen, sowie zahlreiche Plausibilitätsüberprüfungen wie z. B. Leakage-Tests. Diese Tests werden in vielen Häusern zentral gepflegt und oft automatisch mit jedem Modellrelease durchgeführt.

Die Ergebnisüberprüfung erfolgt durch mehrere Kontrollinstanzen. Dazu gehören unter anderem Referenzrechnungen mit vereinfachten Modellen oder vordefinierte Auswertungen, Plausibilitätsprüfungen sowie die Umsetzung des Vier-Augen-Prinzips für Ergebnisanalysen. Die Kombination dieser Maßnahmen sorgt nicht nur für fachliche und mathematische Validität, sondern stellt auch die Reversionssicherheit im Sinne regulatorischer Vorgaben (z. B. MaGo oder Solvency II) sicher.

Insgesamt lässt sich beobachten, dass die Modellierungslandschaft zunehmend durch modulare, skriptgestützte und dokumentationsfreundliche Prozesse geprägt ist. Auch ohne den Einsatz fortgeschrittener KI-Methoden ist damit bereits eine hohe Prozessqualität und Effizienz erreicht – zugleich bilden diese Strukturen jedoch eine exzellente Basis für die zukünftige Integration intelligenter Assistenzsysteme mit dem Ziel mehr Zeit für die Ergebnisanalysen zu schaffen.

4.3. Entwicklung von aktuarieller Software mit KI

Es ist gängige Marktpraxis bei deutschen Versicherern, professionelle aktuarielle Software für die aktuarielle Modellierung einzusetzen, z. B., für die Projektion von Passiv-Cashflows zur Ableitung der ökonomischen Bilanz bzw. zur Ableitung von BEL, Eigenmitteln und schließlich der Kapitalanforderungen.

Die Integration von KI in die aktuarielle Modellierung bietet eine Reihe wertvoller Vorteile, die sowohl die Effizienz als auch die Zuverlässigkeit der aktuariellen Prozesse deutlich erhöhen. Einer der Hauptvorteile ist die Verwendung von KI bei der Optimierung- und Generierung von Programmiercode. Sie kann somit bei der Entwicklung, Verfeinerung und Erweiterung von Modellcode verwendet werden. Dabei kann KI z. B. beim Folgenden unterstützen::

- Routinen für die Programmierung bestimmter Tarife vorschlagen.
- Verbesserungen am bestehenden Code vorschlagen (z. B. Wiederholungen vermeiden, Vereinfachungen vorschlagen, um die Geschwindigkeit des Modells zu erhöhen).
- Codierungsfehler vor der Kompilierung erkennen und korrigieren.

Dies führt unmittelbar zu einer Effizienzsteigerung des gesamten Modellierungsprozesses, da die Entwicklung schneller und die Wartung einfacher wird, während gleichzeitig der manuelle Zeit- und Arbeitsaufwand reduziert wird.

Darüber hinaus hebt der Einsatz moderner KI-Anwendungen bestehende Automatisierungen – wie RPA-gestützte Skripte in VBA, Python oder R auf ein neues Niveau. Durch KI lassen sich nicht mehr nur wiederkehrende, klar definierte Aufgaben abbilden, sondern auch komplexere, kontextabhängige Eingriffe automatisieren. So können beispielsweise KI-gesteuerte Workflows bei Modellläufen die Datenvorbereitung, das Aufsetzen von Testläufen oder die Aggregation und Visualisierung von Outputdaten intelligent anpassen z. B. abhängig vom gewählten Szenario.

Ein weiteres bedeutendes Anwendungsfeld liegt in der automatisierten Unterstützung der Dokumentation (siehe auch Kapitel 7). KI kann auf Basis von Codeänderungen, Parameterupdates und Rechenergebnissen eigenständig standardisierte Dokumentationsbausteine erzeugen, etwa für Release Notes, Änderungsprotokolle oder Validierungsunterlagen, und dabei gleichzeitig konsistente Bezüge zu fachlichen Rahmenbedingungen herstellen.

Besonders hervorzuheben ist auch das semantische Verständnis von Code, Dokumentation und Modellen durch KI. Mithilfe moderner Retrieval-Methoden (wie Graph-RAG) kann ein LLM nicht nur Codezeilen analysieren, sondern auch komplexe Zusammenhänge zwischen Modellbestandteilen erkennen, um etwa bei einem Fehler gezielt auf die betroffene Struktur oder Variable hinzuweisen. Dies beschleunigt die Fehlersuche und reduziert die Abhängigkeit von einzelpersonengebundenem Wissen erheblich.

Auch Machine-Learning-Verfahren finden zunehmend Anwendung in der Modellanalyse. So kann Anomalie-Erkennung dazu genutzt werden, inkonsistente oder unerwartete Ergebnisse in stochastischen Pfaden automatisiert zu identifizieren. Gleichzeitig bieten sich Proxy-Modelle an, um gezielte Vorhersagen und Sensitivitätsanalysen durchzuführen. Diese ermöglichen wiederum ein datenbasiertes Backtesting bestehender Modellannahmen und tragen so zur tieferen Validierung und zum besseren Verständnis der Modelllogik bei.

Die Konsistenz der Input- und Output-Daten über eine Vielzahl von Modellläufen hinweg, z. B. im Rahmen von Szenario-Berechnungen, kann durch KI systematisch überwacht werden. Wie bereits in Kapitel 2 beschrieben, lässt sich so z. B. die Plausibilität von Stresseinstellungen automatisiert überprüfen. Dabei erkennen KI-Modelle Abweichungen von erwartbaren Verläufen oder impliziten Modelllogiken, die mit klassischen Testverfahren mit höherem Aufwand erfassbar wären.

Schließlich lassen sich auch weitere Aspekte aus der Modellvalidierung (siehe Kapitel 6) in den Modellierungsprozess direkt integrieren. So kann etwa die KI-gestützte Bewertung von Modellrisiken, Parameterstabilität oder Sensitivitäten bereits im Entwicklungsprozess eingebunden werden, um frühzeitig Hinweise auf kritische Modellkomponenten zu liefern. Damit trägt die KI nicht nur zur Automatisierung, sondern auch zur vorausschauenden Qualitätssicherung bei.

Vorteile der bisherigen Verfahren

- Effizienz: Der Einsatz von KI (insb. GenAI) hebt das Potenzial der Automatisierung und Prozessunterstützung auf ein neues Niveau.
- Qualität: KI-gestützte Analyseverfahren ermöglichen gezielte Qualitätssicherung und konsistente Validierung.
- Flexibilität und Modularität: KI-Anwendungen lassen sich punktuell in bestehende Modellierungsprozesse integrieren. Sie erlauben so einen schrittweisen Investitionsaufbau mit unmittelbarem Nutzen in konkreten wiederkehrenden Aktivitäten.

Nachteile der bisherigen Verfahren

- Qualität und Verlässlichkeit der KI-Ergebnisse: Insbesondere generative KI-Modelle benötigen hochwertige Trainingsdaten und klare Kontexte, was die Skalierbarkeit der Use-Cases reduziert.
- Integration: Der Einsatz von KI in bestehende Modellierungsplattformen (z. B. Prophet, Moses, RiskAgility FM) erfordert technische Schnittstellen, neue Tools und Governance-Prozesse.
- Kontinuierlicher Pflegeaufwand: KI-Anwendungen müssen regelmäßig gewartet, ggf. neu trainiert und überwacht werden.

4.4. Beispiel Show-Case – Modellierung / Testen eines neuen Produkts

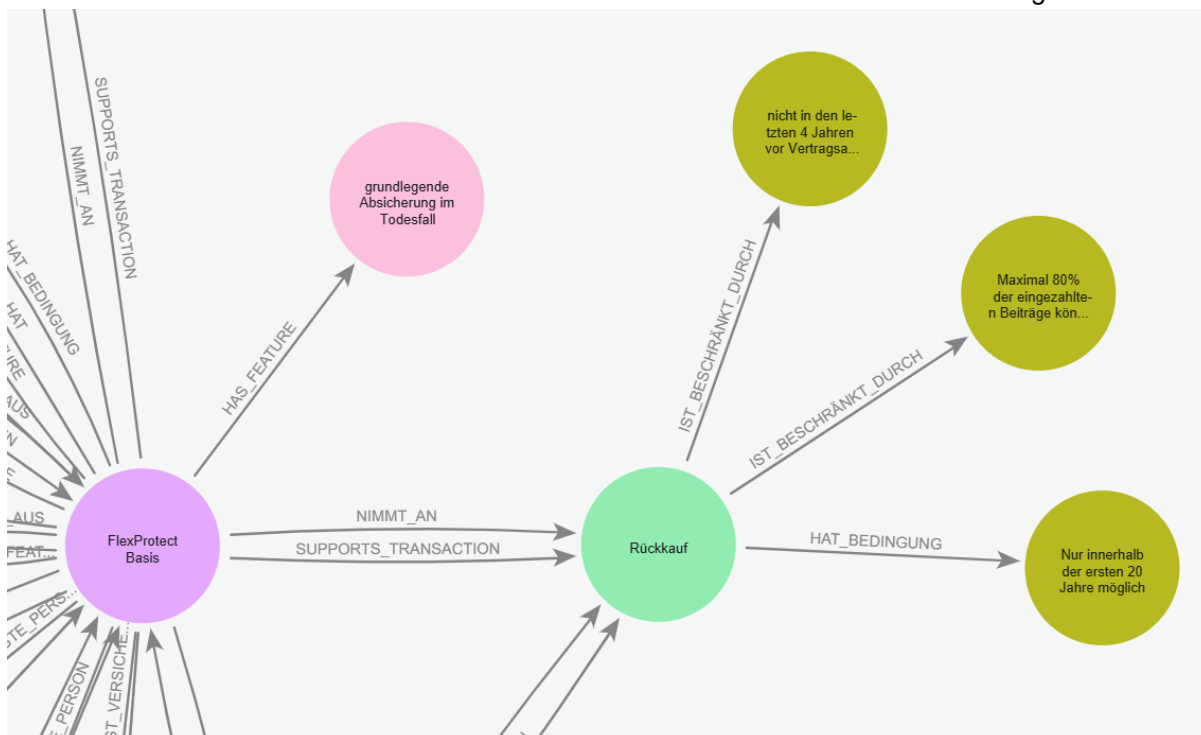
Bei der herkömmlichen Vorgehensweise zur Erstellung synthetischer Modellpoints und Testfälle in der Lebensversicherung muss ein Aktuar das Fachkonzept durcharbeiten und alle relevanten Informationen extrahieren. Dazu gehören Tarifmerkmale, Geschäftsprozesse, Restriktionen und Abhängigkeiten. Anschließend wird basierend auf diesen Informationen ein Plan erstellt, um die Modellpoints und Geschäftsvorfälle korrekt abzubilden. Dieser Prozess ist zeitaufwendig und fehleranfällig, da komplexe Zusammenhänge manuell erfasst und dokumentiert werden müssen.

Mit dem Einsatz von GraphRAG kann dieser Vorgang erheblich optimiert werden. Ein praktisches Beispiel zeigt, dass ein LLM ein Fachkonzept, in dem Tarifmerkmale und Geschäftsvorfälle beschrieben werden, in eine strukturierte Graph-Darstellung umwandeln kann. Im Rahmen einer Analyse wurde ein Fachkonzept für den beispielhaften (fiktiven) Tarif „FlexProtect Basis“ untersucht. Die darin enthaltenen Informationen zu Tarifen, Attributen und Geschäftsvorfällen waren über das gesamte Fachkonzept verstreut. Das LLM konnte automatisch einen Graph erstellen, in dem der Tarif mit den korrekten Geschäftsvorfällen und deren Bedingungen verknüpft wurde, wodurch eine klare und strukturierte Übersicht der relevanten Zusammenhänge entstand.

So wurde z. B. der Geschäftsvorfall Rückkauf identifiziert und mit den entsprechenden Bedingungen verknüpft:

- Nicht in den letzten 4 Jahren vor Vertragsablauf möglich.
- Maximal 80% der eingezahlten Beiträge können erstattet werden.
- Nur innerhalb der ersten 20 Jahre möglich.

Der Graph bietet eine übersichtliche Darstellung der Zusammenhänge zwischen Tarifmerkmalen, Geschäftsvorfällen und Restriktionen. Dadurch entfällt die manuelle Zusammenführung durch den



Aktuar. Es bietet somit automatisch eine strukturierte Basis für die Erstellung synthetischer Modellpunkte und Testfälle. Auch die bestehenden Modellteile ähnlicher Produkte können so identifiziert werden und entweder in ein neues Modell überführt werden, oder einzelne Module in der Implementierung direkt genutzt werden.

5. Berechnungen

5.1. Einleitung

Mit der Einführung von Solvency II sind stochastische Modelle zum Marktstandard in der Bewertung von Lebensversicherungsportfolios geworden. Die Berechnungen, die den gesamten Versicherungsbestand umfassen, verschiedene Stufen der Risikoaggregation beinhalten oder zum Teil auf „Nested-Simulations“ basieren, erfordern eine sehr hohe Rechenintensität und Komplexität der Berechnungsprozesse. Hinzu kommen umfangreiche interne und externe Anforderungen an das Reporting, das verschiedenste Analysen, Szenarien und Sensitivitäten in unterschiedlicher Granularität verlangt. All dies führt zu einer fortschreitenden Digitalisierung und Automatisierung der Berechnungsprozesse bei Lebensversicherungsunternehmen.

Mit der Digitalisierung der Branche seit den 2000er-Jahren beobachten wir eine moderne Automatisierung der Datenverarbeitung und Berechnungsprozessen wie z. B. digitale Datenerfassung, digitale Schnittstellen, die Einführung von Workflow-Systemen sowie Speicherung und Verarbeitung großer Datenmengen in Echtzeit auf Cloud-Plattformen.

Die Entwicklung der Risikomodelle ist eng mit dem technologischen Fortschritt verknüpft und schreitet weiterhin dynamisch voran. Die Künstliche Intelligenz birgt weiteres Potenzial für die Automatisierungen der Prozesse.

5.2. Berechnungen – bisherige Erfahrungen

Die Standardisierung der Prozesse und die automatisierte Datenaggregation ersetzen manuelle Arbeitsschritte. Trotzdem existieren in den Workflows immer noch zahlreiche Excel-Vorlagen, diese werden jedoch oft qualitätsgesichert und automatisiert eingelesen oder erstellt. Hierfür gibt eine Vielzahl von Workflow-Tools, die Unternehmen helfen, Prozesse zu automatisieren.

Die großen Bestandsführungssysteme unterliegen eigenen IT-Qualitätsregeln. Viele der Systeme im Risikomanagement oder im Aktuariat werden als IDV-Tools deklariert (aufbauend auf Excel (VBA), Prophet, Risk Agility FM (ehem. MoSes), PLA.NET, SAS).

Ob vor Ort (on-premise) oder in der Cloud, moderne aktuarielle Software ist in der Lage die Rechenleistung durch Grid-Computing-Cluster zu maximieren und senkt so die Kosten und steigert die Produktivität. Es ist heutzutage möglich einen vollautomatischen Prozess für die Berechnung zu etablieren einschließlich komplexer und aufwendiger Prozeduren wie Nested-Simulations.

Cloud Computing ermöglicht parallele Verarbeitung großer Datenmengen, insbesondere bei datenintensiven oder rechenaufwändigen Aufgaben wie Risikomodelle. Dabei gibt es eine Reihe wichtiger Aspekte zu beachten:

- Datensicherheit und Datenschutz
- Rechenleistung und Skalierbarkeit (Bedarfsgerechte Ressourcen)
- Validierung und Reproduzierbarkeit
- Ausfallsicherheit und Backup

Die Qualitätssicherung kombiniert regelbasierte Validierungen mit Plausibilitätsprüfungen auf Basis standardisierter Auswertungen und Visualisierungen. Seit der Einführung von Solvency II im Jahr 2016 hat sich die Dauer des Abschlussprozesses für Versicherungsunternehmen deutlich verkürzt, sowohl durch regulatorisch bedingte Anpassungen als auch durch technologische Entwicklungen. Die Einführung von digitalen Reporting-Plattformen, automatisierten Datenpipelines und Cloud-basierte Analysewerkzeuge hat sich die Bearbeitungszeit deutlich reduziert.

5.3. Automatisierung durch KI

Innerhalb der Berechnungsprozessen gibt es eine Vielzahl an manuellen Aufgaben, welche sich dazu eignen, völlig automatisiert abgearbeitet zu werden. Vor allem bei der Verarbeitung der stetig wachsenden Menge an Daten, die den Versicherungsunternehmen zur Verfügung stehen sowie bei häufig durchgeführten Berechnungen oder Berechnungen mit hohem Volumen kann der Einsatz von KI erhebliche Effizienzen mit sich bringen.

Inzwischen gibt es eine Vielzahl an Lösungen auf dem Markt, die verschiedene Bereiche der Automatisierung abdecken: Datenaufbewahrung (Cloud), Datenanalyse, Visualisierung oder Abläufe (Workflow).

Beispielweise kann der Prozess der Aufbereitung der Kapitalanalagen für die stochastischen Modelle durch automatisierte Abläufe harmonisiert und beschleunigt werden. Vom Bestandsabzug und Expertenschätzungen bis hin zu verschiedenen Schritten wie Look-Through, Klassifizierung und Konvertierung im Endformat können die Bestände standardisiert und konsistent für verschiedene Berichtserstattungen (Forecast und Planung, Solvency II, IFRS 17, ALM) schneller und qualitätsgesichert erzeugt werden.

Bei Modellen, die den Schritt der Kalibrierung benötigen – z. B. von Kapitalmarktszenarien oder bei den Liability Cashflows – kann KI eingesetzt werden, um die Überprüfung der Güte nach einem Optimierungslauf einzuschätzen und damit die Kalibrierungsroutine effizienter und schneller zu machen.

Bei der Berechnung von Sensitivitäten oder Veränderungsanalyse können KI-Algorithmen automatisierte Läufe mit den geeigneten Parametern durchführen und eine erste Analyse und Visualisierung der Ergebnisse liefern.

Bekannterweise ist die Veränderungsanalyse ein sehr rechenintensiver Prozess, der jedoch unverzichtbar ist und häufig in ausreichender Granularität durchgeführt werden muss, um die Effekte korrekt zu trennen. KI kann nicht nur bei der Steuerung einzelner Berechnungsschritte eingesetzt werden, sondern auch bei der Datensammlung aus verschiedenen Sensitivitäten, den Ergebnissen und Informationen aus vorangegangenen Modelländerungen und sogar bei der Suche nach Erklärungen aus der Historie unterstützen.

KI kann auch für die Optimierung und Transformation von stochastischen Prozessen genutzt werden. Durch den Einsatz von KI in den Berechnungsprozessen verändert sich auch der Ablauf, die Struktur die Art und Weise, wie Daten aufbewahrt und weiterverarbeitet werden sowie wie die Schnittstellen miteinander verknüpft sind. KI kann verwendet werden, um die Prozesse rund um die stochastischen Berechnungen (für Solvency II, IFRS 17, ALM) oder auch im Risikomanagement zu optimieren. Somit ergeben sich Zeitersparnisse und Fachpersonal muss nur noch gezielt eingreifen.

Vorteile bei der Automatisierung von Berechnungsprozessen mit KI

- Automatisierung komplexer Aufgaben: Automatisierte Berechnungen, anstatt manueller Abläufe, reduzieren die Fehleranfälligkeit
- Genauigkeit: Durch KI-optimierte Berechnungen können die Ergebnisse eine höhere Genauigkeit der Vorhersage ausweisen.
- Schnellere und effizientere Berechnungen: Durch KI gesteuerte Berechnungen können wesentliche Zeitersparnisse mit sich bringen und Prozesse optimieren.

Nachteile bei der Automatisierung von Berechnungsprozessen mit KI

- Kostensteigerung: Die Umsetzung von KI-Lösungen kann möglicherweise zu höheren oder zusätzlichen Kosten führen.
- Ressourcen: Der Einsatz von KI erhöht den Bedarf an qualifizierten Mitarbeitern, sowie technische Ressourcen wie z. B. Rechnerkapazitäten.

5.4. KI-proxy bzw. Ersatz inkl. Fortschreibung / Prognose von SCR

Kontext

Solvency II verlangt von den Versicherern, dass sie über genügend Kapital verfügen, um mit einer Wahrscheinlichkeit von 99,5% während eines Jahres solvent zu bleiben.

Die Berechnung dieser Kapitalanforderung (SCR) erfolgt in der Praxis entweder über einen vereinfachten Ansatz (z. B. Standardformel) oder über Monte-Carlo-Berechnungen (internes Modell)¹⁰.

Bei der Standardformel der EIOPA werden die Risiken (und Risikomodule) über einen linearen Korrelationsansatz aggregiert. Bei der Monte-Carlo-Methode wird dagegen eine Vielzahl möglicher Szenarien für die Kapitalmarktentwicklung und den Schadenverlauf generiert, um deren Auswirkungen auf die Eigenmittel des Versicherers zu messen. Das SCR wird dann auf Basis des 99,5%-Quantils der stochastischen Berechnungen bestimmt.

Die Anwendung der Monte-Carlo-Methode zur Bestimmung des SCR für das deutsche gewinnberechtigte Geschäft könnte typischerweise zu einem Nested-Stochastic-Ansatz führen, da für eine große Anzahl von Punkten der Risikoverteilung eine vollständige stochastische Bewertung durchgeführt werden muss. Siehe Abbildung unten.

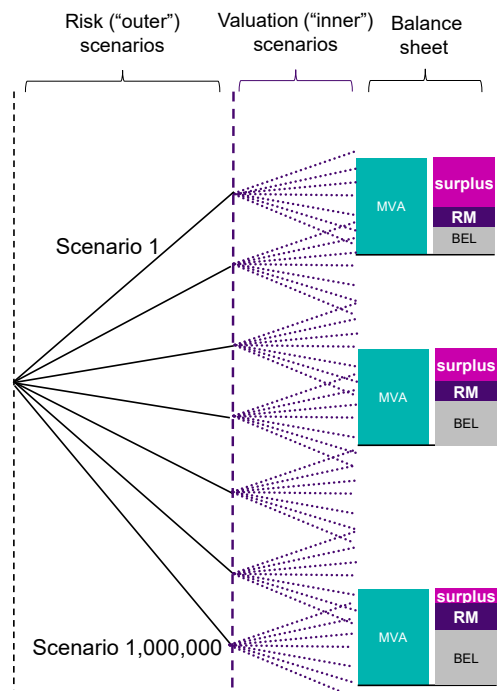


Abbildung 5-1 – Illustration des Nested-Stochastic-Ansatzes für die SCR-Berechnung

Das SCR ist dann als die Differenz zwischen den Eigenmitteln im Basisfall und den Eigenmitteln im 0,5% schlechtesten Szenario definiert.

Proxy-Modelle

Um das Problem der großen Anzahl von Nested-Stochastic-Berechnungen zu lösen, wurden Proxy-Ansätze entwickelt, bei denen der Wert der Verbindlichkeiten oder der Eigenmittel in einem bestimmten Szenario auf der Basis eines Proxys, d.h. einer Näherung, bestimmt wird.

Im europäischen (teilweise auch im deutschen) Markt haben sich drei Proxy-Ansätze etabliert, Curve Fitting, Least Squares Monte Carlo (LSMC) und Replicating Portfolio.

Berechnung der Eigenmittel anhand des Proxys

Wenn ein Proxy bereits kalibriert/gefunden wurde, wird die Höhe der Eigenmittel wie folgt bestimmt – siehe Abbildung unten.

¹⁰ Siehe z. B. DAV, Ergebnisbericht des Ausschusses Investment „Proxy-Modelle für die Risikokapitalberechnung“ vom 8 Juli 2015

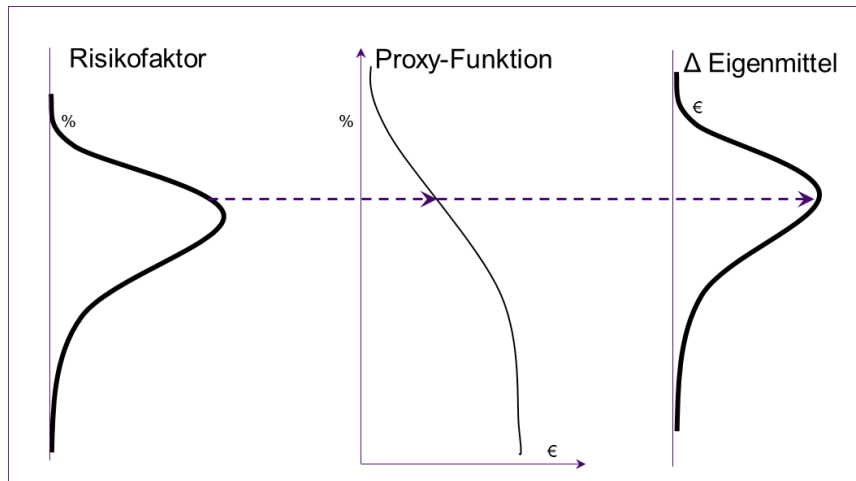


Abbildung 5-2 - Anwendung von Proxy bei der Berechnung der Eigenmittelveränderung

Spezifikation der Proxy-Funktion

Die Kalibrierung der Proxy-Funktion basiert entweder auf ausgewählten Sensitivitäten (Curve Fitting) oder auf einem Regressionsansatz (LSMC - LS = „least square“). Im letzteren Fall wird die Proxy-Funktion auf der Grundlage der zahlreichen potenziell ungenauen Berechnungen des Liability Modells approximiert.

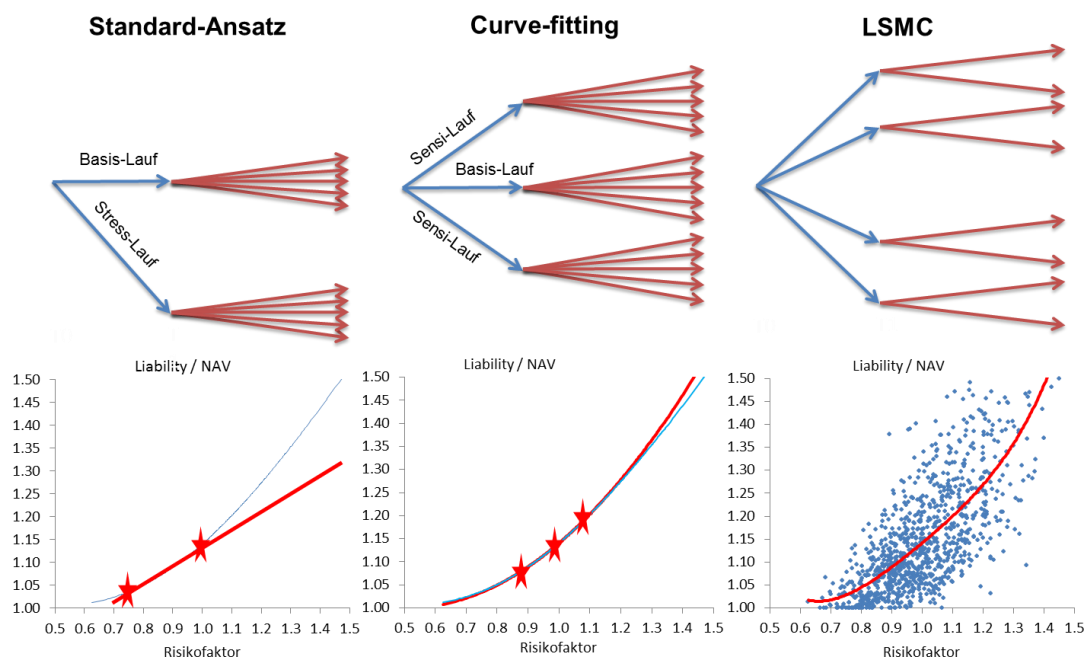


Abbildung 5-3 – Kalibrierung der Proxy-Funktion im Vergleich zu den Berechnungen der Stress Szenarien im Standard-Ansatz - Illustration

LSMC - die in den letzten Jahren am häufigsten eingeführte Proxy Methode

Die Vor- und Nachteile der Proxy-Modellierung sind bekannt, so ist z. B. LSMC - die Methode, die in den letzten Jahren aufgrund ihrer guten mathematischen Basis am häufigsten eingeführt wurde, zeitaufwendig und oft nicht transparent. LSMC verwendet Polynome (eine Funktion der Risikofaktoren „RF“) als Approximation der Verlustfunktion, z. B.

$$\text{Eigenmittel} = \text{Const} + a_1RF_1 + a_2RF_2 + \dots + b_1RF_1^2 + b_2RF_2^2 \dots + \dots + a_{11}RF_1RF_2 + \dots$$

Polynome höheren Grades sind jedoch in der Praxis oft nicht mehr intuitiv erklärbar.

Moderne Proxy-Modelle – KI im Einsatz

Mit der wachsenden Computerleistung fanden KI-Methoden zunehmend Anwendung – vor allem in Bereichen mit großen Datenmengen. Im Falle des LSMC-Ansatzes wird eine große Menge an Daten generiert, um die Polynome zu kalibrieren, die die Verlustfunktion approximieren. Dies hat die Aktuare dazu motiviert, andere moderne Methoden als Proxys in Betracht zu ziehen, wie z. B. GBM oder NN.

Vorteile der KI-Proxy-Modelle

- Qualitätssteigerung: Der Einsatz der KI-Methode führt zu einer signifikanten Qualitätssteigerung der eingesetzten Proxys. Schwächen, die sich aus der Verwendung von Polynomen ergeben, wie beispielsweise die Nichtberücksichtigung von Unregelmäßigkeiten oder gar Sprüngen in der Verlustfunktion, können eliminiert oder deutlich reduziert werden.
- Vielfältige Einsatzmöglichkeiten: Im Gegensatz zu individuellen Szenariorechnungen steckt bei Proxymodellen ein Großteil des Rechenaufwands im Kalibrieren des Modells. Anschließend ist es jedoch verhältnismäßig ressourcenschonend das Modell für verschiedene Use-Cases zu verwenden. Dadurch ergeben sich neue Anwendungsfälle, die sonst aufgrund fehlender Kapazität nicht umsetzbar wären.

Nachteile der KI-Proxy-Modelle

- Herausforderungen bei der Interpretierbarkeit: Das KI-Proxy-Modell wird zu einer Blackbox, deren Parametrisierung, selbst wenn sie bekannt ist, nicht leicht zu interpretieren ist. In diesem Zusammenhang ist anzumerken, dass die Modellparametrisierung bei der Anwendung auf die bestehenden Methoden (Replicating Portfolio, LSMC) ebenfalls oft schwer zu interpretieren ist.
- Validierung: Modelle, die wie Blackboxen aussehen, müssen einer sorgfältigen Validierung unterzogen werden. In diesem Zusammenhang können bereits etablierte Methoden Anwendung finden, wie beispielsweise eindimensionale Darstellungen oder Cross-Terms Darstellungen und auch Out-of-Sample-Tests.
- Human-in-the-Loop: Das Einholen eines Expertenurteils kann weiterhin erforderlich sein, um sicherzustellen, dass die von der KI getroffenen Proxys auch in den Extremszenarien plausibel bleiben.
- Kosten: Die Parametrisierung KI-basierter Proxys erfordert eine große Menge an Daten, die auf Grundlage ökonomischer Szenarien und der Ergebnisse von Liability-Modellen generiert werden müssen. Der Rechenaufwand kann potenziell signifikant sein. Bei der Verwendung der gebräuchlichen LSMC-Methode müssen jedoch ähnliche Kosten berücksichtigt werden.

Alternative Anwendungen der KI-Proxy Modelle

- Verbesserung der bestehenden Proxy-Modelle: Die Erprobung alternativer Modelle für die Proxy-Funktionen kann je nach Komplexität und Ausgereiftheit der untersuchten Modelle zu weiteren Erkenntnissen über die Daten und möglicherweise zu einer Verbesserung der Anpassungen der bestehenden Proxy-Modelle führen. Der Anwender der replizierenden Portfoliomethode kann KI nutzen, um nach alternativen Finanzinstrumenten zu suchen und so die Vorhersagequalität der Proxys zu verbessern. LSMC-Benutzer haben die Möglichkeit, die Schwächen des aktuellen Modells zu testen.
- Fortschreibung / Forecast / Planung / Veränderungsanalyse: Im Rahmen der Diskussion über Proxy-Modelle haben wir uns bislang auf die Anwendung von Proxy-Modellen bei SCR-Berechnungen unter Verwendung eines internen Modells konzentriert. Proxys können auch von Benutzern des Standardmodells verwendet werden. Anwendungsfälle können dabei z. B. Veränderungsanalyse sowie Sensitivitäten des Eigenkapitals und SCRs oder der Forecast und die Planung sein.

6. Validierung

6.1. Einleitung

Die Validierung unter Solvency II ist ein zentraler Bestandteil der regulatorischen Anforderungen für Versicherungsunternehmen.

Die regulatorischen Grundlagen der Validierung sind in der Delegierten Verordnung (EU) 2015/35, insbesondere in Art. 241–264, VAG¹¹, sowie in den EIOPA-Leitlinien konkretisiert

Die Validierung umfasst mehrere Ebenen und ist eng mit dem sogenannten Drei-Säulen-Modell unter Solvency II verknüpft. Die Validierung in Säule 1 betrifft insbesondere die Modellierung der Risiken, u.a. Annahmen, Methoden, Datenqualität, Ergebnisse, einschließlich Backtesting und Sensitivitätsanalysen. Die Ergebnisse müssen dabei plausibel, nachvollziehbar und konsistent mit den zugrunde liegenden Annahmen sein. In Säule 2 umfasst die Validierung die Überprüfung der Governance-Strukturen, der Risikomanagementprozesse und der internen Kontrollsysteme. In Säule 3 müssen die Inhalte in sämtliche Berichte wie z. B. im Solvency and Financial Condition Report (SFCR) und im Regular Supervisory Report (RSR) validiert werden.

Im Kontext von Solvency II unterscheidet sich die Validierung von Modellen deutlich, je nachdem, ob klassische Verfahren oder Methoden der Künstlichen Intelligenz eingesetzt werden. Während klassische Modelle stark auf mathematische Nachvollziehbarkeit und regulatorische Konformität setzen, erfordern KI-Modelle zusätzliche Maßnahmen zur Erklärbarkeit, Transparenz und Datenqualität. Die Validierung wird dadurch komplexer, aber auch potenziell leistungsfähiger – insbesondere bei der Prognose und Risikobewertung.

6.2. Validierung – bisherige Erfahrungen

Durch die Validierung wird sichergestellt, dass die Gültigkeit der Ergebnisse in allen relevanten Aspekten gewährleistet ist.

Heutzutage ist Marktstandard neben der Regelmäßigkeit des Prozesses auch ein Wirksames statistisches Verfahren zu etablieren, das die Angemessenheit des Modells und der Ergebnisse nachweisen kann. Bei der klassischen Modellvalidierung werden oft sowohl quantitative als auch qualitative Kriterien definiert.

Die Validierung findet an unterschiedlichen Stellen im Gesamtberechnungsprozess statt. Eine unmittelbare Validierung des Outputs jedes einzelnen Schritts ist unverzichtbar. Die Validierung unter Solvency II erfolgt systematisch und umfasst mehrere Ebenen: Daten, Annahmen, Methoden, Modelle, Berechnungen und Ergebnisse. Die klassische Validierung kann zum Beispiel so gestaltet werden:

Validierung der Daten

Solvency II verlangt, dass alle Berechnungen – insbesondere der versicherungstechnischen Rückstellungen und der Solvenzkapitalanforderung – auf vollständigen, genauen und angemessenen Daten basieren. Die Validierung prüft, ob diese Anforderungen erfüllt sind durch beispielweise folgende Prüfungen (Siehe auch Kapitel 2. Datenaufbereitung und Qualitätssicherung):

- Plausibilitätsprüfungen: Vergleich mit historischen Daten, Benchmarks oder externen Quellen; Automatisierte Prüfregeln (z. B. „Geburtsdatum darf nicht in der Zukunft liegen“); Prüfung auf fehlende oder doppelte Einträge.
- Vollständigkeits- und Konsistenzprüfungen: Sicherstellung, dass keine Daten fehlen und dass sie logisch zusammenpassen.
- Datenqualitätsindikatoren: Einsatz von KPIs zur Bewertung der Datenqualität.

Validierung der Annahmen

¹¹ Delegierte Verordnung (EU) 2015/35 der Kommission vom 10. Oktober 2014

Die Validierung der Annahmen zielt hauptsächlich darauf ab sicherzustellen, dass die Annahmen realistisch, konsistent und nachvollziehbar sind. In der bisherigen Praxis werden unter anderem folgende Verfahren bei der Validierung der Annahmen eingesetzt:

- Statistische Tests: theoretische und empirische Analyse der Verteilung anhand Histogramme, QQ-Plots und Dichtevergleiche, sowie verschiedene Goodness-of-Fit-Tests.
- Backtesting: Vergleich von Annahmen mit realen (historischen) Entwicklungen (z. B. Sterblichkeit, Storno, Schadenverläufe). Auch Vergleich mit branchentypischen Sterbetafeln oder internen Erfahrungswerten.
- Sensitivitätsanalysen: Untersuchung, wie stark sich Ergebnisse bei Änderung einzelner Annahmen verändern.
- Expertengutachten: Validierung durch unabhängige Aktuarer oder Fachabteilungen.

Vielfältige Hinweise zu Validierung von ESG sowohl für risiko-neutrale als auch für real-world Berechnungen sind im DAV-Ergebnisbericht „Anforderungen an einen ökonomischen Szenariengenerator“¹² zu finden.

Validierung der Berechnungen

Zur Überprüfung der Berechnungen im klassischen Sinne existieren einige etablierten Verfahren:

- Rechenprüfungen: manuelle oder automatisierte Nachrechnungen von Stichproben.
- Vergleich mit alternativen Methoden: z. B. Vergleich interner Modelle mit Standardformel-Ergebnissen.
- Replikationsmethoden: Nachbildung der Ergebnisse mit unabhängigen Tools oder Modellen.

Validierung der Modelle

Die zur Risikobewertung eingesetzte Modelle müssen regelmäßig validiert werden, um sicherzustellen, dass sie die Risiken realistisch abbilden und keine systematischen Fehler enthalten.

Laut EIOPA müssen Versicherer für (interne) Modelle:

- Modellstruktur und -annahmen dokumentieren: insbesondere in Bezug auf die Nachvollziehbarkeit sowie die Angemessenheit für den jeweiligen Zweck.
- Modellimplementierung: technische Korrektheit, z. B. durch Vergleich mit Referenzrechnungen; Prüfung der Modelllogik durch Experten (z. B. ob Renten korrekt berechnet werden); Regressionstests: Sicherstellen, dass Änderungen im Modell keine unbeabsichtigten Effekte haben.
- Regelmäßige Validierungen durchführen: welche mindestens jährlich durchzuführen sind.
- Unabhängigkeit der Validierung sicherstellen: eine Möglichkeit diese Unabhängigkeit sicherzustellen ist z. B. durch separate Teams.
- Modellrisiken identifizieren und bewerten: diese sind unter anderem Angemessenheit des Modells, Robustheit, Präzision sowie weitere Anforderungen.

Falls das Versicherungsunternehmen die Standardformel benutzt, ist es wichtig die Konsistenz der eigene Risikostruktur mit der im Standardmodell gegebenen Strukturen zu überprüfen. Bei der Anwendung des Branchensimulationsmodells (BSM) ist es notwendig die Angemessenheit des Modells für den eigenen Bestand darzulegen und die Vereinfachungen zu analysieren.

Typische Methoden für die allgemeine Modellvalidierung:

- Benchmarking: Vergleich mit Modellen anderer Unternehmen oder Marktstandards.
- Stress- und Szenarioanalysen: Prüfung der Modellreaktion auf Extremereignisse.
- Reverse Stress Testing: Identifikation von Szenarien, die zu Modellversagen führen würden.

¹² DAV, Fachgrundsatz „Anforderungen an einen ökonomischen Szenariengenerator“ vom 27. November 2023

Validierung der Ergebnisse

Gemäß Solvency II (Art. 48 der Richtlinie 2009/138/EG und § 31 VAG) ist die Versicherungsmathematische Funktion (VmF) für die Validierung der Rückstellungen verantwortlich. Die Validierung der SCR-Berechnung umfasst typischerweise Überprüfung der SCR-Struktur, Analyse der Veränderung in den Risikofaktoren und die Exponierung, Korrelation und Aggregationseffekte sowie Wechselwirkungen. Oft werden folgende Validierungsmethoden verwendet:

- Ergebnisanalyse: Vergleich mit Vorjahreswerten, Prognose oder Erwartungswerten, Sensitivitätsanalysen.
- Erklärung von Abweichungen: Ursachenanalyse bei signifikanten Differenzen.
- Management Review: Einbindung der Geschäftsleitung zur Plausibilisierung.
- Backtesting: Vergleich von Modellprognosen mit realen Entwicklungen.
- Benchmarking: Vergleich mit externen Quellen oder anderen Unternehmen.

Weitere Details sind in dem DAV-Ergebnisbericht „Validierung der versicherungstechnischen Rückstellungen unter Solvabilität II“ zu finden¹³.

6.3. Validierung mit KI

Der Einsatz von KI bringt deutlich höhere Komplexität und damit neue Herausforderungen mit sich und erfordert entsprechend angepasste Validierungsansätze. Die Validierung stochastischer Unternehmensmodelle mit KI kann dazu beitragen die Vorhersagekraft und die Robustheit solcher Modelle zu verbessern.

Bei der Validierung unter Einsatz von KI ist besonders wichtig die Erklärbarkeit und Transparenz der Modelle zu gewährleisten. Black-Box-Modelle wie neuronale Netze müssen durch Methoden wie Surrogatmodelle oder Explainable AI (XAI) nachvollziehbar gemacht werden. Der Erklärbarkeit kommt mit steigender Komplexität der Modelle eine zunehmend wichtige Rolle zu. Begriffsdefinition, prinzipienbasierte Kriterien, wann ein Modell als ausreichend erklärbar gilt, sowie verschiedene Erklärbarkeitsmethoden werden im DAV-Ergebnisbericht „Explainable Artificial Intelligence: Ein aktueller Überblick für Aktuarinnen und Aktuare“¹⁴ näher erläutert. Der erforderliche Grad der Erklärbarkeit hängt dabei von zwei Determinanten ab: der Anwendung und den Adressaten.

Nicht zuletzt sind die relevanten regulatorischen Anforderungen an die Validierung zu beachten. Mit der KI-Verordnung wurde die Grundlage der KI-Regulierung gelegt. Die KI-Verordnung legt sieben Grundprinzipien fest. Neben laufende Überwachung und Anpassung der Modelle, was auch für die klassische Modelle und Validierung gilt, Transparenz und Erklärbarkeit, wird die Anforderung „Menschliche Aufsicht“ gestellt. Obwohl KI zunehmend die Automatisierung von Aufgaben und Prozessen ermöglicht, muss immer ein gewisses Maß an menschlicher Beteiligung an der Bereitstellung von KI-Systemen entlang der verschiedenen Phasen des Lebenszyklus von KI-Modellen gewährleistet sein. Weitere Details insb. zur Validierung von KI-Modellen können im DAV-Ergebnisbericht „Regulierung und Validierung von KI-Modellen“¹⁵, Kapitel 4 entnommen werden.

Datenanalysen

KI-Algorithmen können eingesetzt werden, um die Qualität der Daten zu überprüfen, fehlerhafte, unvollständige oder inkonsistente Daten zu identifizieren und zu korrigieren. Dies umfasst zum Beispiel die Erkennung von Duplikaten, die Korrektur von Tippfehlern und die Vervollständigung fehlender Daten. Eine detaillierte Auflistung an Anwendungsfeldern kann Kapitel 2 entnommen werden.

¹³ DAV, Ergebnisbericht des Ausschusses Enterprise Risk Management „Validierung der versicherungstechnischen Rückstellungen unter Solvabilität II“ vom 16. November 2018

¹⁴ DAV, Ergebnisbericht des Ausschusses Actuarial Data Science „Explainable Artificial Intelligence: Ein aktueller Überblick für Aktuarinnen und Aktuare“ vom 27. Mai 2024

¹⁵ DAV, Ergebnisbericht der Ausschüsse Rechnungslegung und Regulierung und Actuarial Data Science „Regulierung und Validierung von KI-Modellen“ vom 26. Februar 2024

Risikofaktoren

Die Auswahl und Transformation von Variablen kann durch den Einsatz von KI-Methoden wie Clustering (z. B. K-Means) oder Dimension-Reduktion (z. B. PCA) erfolgen, um wichtige Merkmale zu extrahieren oder die Validität der eingesetzten Faktoren zu überprüfen.

Validierung von stochastischen Simulationen

- Clustering: Mit Algorithmen wie K-Means oder DBSCAN können die Ergebnisse der Simulation in verschiedene Gruppen oder Cluster eingeteilt werden. Dies hilft, Muster und Ähnlichkeiten in den Daten zu erkennen.
- Fehleranalyse: Bei verschiedenen Kalibrierungsproblemen oder Backtesting können Deep Learning-Modelle, Random Forest (RF) und Support Vector Machines (SVM) benutzt werden, um die Residuen (Differenzen zwischen den tatsächlichen und prognostizierten Ergebnissen) oder systematische Abweichungen eines stochastischen Modells zu analysieren und dabei helfen, systematische Fehler oder Muster zu erkennen. Wenn das stochastische Modell z. B. eine zu hohe Schätzung für eine bestimmte Kundengruppe liefert (z. B. bestimmte Altersgruppen oder Risikoprofile), können RF und SVM-Modelle genutzt werden, um zu analysieren, welche Variablen (wie z. B. Alter, Geschlecht, Gesundheitsstatus) zu den größten Abweichungen führen. Ein Random Forest kann dabei helfen, die wichtigsten Einflussfaktoren auf die Modellfehler zu bestimmen, indem es die Relevanz verschiedener Merkmale misst. Eine SVM kann dann verwendet werden, um zu untersuchen, ob es nicht-lineare Zusammenhänge oder Grenzwerte gibt, die das stochastische Modell möglicherweise nicht berücksichtigt hat.
- Anomalie Erkennung: Algorithmen zur Anomalie Erkennung, wie Isolation Forest oder Autoencoder, können verwendet werden, um ungewöhnliche oder unerwartete Ergebnisse in den Simulationen zu identifizieren. Dies ist besonders nützlich, um potenzielle Fehler oder seltene Ereignisse zu erkennen.
- Visualisierung: Visualisierungstechniken können eingesetzt werden, um die Ergebnisse der Simulation zu analysieren und zu interpretieren. Tools wie t-SNE oder UMAP können helfen, hochdimensionale Daten in verständliche Visualisierungen zu transformieren.
- Zeitreihenanalyse: Stochastische Modelle in der Lebensversicherung versuchen oft, zukünftige Entwicklungen (wie z. B. Sterblichkeit, Morbidität, Prämienveränderungen) basierend auf historischen Daten zu schätzen. Algorithmen zur Zeitreihenanalyse wie Autoregressive Integrated Moving Average- (ARIMA) oder Deep Learning-Modelle wie rekurrente neuronale Netze (RNN) oder Long Short-Term Memory (LSTM)-Netzwerke können verwendet werden, um (langfristige) Trends und Muster über die Zeit zu erkennen und zu überprüfen, ob diese ausreichend gut modelliert sind. Mit Neural Networks (NN) kann die Vorhersagegenauigkeit eines stochastischen Modells getestet werden, indem das NN auf denselben historischen Datensatz trainiert wird und anschließend die Vorhersagen mit den tatsächlichen Werten verglichen werden. Deep Learning-Modelle wie LSTM-Netze (Long Short-Term Memory), die speziell für zeitabhängige Daten geeignet sind, können eingesetzt werden, um Zeitreihendaten wie Sterberaten, Versicherungskosten und Schadensverläufe zu analysieren. Ein Neural Network kann lernen, zukünftige Ereignisse wie Schadenshöhen oder Lebenserwartung besser vorherzusagen als einfache stochastische Modelle, indem es verborgene Muster in den historischen Daten erkennt, die von traditionellen Modellen möglicherweise übersehen werden.

Simulationsbasierte Modellvalidierung

Deep Learning-Modelle können verwendet werden, um simulierte Daten zu erzeugen, die dann als Testdaten für das stochastische Modell verwendet werden. Diese Simulationen können auf Basis historischer Daten trainiert werden und sollen die verschiedenen Unsicherheitsaspekte und -szenarien widerspiegeln, die im stochastischen Modell enthalten sind. Mit einem Deep Learning-Modell lassen sich auch unrealistische oder seltene Szenarien (Outlier-Daten) erzeugen, um zu prüfen, ob das stochastische Modell robust genug ist, um solche Szenarien korrekt zu behandeln.

Unsicherheitsquantifizierung

KI-Methoden wie probabilistische neuronale Netze (Bayesian Networks) können genutzt werden, um die Unsicherheiten der Modellvorhersagen zu quantifizieren und zu validieren.

Ensemble-Methoden

Deep Learning-Modelle können auch im Rahmen von Ensemble-Techniken eingesetzt werden. Hierbei werden mehrere Modelle trainiert und die Ergebnisse kombiniert, um die Validität eines stochastischen Modells zu testen. Die Idee ist, dass ein Ensemble-Modell durch die Kombination der Stärken verschiedener Deep Learning-Architekturen zu einer robusteren Vorhersage führt. Beispielweise könnten ein neuronales Netzwerk, ein Random Forest und ein Support Vector Machine-Modell gemeinsam verwendet werden, um die Konsistenz der Vorhersagen des stochastischen Modells zu überprüfen.

Wichtig bei der Validierung ist es, nicht rein mechanisch vorzugehen, sondern eine nützliche und realitätsnahe Interpretation der Ergebnisse anzustreben. Bestimmte Ansätze von Explainable-AI-Verfahren versuchen die Modellentscheidungen und die Modellarchitektur besser nachvollziehbar zu machen und können hierfür eingesetzt werden.

Die Validierung stochastischer Unternehmensmodelle durch KI erfordert eine Kombination aus traditioneller Modellvalidierung und modernen maschinellen Lerntechniken. Es ist ein iterativer Prozess, der eine enge Zusammenarbeit zwischen mathematischen Modellierern und KI-Experten erfordert, um die Präzision und Relevanz der Ergebnisse sicherzustellen und die Unsicherheiten in den Modellen zu minimieren¹⁶.

Vorteile bei der Verwendung von KI in der Validierung

- Automatisierung und Effizienz: KI kann große Datenmengen schnell und zuverlässig prüfen. Das reduziert den manuellen Aufwand und spart Zeit und Kosten.
- Skalierbarkeit: KI-Modelle lassen sich leicht auf größere Datenmengen oder komplexere Systeme anwenden.
- Konsistenz: Im Gegensatz zu menschlichen Prüfern macht KI keine Flüchtigkeitsfehler und liefert konsistente Ergebnisse.
- Früherkennung von Anomalien: KI kann Muster erkennen, die für Menschen schwer zu identifizieren sind, und frühzeitig auf Fehler oder Ausreißer hinweisen.
- Lernfähigkeit: Durch maschinelles Lernen kann sich die Validierung im Laufe der Zeit verbessern, indem sie aus neuen Daten lernt.

Nachteile und Herausforderungen bei der Verwendung von KI in der Validierung

- Halluzinationen (insb. bei generativer KI/LLMs): Teilweise erzeugen generative Modelle Inhalte bzw. Begründungen, die plausibel wirken, aber sachlich falsch sind oder nicht zum Kontext passen.
- Erklärbarkeit und Transparenz: Viele KI-Modelle (z. B. Deep Learning) sind schwer zu interpretieren (Black Box), was die Nachvollziehbarkeit der Validierung erschwert.
- Datenabhängigkeit: Die Qualität der Validierung hängt stark von der Qualität und Repräsentativität der Trainingsdaten ab.
- Bias und Diskriminierung: Wenn die Trainingsdaten verzerrt sind, kann KI falsche oder diskriminierende Entscheidungen treffen.
- Kosten und Komplexität der Implementierung: Die Entwicklung und Integration von KI-Systemen kann teuer und technisch anspruchsvoll sein.
- Overfitting: Ein KI-Modell kann zu stark auf Trainingsdaten angepasst sein und dadurch bei neuen Daten versagen.

¹⁶ DAV, Ergebnisbericht der Ausschüsse Rechnungslegung und Regulierung und Actuarial Data Science „Regulierung und Validierung von KI-Modellen“ vom 26. Februar 2024

- Regulatorische und ethische Fragen: In sensiblen Bereichen (z. B. Medizin, Justiz) ist der Einsatz von KI zur Validierung rechtlich und ethisch umstritten.

7. Dokumentation sowie Berichtswesen

7.1. Einleitung

Eine klare und umfassende Dokumentation ist essenziell für die Nachvollziehbarkeit und Transparenz von Prozessen in der Lebensversicherung. Der Einsatz von KI kann hierbei maßgeblich unterstützen, sowohl bei der initialen Erstellung als auch bei der kontinuierlichen Pflege von Dokumentationen.

Versicherungsunternehmen müssen eine Vielzahl an Berichten auf Einzelversicherungsebene als auch Gruppenebene für verschiedene Stakeholder und im Rahmen diverser Rechnungslegungsstandards und aufsichtsrechtlicher Anforderungen erstellen. Die Berichte müssen regelmäßig erstellt und qualitätsgesichert werden. Zudem muss sichergestellt werden, dass die Berichte untereinander konsistent sind.

7.2. Dokumentation und Berichtswesen – bisherige Erfahrungen

Dokumentationen und Berichte werden überwiegend manuell unter Nutzung von Tools zur Kollaboration erstellt. Daten werden manuell übertragen und geprüft. Qualitätssicherung, Kontrollen und Konsistenzchecks erfolgen häufig manuell. Daher verursachen Dokumentationsanforderungen und Berichtswesen hohe Aufwände und Kosten.

7.3. Erstellung von Dokumentationsentwürfen durch KI

Generative KI, insbesondere in Form von LLMs, bietet die Möglichkeit, erste Entwürfe für Dokumentationen effizient zu erstellen. Diese Technologie nutzt vorhandene Daten und Dokumente, um kohärente und strukturierte Inhalte zu generieren, die anschließend von Experten überprüft und verfeinert werden können.

- Erstellung von Dokumentationsentwürfen: Basierend auf technischen Spezifikationen, Modelländerungen oder Prozessbeschreibungen erstellt die KI automatisch Vorlagen für neue Dokumentationen. Diese Entwürfe enthalten standardisierte Abschnitte, wie etwa Zielsetzungen, methodische Ansätze und technische Details.
- Verwendung von Output-Schemata: LLMs können an vordefinierte Templates angepasst werden, sodass die generierten Dokumentationen ein einheitliches Format und Layout einhalten. Diese Output-Schemata gewährleisten, dass alle wichtigen Abschnitte enthalten sind und die Inhalte in der gewünschten Reihenfolge präsentiert werden.
- Konsistenz- und Aktualitätsprüfung: Zusätzlich kann KI bestehende Dokumentationen durch einen Quercheck auf Inkonsistenzen oder veraltete Informationen analysieren. Dadurch wird sichergestellt, dass alle relevanten Inhalte aufeinander abgestimmt und aktuell sind.
- Integration von KI-Agenten: KI-Agenten können den gesamten Prozess überwachen, indem sie automatisch Informationen aus verschiedenen Quellen extrahieren und in die Entwürfe integrieren. Dies spart nicht nur Zeit, sondern minimiert auch die Gefahr menschlicher Fehler.

Die Effizienz bei der Dokumentationserstellung wird durch automatisierte Entwürfe gesteigert, während die Konsistenz und Qualität der Inhalte gewährleistet bleiben. Dies ermöglicht eine beschleunigte Bearbeitung ohne Abstriche bei der Genauigkeit.

7.4. Automatisiertes Aktualisieren von Dokumentationen und Berichten

Die regelmäßige Aktualisierung von Dokumentationen (z. B. bei häufigen Modelländerungen) sowie Berichten stellt eine zeitintensive Aufgabe dar. KI kann diesen Prozess durch automatisierte Mechanismen erheblich erleichtern.

- Regelmäßige Überprüfung und Anpassung: KI-Systeme können kontinuierlich Änderungen in Modellen, Parametern oder Prozessen überwachen und automatisch die relevanten Abschnitte der Dokumentation aktualisieren. Beispielsweise könnten bei einer Änderung von Annahmen oder Modellparametern die entsprechenden Kapitel automatisch angepasst werden.
- Einsatz von KI-Agenten für Automatisierung: Agenten können nicht nur Aktualisierungen vornehmen, sondern auch konsistente Verknüpfungen zwischen verschiedenen Dokumenten sicherstellen. Beispielsweise könnten Verweise auf ältere Versionen durch die neuesten Informationen ersetzt werden, ohne dass manuelle Eingriffe erforderlich sind.
- Integration von Überwachungsfunktionen: Die Agenten können Warnungen ausgeben, wenn eine Diskrepanz zwischen den aktuellen Daten und der bestehenden Dokumentation erkannt wird. Dies unterstützt die kontinuierliche Qualitätssicherung.

Durch den automatisierten Aktualisierungsprozess wird eine erhebliche Zeitersparnis erzielt. Gleichzeitig wird die Konsistenz und Vollständigkeit der Dokumentation gewährleistet, was zu einer höheren Verlässlichkeit und Nachvollziehbarkeit führt.

7.5. Automatisierte und KI-gestützte Berichtserstattung unter Solvency II

Die Integration von KI in den Reporting-Prozess bietet Möglichkeiten zur Effizienzsteigerung und Qualitätsverbesserung. Insbesondere der Einsatz von GenAI kann im Reporting-Prozess sinnvoll sein.

Der Einsatz von GenAI kann bei der Erstellung, Weiterverarbeitung und Auswertung von Berichten unterstützen und den Zeitaufwand für die Bearbeitung reduzieren:

- Automatisierte Berichtserstellung: KI kann Berichte automatisch generieren, indem sie relevante Daten zusammenführt (u.a. quantitative).
- Unterstützung bei der Berichtserstellung: GenAI kann Textaktualisierungen basierend auf vorhandenen Vorunterlagen vornehmen, Daten analysieren und aufbereiten sowie diese in den Berichtstext integrieren.
- Aggregierte Berichterstattung: KI kann Berichte auf Gruppenebene erstellen, indem sie Informationen aus den Berichten einzelner Versicherungsunternehmen einer Gruppe zusammenführt.
- Analyse von Berichten: KI ermöglicht Zeitreihenanalysen, Vergleiche von Berichtsformaten über verschiedene Unternehmen oder über verschiedene Berichte eines Unternehmens.
- Weiterverarbeitung von Berichten: GenAI kann bei der Erstellung von Präsentationen helfen, Maßnahmen ableiten und die Nachverfolgung unterstützen.
- Zusammenfassung: Die Management-Summary des Berichtes kann automatisiert durch KI erstellt werden und die wesentlichen Aussagen des Berichtes können je nach Stakeholder herausgestellt werden.
- Analysen für die Berichte: KI kann auch genutzt werden, um Daten zu visualisieren und auszuwerten. So kann KI beispielsweise auch unterstützen Trends abzuleiten und Veränderungen zu analysieren.
- Übersetzung für Gruppenfunktion: Bei multi-nationalen Unternehmen kann KI bei der Übersetzung von Berichten hilfreich sein.

Der Einsatz von KI kann auch zu einer Qualitätsverbesserung führen:

- Fehlerreduktion: Manuelle Übertragungsfehler können durch automatisierte Prozesse minimiert werden.
- Konsistenzsicherung: KI kann die Konsistenz zwischen verschiedenen Berichtsformaten und innerhalb eines Konzerns gewährleisten.
- Erhöhung der Lesbarkeit und Verständlichkeit: KI kann Texte so optimieren, dass sie klarer und verständlicher werden.
- Überprüfung von Anforderungen: KI kann überprüfen, ob alle regulatorischen und internen Anforderungen, wie Arbeitsanweisungen, Normen und Kontrollen an den jeweiligen Bericht

eingehalten wurden. Es können automatisiert Checklisten aller Anforderungen erstellt werden und deren Erfüllung dokumentiert werden.

Der Einsatz von KI im Reporting-Prozess erfordert einige Voraussetzungen, um diese auch effizient und sicher einsetzen zu können.

- **Datenintegration:**
 - Die GenAI Anwendung muss zusätzlich zu den Trainingsdaten weitere Daten einbeziehen, wie beispielsweise Berechnungsergebnisse, Analysen, interne Berichte, regulatorische Anforderungen und interne Anforderungen an die Berichtserstellung.
 - Die GenAI Anwendung muss darauf beschränkt werden, ausschließlich die vorgegebenen Daten zu verwenden, um ein Halluzinieren (d.h. das Erfinden oder Generieren von inhaltlich falschen oder nicht belegbaren Informationen) zu vermeiden.
- **Datensicherheit:**
 - GenAI Anwendungen, die für die Verarbeitung vertraulicher Daten verwendet werden, müssen so konzipiert sein, dass Datenschutzverstöße vermieden werden (eigene Implementierung im Unternehmen, interne Daten dürfen nicht für Trainingszwecke an externe Unternehmen übermittelt werden, EU DSGVO, Schutz sensibler Daten).
 - Datenverfügbarkeit: Daten müssen aufbereitet vorliegen, um von GenAI Anwendung verwendet werden zu können (Anforderungen an Datenhaltung im Unternehmen – Datenqualität, Verfügbarkeit von Daten, Datenaufbereitung, Datenhistorie, Datenintegrität).
- Erklärbarkeit / Transparenz
- Die Ausgaben der KI müssen nachvollziehbar und erklärbar sein
- Regulatorische Anforderungen: EU AI Act und weitere regulatorische Anforderungen (DSGVO, DORA, SII, ...) sind bei Einführung eines KI-Systems zur Unterstützung im Berichtsprozess zu beachten.

Vorteile automatisierter und KI-gestützte Berichtserstattung

- **Produktivität:** Effizienzsteigerung und Kostenreduktion bei Erstellung, Aktualisierung, Weiterverarbeitung und Auswertung.
- **Qualitätsverbesserung:** Fehlerreduktion, Konsistenzprüfungen, Erhöhungen der Lesbarkeit und Abgleich zu internen und externen Anforderungen an die Berichtserstellung.

Nachteile und Herausforderungen automatisierter und KI-gestützte Berichtserstattung

- **Datenhaltung:** Hohe Anforderungen an Datenhaltung, um GenAI nutzen zu können (Datenintegration in KI-Tool, Datensicherheit, Datenverfügbarkeit).
- **Fehleranfälligkeit:** KI-Anwendungen können halluzinieren und Daten und Texte frei erfinden oder verfälschen. Die Berichte müssen jedoch vollständig und richtig sein.
- **Technische Einschränkungen:** Bisherige GenAI Anwendungen sind oft hinsichtlich der Ausgabengröße oder des Ausgabeformats beschränkt: Daher können Berichte bisher nur abschnittsweise erstellt werden können.

8. Fazit und Ausblick

Das Papier zeigt: KI stiftet entlang der gesamten Wertschöpfung der Risikomodellierung spürbaren Nutzen – von der Datenaufbereitung über Annahmen und Modellierung bis hin zu Validierung und Berichtswesen. Insbesondere dort, wo heute noch viel manuell geprüft, dokumentiert und konsistent gehalten werden muss, kann KI-Struktur, Tempo und Nachvollziehbarkeit erhöhen.

Damit diese Wirkung nicht in Einzelprojekten verpufft, braucht es ein gemeinsames technisches Fundament statt mehrerer Insellösungen: eine unternehmensweite Basis, auf der Use Cases standardisiert angebunden, betrieben und weiterentwickelt werden können. Kerngedanke ist eine ska-

lierbare Architektur mit klaren Schnittstellen, wiederverwendbaren Bausteinen und zentral gepflegtem Wissen – so werden neue Anwendungsfälle nicht jedes Mal neu erfunden, sondern „ein-gesteckt“ und hochskaliert.

Ausblick: Priorität hat der Aufbau dieser Plattform mit Governance-by-Design (Rollen, Freigaben, Audit-Trails) und einem fokussierten Pfad für die Weiterentwicklung von Use Cases entlang der Prozesse. Auf dieser Basis lassen sich schnelle, risikoarme Einstiege (z. B. agentenbasierte Assistenz) realisieren und schrittweise in die Fläche bringen – strategisch eingebettet, sicher betrieben und fortlaufend erweitert. So wird aus punktuellen Effizienzgewinnen ein skalierbarer, prüfbarer KI-Betrieb.

Anhang

8.1. Abkürzungen

ALM	Asset-Liability Management
ARIMA	Autoregressive Integrated Moving Average
BEL	Besten Schätzwert der Versicherungstechnischen Rückstellungen
BERT	Bidirectional Encoder Representations from Transformers
BMI	Body-Mass-Index
BSM	Branchensimulationsmodells
CF Cashflow	Cashflow
DAV	Deutschen Aktuarvereinigung e. V.
DBSCAN	Density-Based Spatial Clustering of Applications with Noise
DL	Deep Learning
DORA	Digital Operational Resilience Act
DSGVO	Datenschutz Grundverordnung
ESG	Economic Scenario Generator / Ökonomischer Szenariengenerator
GBM	Gradient Boosting Machines
GenAI	Generative AI
GLM	Generalized Linear Model
GPT	Generative Pre-Trained Transformer
Graph-RAG	Retrieval-Augmented Generation with Graphs
IDV	Individuelle Datenverarbeitung
IF	Isolation Forest
IFRS	International Financial Reporting Standards
IKS	Internes Kontrollsystem

KI/AI	Künstliche Intelligenz / Artificial Intelligence
KPI	Key Performance Indicator
KRI	Key Risk Indicator
LLaMA	Large Language Model Meta AI
LLM	Large Language Model
LSMC	Least-Squares Monte Carlo
LSTM	Long Short-Term Memory
MaGo	Mindestanforderungen an die Geschäftsorganisation von Versicherungsunternehmen
MindZV	Mindestzuführungsverordnung
ML	Machine Learning
NLP	Natural Language Processing
NN	Neural Network
O&G	Optionen und Garantien
PCA	Principal Component Analysis
QRT	Quantitative Reporting Templates
RAG	Retrieval-Augmented Generation
RF	Random Forest
RfB	Rückstellung für Beitragsrückerstattung
RM	Reasoning-Modelle
RNN	Rekurrente Neuronale Netze
RPA	Robotic-Process-Automation
RSR	Regular Supervisory Report
SCR	Solvency Capital Requirement / Solvenzkapitalanforderung
SFCR	Solvency and Financial Condition Report
SVM	Support Vector Machines
t-SNE	t-Distributed Stochastic Neighbour Embedding

UMAP	Uniform Manifold Approximation and Projection
VmF	Versicherungsmathematische Funktion
VN	Versicherungsnehmer
XAI	Explainable AI

8.2. Literaturverzeichnis

1. **Verordnung (EU) 2024/1689 des Europäischen Parlaments und des Rates vom 13. Juni 2024** über harmonisierte Vorschriften für künstliche Intelligenz (AI-Act). Amtsblatt der Europäischen Union, L 2024/1689 vom 13.06.2024.
Verfügbar unter: [L_202401689DE.000101.fmx.xml](https://eur-lex.europa.eu/eli/reg/2024/1689/oj)
2. **Delegierte Verordnung (EU) 2015/35 der Kommission vom 10. Oktober 2014** zur Ergänzung der Richtlinie 2009/138/EG des Europäischen Parlaments und des Rates über die Aufnahme und Ausübung der Versicherungs- und Rückversicherungstätigkeit (Solvabilität II), insbesondere Artikel 241–264.
Amtsblatt der Europäischen Union, L 12 vom 17. Januar 2015.
3. **Deutsche Aktuarvereinigung e. V. (DAV): Der Artificial Intelligence Act im aktuariellen Kontext.**
Ergebnisbericht, 11. März 2025.
Verfügbar unter: https://aktuar.de/content/PDF/Fachwissen/2025-03-11_DAV_EB_AI_Act_im_aktuariellen_Kontext.pdf
4. **Deutsche Aktuarvereinigung e. V. (DAV): Anforderungen an einen ökonomischen Szenariengenerator.**
Fachgrundsatz, 27. November 2023.
Verfügbar unter: https://aktuar.de/content/PDF/Fachwissen/2023-11-27_DAV-Hinweis_Oekonomischer_Szenariengenerator.pdf
5. **Deutsche Aktuarvereinigung e. V. (DAV): Best Estimate in der Lebensversicherung.**
Fachgrundsatz, Köln, 27. Juni 2019.
Verfügbar unter: https://aktuar.de/content/PDF/Fachwissen/2019-06-27_DAV-Hinweis_Best_Estimate_Lebensversicherung.pdf
6. **Deutsche Aktuarvereinigung e. V. (DAV): Explainable Artificial Intelligence: Ein aktueller Überblick für Aktuarinnen und Aktuare.**
Ergebnisbericht, 27. Mai 2024.
Verfügbar unter: https://aktuar.de/content/PDF/Fachwissen/2024-05-27_DAV-Ergebnisbericht_Explainable_Artificial_Intelligence.pdf
7. **Deutsche Aktuarvereinigung e. V. (DAV): Proxy-Modelle für die Risikokapitalberechnung.**
Ergebnisbericht, 08. Juli 2025.
Verfügbar unter: https://aktuar.de/content/PDF/Fachwissen/2015-07-08_DAV_Ergebnisbericht_AG_Aggregation.pdf
8. **Deutsche Aktuarvereinigung e. V. (DAV): Regulierung und Validierung von KI-Modellen.**
Ergebnisbericht, 26. Februar 2024.
Verfügbar unter: https://aktuar.de/content/PDF/Fachwissen/2024-02-26_DAV-Ergebnisbericht_Regulierung_und_Validierung_von_KI-Modellen.pdf
9. **Deutsche Aktuarvereinigung e. V. (DAV): Validierung der versicherungstechnischen Rückstellungen unter Solvabilität II.**
Ergebnisbericht, 16. November 2018.
Verfügbar unter: https://aktuar.de/content/PDF/Fachwissen/2018-11-16_DAV_Ergebnisbericht_Validierung_vt_Rueckstellungen.pdf
10. **Deutsche Aktuarvereinigung e. V. (DAV): Vermeidung von Bias und Diskriminierung im Bereich Data Science.**
Ergebnisbericht, 11. März 2025.
Verfügbar unter: https://aktuar.de/content/PDF/Fachwissen/2025-03-11_DAV_EB_Bias-Diskriminierung.pdf

Alle online verfügbaren Quellen wurden zuletzt geprüft am 12.01.2026.