



Rechengenauigkeit vs. fundierte Entscheidungsunterstützung: Grenzen generativer KI im Controlling

Christoph Eisl / Werner Gleißner

RMA



Summary

Generative KI-Systeme wie ChatGPT lösen betriebswirtschaftliche Rechenaufgaben formal korrekt – etwa Break-even-Analysen –, erzeugen dabei aber eine gefährliche Scheingenauigkeit. Sie übernehmen implizit vereinfachende Annahmen (z. B. Sicherheit aller Plangrößen) und ignorieren reale Unsicherheiten und Risiken. Rechengenauigkeit wird so fälschlich mit Entscheidungssicherheit gleichgesetzt. Für den professionellen Einsatz im Controlling braucht es daher hybride Lösungen: KI, spezialisierte Software und menschliche Experten, die Ergebnisse kritisch prüfen und risikoorientiert interpretieren.

Von einfachen KI-Demos zu realen Controlling-Aufgaben

Der Umgang mit wichtigen betriebswirtschaftlichen Aufgabenstellungen, z.B. bei Strategieentwicklung, Unternehmensplanung oder Risikoanalyse, erfordert meist hohe fachliche Kompetenz und damit den Einsatz von Spezialisten, insbesondere Controllern. Zudem erscheint es naheliegend, über die Nutzung von generativer künstlicher Intelligenz (KI) nachzudenken, weil schon die unmittelbar am Markt verfügbaren Tools, wie Gemini, ChatGPT oder Perplexity, ein breites Wissen auch in betriebswirtschaftlichen Themen aufweisen.¹ Die Anwendung generativer KI-Modelle (Large Language Models, LLM) auf klassische betriebswirtschaftliche Aufgaben erzeugt häufig den Eindruck hoher Präzision und Eindeutigkeit. Dieser Eindruck entsteht unter anderem, weil sol-

che Aufgaben in Lehrbüchern und Übungsformaten bewusst als deterministische Rechenprobleme formuliert werden.² In den letzten Ausgaben des Controller Magazins wurden mehrere derartige Beispiele zum Einsatz von GenAI-Tools vorgestellt. Die Zielsetzung lag darin, aufzuzeigen, dass LLM bereits heute einfach gehaltenen Rechenaufgaben im Controlling korrekt und transparent lösen und erläutern können. Zudem sollte die Leserschaft ermutigt werden, diese Beispiele selbst auszuprobieren und damit Erfahrung im Umgang mit GenAI-Tools zu sammeln. Ergänzend zu den in diesen Veröffentlichungen bereits aufgezeigten Potenzialen werden hier nun die zentralen Gefahren und Limitationen einer unreflektierten Nutzung vertiefend analysiert. Zur Illustration der diskutierten Problembereiche wird erneut die Aufgabenstellung der Break-even-Analyse aus Leitner-Hanetseder et al. (2026) aufgegriffen.

Ausgangspunkt: Eine scheinbar eindeutige betriebswirtschaftliche Standardaufgabe

Break-even-Analysen gehören zu den klassischen Instrumenten des Controllings. Sie dienen der Ermittlung jener Absatzmenge oder jenes Umsatzes, ab dem ein Unternehmen seine variablen und fixen Kosten deckt und beginnt, Gewinne zu erzielen. In Ausbildung, Praxis und Softwareanwendungen werden sie häufig als deterministische Rechenaufgaben behandelt, bei denen aus gegebenen Kosten- und Preisparametern ein eindeutiger Schwellenwert abgeleitet wird.

Im hier betrachteten Beispiel³ wird diese Methodik konsequent angewendet. Für eine Automotive-Komponente liegen folgende Planannahmen vor:

- Materialpreis: 15,00 € pro Kilogramm
- Materialverbrauch: 0,5 Kilogramm je Einheit
- Fertigungszeit: 3,0 Minuten je Einheit
- Variabler Verrechnungssatz der Kostenstelle Fertigung: 90 € pro Stunde
- Geplante Fixkosten: 2.000.000 €
- Zielvorgabe: Deckungsbeitrags-Umsatz-Quote (DBU-Quote) von 25 %

Auf dieser Basis wird der Verkaufspreis so festgelegt, dass ein Viertel des Umsatzes als Deckungsbeitrag zur Fixkostendeckung und zum Gewinn beiträgt.

Die Break-even-Rechnung der KI (ChatGPT): formal korrekt und nachvollziehbar

- „Zunächst werden durch die KI die variablen Stückkosten ermittelt.⁴ Die variablen Materialkosten betragen: $15,00 \text{ €} \times 0,5 \text{ kg} = 7,50 \text{ € je Einheit}$.
- Die variablen Fertigungskosten ergeben sich aus der Bearbeitungszeit von 3 Minuten, was 0,05 Stunden entspricht: $0,05 \text{ h} \times 90 \text{ €} = 4,50 \text{ € je Einheit}$.
- Damit ergeben sich gesamte variable Stückkosten von: $v = 7,50 \text{ €} + 4,50 \text{ €} = 12,00 \text{ € je Einheit}$.

- Die Ziel-DBU-Quote von 25 % wird definiert als Verhältnis von Deckungsbeitrag zum Umsatz. Formal gilt damit: $(DB / \text{Umsatz}) = (p - v) / p = 0,25$.
- Durch Umformung ergibt sich ein Verkaufspreis von: $p = 16,00 \text{ € je Einheit}$.
- Der Deckungsbeitrag pro Stück beträgt folglich: $DB = 16,00 \text{ €} - 12,00 \text{ €} = 4,00 \text{ € je Einheit}$.
- Der Break-even-Point in Stück ergibt sich klassisch aus dem Quotienten von Fixkosten und Stückdeckungsbeitrag: $Q_{BE} = 2.000.000 \text{ €} / 4,00 \text{ €} = 500.000 \text{ Stück}$.
- Der zugehörige Break-even-Umsatz beträgt: $U_{BE} = 500.000 \times 16,00 \text{ €} = 8.000.000 \text{ €}$.

Die kritische Nachfrage des Nutzers: Sicherheit der Rechnung vs. Sicherheit der Entscheidung

Diese Rechnung ist in sich konsistent, vollständig nachvollziehbar und mathematisch exakt. Auf eine explizite Nachfrage hin erklärte ChatGPT das ermittelte Ergebnis – den Break-even-Umsatz – sogar als „zu 100 % sicher korrekt“. Erst die weiterführende Frage nach der Bedeutung dieser vermeintlichen „100-%-Sicherheit“ führte zu einer Relativierung durch ChatGPT. ChatGPT räumte ein, dass das Ergebnis gar nicht sicher sein kann – insbesondere dann nicht, wenn es – wie in der Controllingpraxis in den allermeisten Fällen – auf Plandaten beruht. Die mathematische Sicherheit der Rechnung impliziert stillschweigend, dass sämtliche zugrunde liegenden Annahmen ebenfalls sicher seien. Genau hier liegt ein zentraler Denkfehler, der nicht nur bei KI-Systemen, sondern auch in der klassischen Controlling-Praxis weit verbreitet ist.⁵

Rechengenauigkeit wird fälschlicherweise mit Entscheidungssicherheit verwechselt. In der Realität existieren aber keine sicheren Preise, keine sicheren Kosten, keine sicheren Mengen. Materialpreise unterliegen Markt- und Lieferantenrisiken, Materialverbräuche schwanken durch Ausschuss und Qualitätsabweichungen, Fertigungszeiten variieren aufgrund von Störungen, Personalverfügbarkeit oder Lernkurven. Absatzpreise und Absatzmengen sind Ergebnis von Nachfrage, Wettbewerb und Marktdynamik. Selbst Fixkosten sind häufig nur kurzfristig fix und reagieren mittelfristig auf Beschäftigungsänderungen. Damit sind defacto alle Eingangsgrößen der Break-even-Analyse unsicher. Der Break-even-Umsatz ist folglich keine feste Zahl, sondern eine Zufallsvariable.

Implizite Modellannahmen und ihre praktische Irrelevanz

Die klassische Break-even-Analyse beruht implizit auf einer Reihe stark vereinfachender Annahmen:

- linearer Zusammenhang zwischen Menge, Kosten und Erlös,
- keine Unsicherheit der zugrundeliegenden Parameter,
- keine Abhängigkeiten zwischen Einflussgrößen sowie
- keine zeitliche Dynamik.



Prof. Dr. Christoph Eisl

FH Oberösterreich, Fakultät für Wirtschaft und Management, Steyr. Professor für Controlling und Koordinator des Masterstudiengangs Controlling, Rechnungswesen und Finanzmanagement. christoph.eisl@fh-steyr.at



Prof. Dr. Werner Gleißner

lehrt an der TU Dresden (Betriebswirtschaft, insb. Risikomanagement) und ist Vorstand der FutureValue Group AG und der EACVA (European Association of Certified Valuators and Analysts). Er forscht zur Integration der Methoden aus Risikomanagement, Rating und Bewertungstheorie. w.gleissner@futurevalue.de

Diese Annahmen sind didaktisch hilfreich, aber ökonomisch realitätsfern, gerade wenn die Break-even-Analyse als Grundlage für unternehmerische Entscheidungen⁶ dienen soll. LLMs übernehmen diese Annahmen automatisch, weil sie ausschließlich auf Basis der explizit formulierten Aufgabenstellung arbeiten. Risiken, die im Prompt nicht genannt sind, werden nicht berücksichtigt – selbst dann nicht, wenn sie in der realen Unternehmenspraxis offensichtlich sind. Diese Überlegungen gelten natürlich für alle Aufgabenstellungen in der Betriebswirtschaftslehre (speziell im Controlling), bei denen Chancen und Gefahren (Risiken) eine Rolle spielen. Weil diese in der Praxis häufig nur unzureichend berücksichtigt werden, neigen auch KI-Systeme dazu, sie auszublenden. Dabei wird leicht übersehen, dass die verwendeten Planwerte oftmals nicht den entscheidungsrelevanten, erwartungstreuen Größen entsprechen.⁷

Noch allgemeiner gilt: Sobald ein betriebswirtschaftliches Problem nicht explizit als unsicher formuliert wird, trifft ein LLM automatisch eine Vielzahl impliziter Annahmen, die mit der Realität eines Unternehmens regelmäßig nicht vereinbar sind. Dem Modell fehlt das notwendige Hintergrund- und Spezialwissen, um betriebswirtschaftliche Konzepte auf einem für Unternehmensentscheidungen angemessenen Niveau anzuwenden. Dies unterstreicht die Notwendigkeit menschlicher Expertise – ohne dass dies die grundsätzliche Nützlichkeit von KI Systemen infrage stellt. Genau hier liegt ein strukturelles Problem beim Einsatz von KI (LLM) in betriebswirtschaftlichen Anwendungen: Die Modelle sind darauf optimiert, konsistente (sprachliche) Antworten auf Basis gegebener Informationen zu erzeugen – nicht jedoch darauf, implizite Unsicherheiten kritisch zu hinterfragen oder systematisch zu modellieren. Sie produzieren Ergebnisse, die für typische Nutzer plausibel erscheinen, aber keine Garantie für Wahrheit oder Entscheidungsrelevanz bieten. Und letztlich stellt sich die Frage: Warum sollten sie dies auch leisten, wenn ihre Antworten häufig keine unmittelbaren Entscheidungen auslösen?

Warum scheinbar korrekte KI-Ergebnisse in die Irre führen können

Für reale unternehmerische Entscheidungen ist nicht entscheidend, wo der Break-even-

Point unter idealisierten Annahmen liegt, sondern mit welcher Wahrscheinlichkeit dieser Punkt erreicht oder überschritten wird (und wo dessen Erwartungswert liegt).⁸ Eine deterministische Break-even-Zahl liefert hierfür keinerlei nützlichen Informationen und kann im Entscheidungsprozess sogar in die Irre führen. Im ungünstigsten Fall führt eine scheinbar exakte, tatsächlich, aber kontextfreie, Zahl zu systematischen Fehlentscheidungen, etwa zu:

- überoptimistischen Investitionsentscheidungen,
- falschen Preisuntergrenzen,
- unzureichenden Sicherheitsmargen sowie
- Fehleinschätzungen der Verlustwahrscheinlichkeit.

Die Gefahr liegt nicht in der Rechnung selbst – diese ist korrekt –, sondern in der Scheinsicherheit, die aus unrealistischen Annahmen entsteht. Ein vereinfachtes Modell wird nicht mit der Realität abgeglichen; wichtige Unsicherheiten bleiben außen vor, obwohl sie für Entscheidungen zentral sind. So entsteht ein trügerischer Eindruck von Präzision: Das Ergebnis wirkt eindeutig, ist jedoch inhaltlich unvollständig und kann zu falschen Schlussfolgerungen führen.

Die KI erzeugt Lösungen, die sie auf Basis ihrer Trainingsdaten für plausibel hält. Die Break-even-Lösung entspricht daher genau dem, was in einführenden Controlling-Lehrveranstaltungen üblich ist. Dieses Muster zeigt sich auch bei anderen betriebswirtschaftlichen Fragen: In der Praxis und im (einfachen) Schrifttum wird „Risiko“ noch immer häufig als erwarteter Schaden multipliziert mit der Eintrittswahrscheinlichkeit definiert – eine für professionelle Risikoquantifizierung unzureichende Vereinfachung.⁹ LLMs reproduzieren solche Muster, weil sie statistisch typische Formulierungen wiedergeben und nicht, weil sie deren methodische Qualität einschätzen können.¹⁰ Dadurch macht die KI gerade in anspruchsvolleren Bereichen – etwa bei der risikoadäquaten Bewertung von Handlungsoptionen – genau jene Fehler, vor denen Experten regelmäßig warnen.

Im vorliegenden Beispiel liegt das Problem nicht in einem Fehler im engeren Sinne, sondern in einer Reihe impliziter Annahmen, die das Modell – basierend auf Trainingsinformationen – trifft und die mit der Realität nicht übereinstimmen. Mangels gegenteili-

ger Informationen unterstellt die KI, dass alle in der Aufgabenstellung bereitgestellten Informationen, obwohl sie die Zukunft betreffen, sicher seien. Diese Annahme ist offenkundig falsch und realitätsfern. Hier zeigt sich das fehlende betriebswirtschaftliche Verständnis von LLMs: Sie erkennen typische Fehlerquellen nicht, können unrealistische Annahmen nicht identifizieren und verstehen nicht, dass die Aufgabenstellung lediglich ein stark vereinfachtes Abbild betrieblicher Realität ist. Und dieses Problem, das hier anhand der Break-even-Analyse erläutert wurde, ist grundsätzlich bei allen betriebswirtschaftlichen Aufgaben vorhanden. Das Ignorieren von Risiken ist dabei nur ein besonders sichtbares Beispiel – und es entsteht nicht zufällig: Auch viele Menschen ignorieren Risiken. LLMs replizieren diese menschliche Risikoblinkheit – und genau das macht sie gefährlich.¹¹ Denn wenn KI und Mensch denselben Denkfehler teilen, bleibt er häufig unentdeckt, zumal das Vertrauen in KI-Systeme in vielen Fällen hoch ist.

Konsequenzen für den Einsatz generativer KI im Controlling

Das Beispiel zeigt exemplarisch, dass generative KI zwar geeignet ist, strukturierte Rechenaufgaben zu lösen, dass ihre Ergebnisse jedoch ohne explizite Berücksichtigung von Risiken nicht entscheidungsfähig sind. Besonders kritisch ist dabei, dass die sprachliche Sicherheit und mathematische Präzision der Ausgabe eine Scheingenauigkeit erzeugen, die bestehende Unsicherheiten verdeckt, statt sie transparent zu machen.

Für den professionellen Einsatz von LLM im Controlling ergeben sich daraus mehrere zentrale Konsequenzen:

- Deterministische Ergebnisse dürfen nicht als Entscheidungsgrundlage verwendet werden.
- Risiken müssen explizit identifiziert, quantifiziert und aggregiert werden.
- Break-even-Analysen müssen probabilistisch interpretiert werden, etwa über Bandbreiten oder Wahrscheinlichkeitsverteilungen.
- KI-Systeme benötigen eine explizite methodische Einbettung.

Fazit und Implikationen

Die Diskussion um den Break-even-Point in diesem Beispiel macht deutlich, dass die

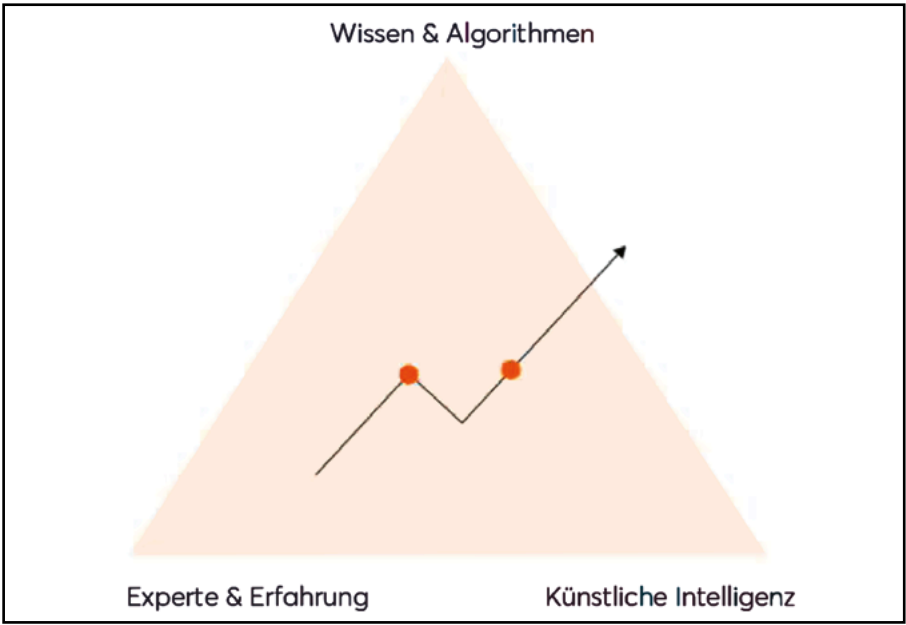


Abb. 1: Triangel (KI-generiert)

größte Gefahr beim Einsatz von LLM nicht in möglichen Rechenfehlern liegt, sondern in der unreflektierten Übertragung didaktischer Modelle auf reale Unternehmen und deren Entscheidungssituationen. Die formale Korrektheit einer KI-gestützten Rechnung ersetzt nicht das Verständnis der zugrunde liegenden ökonomischen Unsicherheiten. Für Controller, Risikomanager und Geschäftsleiter folgt daraus eine klare Schlussfolgerung: KI-gestützte Analysen können klassische Instrumente sinnvoll und effizient unterstützen, ersetzen jedoch keine risikoorien-

tierte Entscheidungslogik und auch nicht das fachliche Urteilsvermögen menschlicher Experten. Ohne die explizite Berücksichtigung relevanter Unsicherheiten bleibt jede scheinbar präzise Zahl eine potenzielle Quelle falscher Sicherheit – und da-mit ein Risiko für das Unternehmen.

In der Konsequenz werden sich bei anspruchsvollen Aufgaben, wie der Quantifizierung des Gesamtrisikoumfangs oder der Vorbereitung unternehmerischer Entscheidungen, „hybride“ Lösungen etablieren¹² Die-

se ergeben sich aus einer intelligenten Kombination (Abb. 1) von:

- a) KI-Systemen (insbesondere auf der Input- und Output-Seite), etwa bei Datenaufbereitung und Dateninterpretation,
- b) Software mit Algorithmen, die betriebswirtschaftliches Wissen abbilden (wie den Strategie Navigator¹³ mit Planungslogik, Monte-Carlo-Simulation und Rating-Modellen), sowie
- c) Experten, die die Anpassung und das Training der KI übernehmen und vor allem die Qualitätssicherung der Ergebnisse gewährleisten.

Das in Abb. 1 skizzierte Triangel-Modell ist dabei von zentraler Bedeutung. Bei realen betriebswirtschaftlichen Fragestellungen muss man davon ausgehen, dass KI systematische Fehler macht – nicht nur gelegentliche Rechenfehler, sondern vor allem Fehler in Bezug auf die ökonomische Realität. Die unzureichende Berücksichtigung von Risiken ist nur ein Beispiel. Problematisch ist insbesondere, dass KI häufig genau jene Denkfehler reproduziert, die auch typische Anwender (z. B. Controller) machen würden – und diese bleiben daher oft unentdeckt. Die Konsequenz ist klar: KI kann im Controlling sinnvoll eingesetzt werden, aber nur gemeinsam mit Fachexperten, die wissen, wo Menschen und KI-Systeme typischerweise irren, und die Ergebnisse kritisch prüfen und realitätsgerecht interpretieren können. ■

Literatur

Gleißner, W. (2020): Risikoblinkheit. Facetten, Ursachen, Auswirkungen und Gegenmaßnahmen, in: Zeitschrift für Risikomanagement (ZfRM), 1. Jg., Heft 1.20, S. 10-14
Gleißner, W. (2021): Unternehmerische Entscheidungen. Haftungsrisiken vermeiden (§ 93 AktG, Business Judgement Rule), in: Controller Magazin, Heft 1, Januar/Februar 2021, S. 16-23
Gleißner, W. (2022): Grundlagen des Risikomanagements. Handbuch für ein Management unter Unsicherheit, 4. Aufl., Vahlen Verlag München
Gleißner, W. (2023): Finanzwirtschaft und Risiko – Finanzierung, Kapitalkostenberechnung und Investitionsbewertung mit Methoden des Risikomanagements, in: BFuP, 75. Jg., Heft 5, S. 598-619
Gleißner, W. (2026): Künstliche Intelligenz (KI) im Risikomanagement, in: Controller Magazin, 51. Jg., Heft 2 (März/April 2026), S. 26-31
Gleißner, W./Rieg, R. (2026): KI-gestützte Prüfung von Unternehmensplanungen nach den GoP 3.0. Ein Anwendungskonzept für generative Sprachmodelle, in: Controlling, Themenheft 4/2026 „Künstliche Intelligenz im Controlling“ (erscheint in Kürze)
Gleißner, W./Rieg, R./Vanini, U. (2021): Risikoblinkheit im Controlling und wie man sie überwindet. Integration

von Risiko-Informationen in Controlling und Management, in: Controlling, Heft 5/2021, S. 66-72
Gleißner, W./Wolfrum, M. (2026): Der Einsatz Künstlicher Intelligenz für die Risikoidentifikation. Wie KI Aufsichtsräten hilft, Risiken schneller und strukturierter zu erkennen und zu priorisieren, in: Der Aufsichtsrat, Heft 2.2026 (Februar 2026), S. 32-35
Kamaras, E. (2025): Vortrag zum Thema AI in Corporate Finance Limits of AI in (quantitative) risk management and business decisions - Why an expert is still indispensable an der TU Dresden am 27.09.2025
Kaplan, R. S./Anderson, S. R. (2007): Time-Driven Activity-Based Costing. Boston
Leitner-Hanetseder, S./Eisl, Ch./Frenkenberger, S. (2026a): GenAI Chatbots im Controlling. Grundlagen und Praxistipps für den Einsatz von ChatGPT & Co, in: Controller Magazin, 51. Jg., Heft 1, S. 54-55
Leitner-Hanetseder, S./Eisl, Ch./Frenkenberger, S. (2026b): GenAI Chatbots im Controlling: Praxisbeispiel 1: Break Even Analyse, in: Controller Magazin, 51. Jg., Heft 1, S. 56-57
Rieg, R. (2015): Break-Even-Analyse im Mehrproduktfall unter Unsicherheit und Risiko, in: Controller Magazin, 40. Jg. (2015), Heft 4, S. 76-82
Rieg, R./Gleißner, W. (2022): Was ist ein erwartungstreuer Plan?, in: WPg, Heft 24.2022, S. 1407-1414

von Daacke, M./Gleißner, W./Eisl, Ch./Kleinhietaß, G. (2026): Die Welten der Unternehmensplanung: Managementplanung versus erwartungstreue Zukunftserfolgsplanung – worin besteht (vermeintlich) der Unterschied? (erscheint in Kürze)
Wood, D. (2025): Rewiring Your Mind for AI: Wie man im Zeitalter der Künstlichen Intelligenz denkt, arbeitet und erfolgreich ist, Technics Publications Sedona, Arizona

Fußnoten

1 Siehe z. B. Gleißner, 2026 zu Anwendungsmöglichkeiten im Risikomanagement und grundsätzlich bestehenden Limitierungen
2 Vgl. Leitner-Hanetseder et al., 2026a
3 Vgl. Leitner-Hanetseder et al., 2026b, S. 56-57
4 Nachfolgender Text direkt durch ChatGPT 5.2 generiert
5 Gleißner/Rieg/Vanini, 2021; 6 Gleißner, 2021
7 Siehe dazu von Daacke/Gleißner/Eisl/Kleinhietaß, 2026
8 Rieg/Gleißner, 2022; 9 Siehe Gleißner, 2022
10 Siehe dazu auch Wood, 2025
11 Siehe Gleißner, 2020 und speziell für das Controlling Gleißner/Rieg/Vanini, 2021
12 In enger Anlehnung an Gleißner, 2026
13 Kostenloser Download unter <https://futurevalue.de/leistungenprodukte/strategienavigator/>