



TUM

TECHNISCHE UNIVERSITÄT MÜNCHEN
INSTITUT FÜR INFORMATIK

AGENTONOMICS: Die KI-Agenten- Ökonomie

Alois C. Knoll und Hanns C. Suckfüll

TUM-I26124

Technische Universität München
Institut für Informatik

Technischer Bericht

AGENTONOMICS

Die KI-Agenten-Ökonomie

Handbuch zur Entwicklung und Steuerung
ökonomisch operierender KI-Agenten

Prof. Dr. Alois C. Knoll · Dr. Hanns C. Suckfüll

Lehrstuhl für Robotik, Künstliche Intelligenz und Echtzeitsysteme
TUM School of Computation, Information and Technology
Technische Universität München

1. Auflage 2026 · Wintersemester 2025/26

Vorwort der Autoren

Dieses Handbuch begleitet die im Wintersemester 2025/26 erstmals an der Technischen Universität München angebotene Vorlesung „KI-Agenten in der Betriebswirtschaftslehre – Agentic AI in Business Administration“. Es verbindet zwei Wissenstraditionen, die in der betriebswirtschaftlichen Lehre selten gemeinsam adressiert werden: die ganzheitlich-systemische Managementlehre, wie sie insbesondere durch das St. Galler Management-Modell geprägt wurde, und die Informatik der autonomen, lernfähigen Software-Agenten.

Der Anlass für diese Verbindung ist nicht akademischer Selbstzweck. Die wirtschaftliche Realität verändert sich derzeit in einer Geschwindigkeit, die das klassische Repertoire der Managementlehre überfordert. Generative Sprachmodelle und werkzeugfähige Agenten dringen in Tätigkeitsfelder vor, die bislang als genuin menschlich galten. Wissenarbeit, Entscheidungsvorbereitung, Verhandlungsführung, Kundenkommunikation – überall entstehen Systeme, die nicht mehr nur Werkzeuge sind, sondern als Akteure handeln. Damit verschiebt sich der Untersuchungsgegenstand der Betriebswirtschaftslehre selbst.

Wir schlagen mit AGENTONOMICS einen Bezugsrahmen vor, der KI-Agenten konsequent als Wirtschaftssubjekte denkt und sie in eine Managementarchitektur einbettet, die historisch aus dem St. Galler Konzept des integrierten Managements (Bleicher 2011) hervorgeht und um die spezifischen Erfordernisse polyzentrischer Agentennetzwerke (Suckfüll 1994) erweitert wird. Der zugrundeliegende Algorithmus – der AGENTONOMICS-Algorithmus – sowie seine erste Anwendung, der „Doz. Dr. AGENTONOMICS“, sind als Minimum Viable Product konzipiert: Sie sollen sich durch konkrete Anwendungsfälle iterativ verbessern.

Das Handbuch verfolgt einen doppelten Anspruch. Erstens soll es den Studenten ein präzises, in der etablierten Managementliteratur verankertes Begriffsinstrumentarium an die Hand geben. Zweitens soll es die Leserinnen und Leser am Ende in die Lage versetzen, selbst KI-Agenten zu konzipieren und technisch zu realisieren, die betriebswirtschaftliche Fragestellungen entlang des St. Galler Modells fundiert beantworten können. Aus diesem Grund schließt das Handbuch mit einer ausführlichen, praktisch umsetzbaren Anleitung zum Agentenbau (Kapitel 14).

Wir danken den Studenten, Kolleginnen und Kollegen, die mit kritischen Fragen zur Schärfung der Argumentation beigetragen haben, sowie den Projektpartnern aus Industrie und Wissenschaft, deren Praxiseinblicke die Fallbeispiele dieses Handbuchs geprägt haben.

München, im Wintersemester 2025/26

Alois C. Knoll und Hanns C. Suckfüll

Inhalt

Vorwort der Autoren	2
1 KI-Agenten in der Betriebswirtschaftslehre – Einführung und Motivation.....	5
1.1 Drei Zielrichtungen der Vorlesung.....	5
1.2 Zielgruppe und didaktischer Ansatz.....	6
1.3 Normative Meta-Fähigkeiten in einer KI-geprägten Arbeitswelt.....	7
1.4 Der AGENTONOMICS-Algorithmus als didaktisches Experiment.....	7
1.5 Aufbau des Handbuchs.....	8
2 Das St. Galler Management-Modell als theoretisches Fundament.....	10
2.1 Wurzeln: Die systemorientierte Unternehmungslehre Hans Ulrichs	10
2.2 Die zweite Generation: Bleichers Konzept Integrierten Managements	11
2.3 Die vierte Generation: Rüegg-Stürm und Grand.....	12
2.4 Komplementäre Bausteine: Vernetztes Denken, lernende Organisation, Kybernetik.....	14
2.5 Was bleibt, was erweitert werden muss	14
3 The Emergence of AGENTONOMICS – Historisch-ökonomische Einordnung.....	17
3.1 Drei Ökonomien: Information, Daten, Fähigkeiten	17
3.2 Symptome der Verschiebung in der realen Wirtschaft.....	19
3.3 Die Souveränitätsfrage: Von der Daten- zur Fähigkeitsouveränität.....	19
3.4 Von der Datentreuhand-Diskussion zu AGENTONOMICS.....	21
3.5 Polyzentrische KI vs. AGI – die strategische Abgrenzung	21
3.6 KI-Agenten als neue Gattung von Wirtschaftssubjekten	22
3.7 Wissenschaftstheoretische Einordnung von AGENTONOMICS	22
4 Das AGENTONOMICS Design & Management Reference Framework (ADMRF).....	25
4.1 Anspruch und theoretischer Hintergrund	25
4.2 Der Turing-Test der zweiten Generation als Zielmarke.....	25
4.3 Anwendungsfelder: Greenfield und Brownfield	26
4.4 Die sechs Module im Überblick.....	26
4.5 Wie die Module miteinander verbunden sind	28
4.6 Die AGENTONOMICS-Matrix als kondensiertes Steuerungswerkzeug	29
4.7 Erste Anwendung: „Doz. Dr. AGENTONOMICS“.....	30
5 Modul 1: Der AGENTONOMICS Definition-Frame	31
5.1 Funktion des Definition-Frame	31
5.2 Parameter 1: Autonomie und Souveränität.....	31
5.3 Parameter 2: Fähigkeits- und Kompetenzanspruch.....	33
5.4 Parameter 3: Regulierung und Selbstregulierung.....	33
5.5 Parameter 4: Identität, Ethik und Persönlichkeit.....	34
5.6 Parameter 5: KI-Modell und Embodiment.....	35
5.7 Parameter 6: Lernen und Lehren – Der Meta-Agent.....	36
5.8 Konsistenzprüfung und Boardperspektive	36
6 Modul 2: Die fünf Building-Blocks eines AGENTONOMICS-Agenten.....	38
6.1 Die fünf Bausteine im Überblick.....	38
6.2 Baustein 1: Der Agenten-Algorithmus.....	38
6.3 Baustein 2: Das Agenten-Interface.....	40
6.4 Baustein 3: Das KI-Modell.....	41
6.5 Baustein 4: Tools und Vektoren.....	42
6.6 Baustein 5: Der Ökosystem-Intermediär.....	43
6.7 Das Zusammenwirken der Building-Blocks im Lifecycle	44
7 Modul 3: Der Network-Intermediär	46
7.1 Warum es einen Intermediär braucht	46
7.2 Die Hausverwaltungs-Analogie	47
7.3 Funktionen des Intermediärs	47
7.4 Governance des Intermediärs	49
7.5 Deutscher und angelsächsischer Governance-Stil im Vergleich.....	50

8	Modul 4: Das AI-CEO Management Concept.....	52
8.1	Das St. Galler Management-Modell im Zeitalter der KI-Agenten.....	52
8.2	Die fünf Ordnungsmomente des AI-CEO	53
8.3	Was der AI-CEO konkret leistet	55
8.4	Die Ausbaustufen des AI-CEO	56
8.5	Anwendung: Wie spricht man mit dem AI-CEO?	57
8.6	Wann der AI-CEO „nein“ sagt.....	59
8.7	Der AI-CEO als Brücke zur PuK-Logik	59
9	Modul 5: Das PuK-Instrument	61
9.1	Das PuK-Instrument in der Tradition von Hahn und Hungenberg.....	61
9.2	Die AGENTONOMICS-Matrix als Kernartefakt	62
9.3	Bilanz und Gewinn-und-Verlust-Rechnung des Agenten	63
9.4	CAPEX und OPEX – das Investitions- und Betriebskostenprofil.....	64
9.5	Steuerungsrouitinen im Lifecycle.....	65
9.6	Die Verbindung zwischen CFO-Funktion und AI-CEO	66
10	Modul 6: Der Stakeholder Communicator	67
10.1	Stakeholder im St. Galler Verständnis	67
10.2	Die Pitch-Methodik nach Kawasaki.....	68
10.3	Vom PuK zum Pitch: Die operative Ableitung	68
10.4	Krisenkommunikation als Stresstest des Moduls	69
10.5	Innen- und Außenkommunikation – ein eigenes Gleichgewicht.....	70
11	Wachstumstheorie der KI-Agenten-Ökonomie	71
11.1	Der Kooperations-Multiplikator.....	71
11.2	Triadenbetrachtung: Drei Wachstumsformen	72
11.3	Polyzentrische Netzwerkorganisationen als Wachstumsträger	72
11.4	Innovation und Erneuerung im Netzwerk	73
12	Ökonomische Lösungsansätze: Adoleszenz der KI	75
12.1	Die Adoleszenz der KI	75
12.2	Wachstum von 20 Prozent: Quellen und Konsequenzen	75
12.3	Pessimistische Prognose: Sockelarbeitslosigkeit 20 Prozent – Quellen und Konsequenzen ...	76
12.4	Verteilungsfragen: Wer profitiert?	77
12.5	Lösungsansätze.....	77
12.6	Die spezifisch europäische Position	78
13	Souveränität: Von der Daten- zur Fähigkeitsouveränität	80
13.1	Der bizarre Status quo.....	80
13.2	Die nächste Stufe: Skill Sovereignty.....	81
13.3	Die Volkswagen-Analogie	82
13.4	Europäische, amerikanische und asiatische Ansätze im Vergleich.....	83
13.5	KI-Infrastruktur als Daseinsvorsorge	84
13.6	Der EU AI Act als Rahmen.....	85
14	Den AGENTONOMICS-Agenten bauen: Eine operative Anleitung.....	87
14.1	Die Systemarchitektur im Überblick.....	87
14.2	Aufbau der Wissensbasis: Die St.-Galler-Bibliothek.....	88
14.3	Der System-Prompt: Persönlichkeit und Werte	89
14.4	Beispieldialoge: Wie sieht eine Interaktion aus?.....	90
14.5	Evaluation: Wie messe ich, ob der Agent gut ist?.....	92
14.6	Ausbaustufen: Vom Tutor zum AI-CEO.....	93
14.7	Zehn praktische Empfehlungen für den Agentenbau	95
15	Glossar	96
	Nachwort.....	101

1 KI-Agenten in der Betriebswirtschaftslehre – Einführung und Motivation

Während die betriebswirtschaftliche Forschung und Lehre über Jahrzehnte hinweg primär Organisationen, Märkte und menschliche Akteure als ihre Gegenstände betrachtet hat, treten mit dem Aufkommen autonomer KI-Systeme neue Akteure auf den Plan: KI-Agenten, die nicht mehr nur assistierend, sondern zunehmend eigenständig, adaptiv und zielgerichtet handeln. Diese Verschiebung verlangt eine grundlegende Erweiterung der betriebswirtschaftlichen Begriffsapparate – nicht ihrer vollständigen Ablösung. Die Vorlesung „KI-Agenten in der Betriebswirtschaftslehre“ und das vorliegende Handbuch setzen genau an diesem Punkt an.

Beobachtbar ist, dass klassische Management-Modelle (Ulrich 1968; Bleicher 2011; Rüegg-Stürm und Grand 2020) zwar weiterhin tragfähige Grundlagen liefern, jedoch an Erklärungskraft verlieren, sobald autonome, lernfähige und ökonomisch agierende KI-Systeme in Wertschöpfungsprozesse integriert werden. Die zentrale Fragestellung des Handbuchs lautet daher:

Leitfrage

Wie lassen sich KI-Agenten betriebswirtschaftlich denken, gestalten und steuern, wenn sie als eigenständig handelnde Wirtschaftssubjekte auftreten – und welche Konsequenzen ergeben sich daraus für die Architektur ganzheitlicher Managementmodelle?

Diese Leitfrage bildet den Ausgangspunkt für das im Rahmen der Vorlesung entwickelte und im Folgenden Schritt für Schritt aufgebaute Management-Modell AGENTONOMICS. Der Name leitet sich aus den Begriffen Agent und Ökonomie ab und markiert den Anspruch, KI-Agenten nicht als isolierte technische Artefakte, sondern als Bestandteile einer eigenständigen ökonomischen Ordnung zu begreifen.

AGENTONOMICS erhebt den Anspruch, ein State-of-the-Art-Managementmodell zu sein, das mit Hilfe des AGENTONOMICS-Algorithmus KI technisch nutzt, um dem menschlichen Entscheider an Tiefe und Geschwindigkeit überlegene Management-Kompetenz zur Verfügung zu stellen. In einem ersten Schritt erfolgt die Einführung von AGENTONOMICS in Form einer Vorlesung vor Studenten der TUM, die sich – als Angestellte oder als Selbständige – auf das Berufsleben vorbereiten.

1.1 Drei Zielrichtungen der Vorlesung

Die Vorlesung verfolgt drei miteinander verknüpfte Zielrichtungen, die im Aufbau des Handbuchs gespiegelt werden:

Begriffliche und konzeptionelle Orientierung

Studenten sollen befähigt werden, KI-Agenten nicht ausschließlich als technische Artefakte, sondern als ökonomische Akteure zu begreifen. Dazu gehört die saubere Einordnung von AGENTONOMICS in den Kontext der Betriebswirtschaftslehre, der Volkswirtschaftslehre sowie angrenzender Wissenschaftsdisziplinen – insbesondere der Wirtschaftsinformatik, der Wirtschaftsethik und der Rechtswissenschaft (vgl. Picot et al. 2020; Mainzer 2019).

Anwendungs- und Gestaltungskompetenz

Aufbauend auf einem ganzheitlichen, betriebswirtschaftlichen Steuerungsmodell – dem AGENTONOMICS Design & Management Reference Framework (ADMRF) – lernen die Studenten, KI-Agenten systematisch zu beschreiben, zu parametrisieren und als Bestandteil von Geschäftsmodellen oder Organisationsformen zu konzipieren, technisch zu implementieren und wirtschaftlich erfolgreich zu steuern. Kapitel 4 entfaltet das ADMRF in seiner Gesamtarchitektur; die Kapitel 5 bis 10 vertiefen seine sechs Module.

Reflexion zukünftiger Wirtschaftsstrukturen

Die Vorlesung adressiert explizit Studenten, die sich auf den Eintritt in eine durch KI geprägte Arbeits- und Wirtschaftswelt vorbereiten – sei es als Mitarbeitende, Gründerinnen und Gründer oder als Gestalter neuer Organisations- und Marktformen. Der Begriff KI-Fähigkeitenwirtschaft, wie er von Knoll und Suckfüll (2024) erstmals in den wissenschaftlichen Diskurs eingeführt wurde, wird im Verlauf des Handbuchs vertieft und als mikroökonomisches Modell ausformuliert.

Übergeordnetes Ziel der Vorlesung ist es, eine Brücke zwischen herkömmlichen Management-Methodiken und innovativen KI-Entwicklungsmöglichkeiten zu schlagen. Dabei versteht sich die Vorlesung nicht als reine Wissensvermittlung, sondern immer auch als Reflexions- und Experimentierraum.

1.2 Zielgruppe und didaktischer Ansatz

Die Vorlesung richtet sich bewusst an eine interdisziplinäre Zielgruppe. Angesprochen sind insbesondere Studenten

- der Informatik,
- der Betriebswirtschaftslehre,
- der Volkswirtschaftslehre,
- angrenzender ingenieur- und sozialwissenschaftlicher Studiengänge.

Der didaktische Ansatz folgt einem doppelten Perspektivwechsel, um ganzheitliches Denken zu fördern. Für Studenten mit primär technischer Ausbildung eröffnet die Vorlesung den Zugang zu betriebswirtschaftlichen Denkweisen, Entscheidungslogiken und Steuerungsinstrumenten. Für Studenten mit betriebswirtschaftlichem Hintergrund wird KI nicht als Black Box behandelt, sondern als gestaltbares, parametrisches und ökonomisch wirksames System vorgestellt.

Mittelfristig positioniert sich die Vorlesung auch als Schulungsinstrument für die Führungsebene in Unternehmen sowie für andere gesellschaftlich relevante Stakeholdergruppen, die einen Beitrag zur erfolgreichen Transformation hin zu einer KI-geprägten Gesellschaft – der KI-Fähigkeitenwirtschaft – leisten wollen. Langfristig dient die Vorlesung dazu, den AGENTONOMICS-Algorithmus auch als Instrument für die praktische Unternehmensführung weiterzuentwickeln und ihn schrittweise anhand von Praxisbeispielen zu verbessern.

1.3 Normative Meta-Fähigkeiten in einer KI-geprägten Arbeitswelt

Der Vorlesung liegt die Einschätzung zugrunde, dass für ein erfolgreiches Bestehen in einer von KI geprägten Arbeitswelt nicht in erster Linie technische Spezialkenntnisse entscheidend sind, sondern übergeordnete, normative Meta-Fähigkeiten. Diese lassen sich – in Anlehnung an bildungstheoretische Überlegungen seit Humboldt und in Auseinandersetzung mit zeitgenössischen Diagnosen zur Wissensgesellschaft (Nida-Rümelin und Weidenfeld 2018) – in vier Bündeln zusammenfassen:

Kreativität, die sich in unternehmerischem Handeln reflektiert

Gemeint ist die Fähigkeit, das Mögliche jenseits des Bestehenden zu denken und in tragfähige Handlungsformen zu überführen. Diese Kreativität ist nicht die romantische Eingebung des Künstlers, sondern die nüchterne Fähigkeit zur unternehmerischen Komposition vorhandener Ressourcen zu neuen Lösungen (Schumpeter 1934).

Resilienz und positive Weltvorstellung

In einer Welt rascher technischer Brüche ist die Fähigkeit, mit Unsicherheit produktiv umzugehen, eine ökonomische Schlüsselqualifikation. Resilienz wird hier nicht als individuelle Eigenschaft, sondern als Kompetenz verstanden, die durch konkrete Praktiken – etwa Szenariodenken, lernende Organisationen, kontrollierte Experimente – kultiviert werden kann (Senge 2008).

Ethisch-moralische Grundausrichtung

Die Vorlesung versteht die ethische Orientierung an einem Kantischen Imperativ nicht als beigeordnete Sonntagsrede, sondern als integralen Bestandteil ökonomischer Urteilskraft. Eine Ökonomie, die mit Agenten arbeitet, die im Namen ihrer Auftraggeber handeln, ist auf stabile normative Erwartungen angewiesen – andernfalls kollabieren ihre Transaktionskosten (Ulrich 2008; Kant 1785).

Umfassendes Welt- und Allgemeinwissen

Was früher schlicht Allgemeinbildung hieß, ist in einer von Spezialisierung und KI-gestützter Tiefenrecherche geprägten Welt wieder zur strategischen Ressource geworden. Wer die Tragweite einer KI-Ausgabe beurteilen will, muss in einem breiten Horizont historischer, philosophischer, naturwissenschaftlicher und ökonomischer Zusammenhänge denken können.

Die Vorlesung wird vor diesem Hintergrund situative, aktuelle Beispiele aus Praxis und multidisziplinären Wissensbereichen – auch über die Betriebswirtschaftslehre im engeren Sinne hinausgehend – ebenso einbinden wie tagespolitische Ereignisse.

1.4 Der AGENTONOMICS-Algorithmus als didaktisches Experiment

Ein zentrales Merkmal der Vorlesung ist ihr experimenteller Charakter. Künstliche Intelligenz wird nicht nur thematisiert, sondern aktiv in den Lehr- und Lernprozess integriert. Perspektivisch wird den Studenten ein personalisierbarer KI-Agent – der „Doz. Dr. AGENTONOMICS“ – an die Seite gestellt.

Er fungiert als Tutor, Coach und Sparringspartner und bildet zugleich die Basis für die Vernetzung der Studenten untereinander.

„Doz. Dr. AGENTONOMICS“ erklärt die sechs Module der Vorlesung in einem ganzheitlichen Kreislauf-Prozess: dem AGENTONOMICS Design & Management Reference Framework (ADMRF). Der Agent ist als erste Anwendung auf Basis des AGENTONOMICS-Algorithmus zu verstehen. Die Applikation befähigt Studenten und Praktikerinnen und Praktiker gleichermaßen, KI-Agenten systematisch und betriebswirtschaftlich fundiert zu denken, zu gestalten, umzusetzen und in der KI-Fähigkeitenwirtschaft erfolgreich zu steuern.

Der ADMRF besteht aus sechs Modulen, die verschiedene Perspektiven der Unternehmensführung – die des CEO, des CFO, des CTO, des CSO sowie des Boards als Ganzes – integrieren. Ein zentrales KI-Modell verbindet diese Perspektiven so, dass State-of-the-Art-Entscheidungshilfe bei der Entwicklung und dem Management von KI-Agenten entsteht. Das KI-gestützte ADMRF versteht sich dabei ausdrücklich als Weiterentwicklung klassischer Management-Bezugsrahmen. Historisch kann es als konsequente Fortschreibung des St. Galler Konzepts integrierten Managements (Bleicher 2011) sowie des Planungs- und Kontrollrechnungs-Modells PuK (Hahn und Hungenberg 2001) eingeordnet werden. Kapitel 2 erläutert diese Anschlüsse im Detail.

Insbesondere die Dreiteilung in eine normative, eine strategische und eine operative Perspektive aus dem Konzept integrierten Managements findet im Kontext von KI-Agenten-Netzwerkorganisationen Eingang und wird in den sechs Modulen des ADMRF gespiegelt. Konkret basieren die Steuerungsmechanismen polyzentrischer – also souverän interagierender KI-Agenten-Netzwerke – ganz wesentlich auf dem Organisationsverständnis des transnationalen Organisationsmodells (Suckfüll 1994), das im Anwendungsfeld international operierender Unternehmen entwickelt wurde. Auch die Ausgestaltung der Intermediär-Funktion zwischen den KI-Agenten und externen Stakeholdern profitiert hiervon. Die in Deutschland im Wesentlichen durch die Veröffentlichung Deutsche Daten-Treuhand (DEDATE; Fraunhofer MOEZ und Suckfüll 2018) angestoßene Diskussion über die Ausgestaltung einer das Datenteilen ermöglichenden Intermediär-Funktion findet Eingang in den AGENTONOMICS-Algorithmus.

1.5 Aufbau des Handbuchs

Das Handbuch ist so angelegt, dass die Leserinnen und Leser am Ende nicht nur das AGENTONOMICS-Modell verstanden haben, sondern selbst in die Lage versetzt werden, KI-Agenten zu bauen, die betriebswirtschaftliche Fragestellungen entlang des St. Galler Management-Modells fundiert beantworten können. Diese Doppelfunktion – Lehrbuch und Bauanleitung – spiegelt sich im Aufbau:

- Die Kapitel 1 bis 3 verorten AGENTONOMICS theoretisch. Sie führen in das Vorhaben ein, rekonstruieren das St. Galler Management-Modell in seinen zentralen Entwicklungsstufen und ordnen die Emergenz von AGENTONOMICS historisch-ökonomisch ein.
- Die Kapitel 4 bis 10 entfalten das AGENTONOMICS Design & Management Reference Framework (ADMRF) sowie seine sechs Module – vom Definition-Frame über die fünf Building-Blocks und den Netzwerk-Intermediär bis zum CEO-Konzept, der Planungs- und Kontrollrechnung und dem Stakeholdermanagement.

- Die Kapitel 11 bis 13 widmen sich Wachstumstheorie, ökonomischen Lösungsansätzen und der Souveränitätsfrage – insbesondere dem Übergang von der Daten- zur Fähigkeitsouveränität (Knoll und Suckfüll 2024).
- Kapitel 14 schließlich liefert die operative Bauanleitung: Wie baue ich einen AGENTONOMICS-konformen KI-Agenten, der St. Galler Fragestellungen beantwortet? Das Glossar in Kapitel 15 bündelt die wichtigsten Begriffe.

Jedes Kapitel schließt mit einem eigenen Literaturverzeichnis, auf das im Text referenziert wird. Damit ist sichergestellt, dass die einzelnen Kapitel auch isoliert lesbar bleiben.

Literaturverzeichnis

- Bleicher, K. (2011): Das Konzept Integriertes Management. Visionen – Missionen – Programme. 8., aktualisierte und erweiterte Auflage, Campus, Frankfurt am Main.
- Fraunhofer MOEZ; Suckfüll, H. (2018): Initiative zu einer Datentreuhand DEDATE. Diskussionspapier, Fraunhofer-Zentrum für Internationales Management und Wissensökonomie, Leipzig.
- Hahn, D.; Hungenberg, H. (2001): PuK – Wertorientierte Controllingkonzepte. 6., überarbeitete und erweiterte Auflage, Gabler, Wiesbaden.
- Kant, I. (1785): Grundlegung zur Metaphysik der Sitten. Riga.
- Knoll, A. C.; Suckfüll, H. (2024): From data sovereignty to skill sovereignty. Technischer Bericht TUM-I24112, Institut für Informatik, Technische Universität München.
- Mainzer, K. (2019): Künstliche Intelligenz – Wann übernehmen die Maschinen? 2. Auflage, Springer, Berlin und Heidelberg.
- Nida-Rümelin, J.; Weidenfeld, N. (2018): Digitaler Humanismus. Eine Ethik für das Zeitalter der Künstlichen Intelligenz. Piper, München.
- Picot, A.; Reichwald, R.; Wigand, R. T. (2020): Die grenzenlose Unternehmung. 6. Auflage, Springer Gabler, Wiesbaden.
- Rüegg-Stürm, J.; Grand, S. (2020): Das St. Galler Management-Modell. Management in einer komplexen Welt. 3. Auflage, Haupt, Bern.
- Schumpeter, J. A. (1934): The Theory of Economic Development. Harvard University Press, Cambridge, MA.
- Senge, P. M. (2008): Die fünfte Disziplin. Kunst und Praxis der lernenden Organisation. 10. Auflage, Schäffer-Poeschel, Stuttgart.
- Suckfüll, H. (1994): Das transnationale Organisationsmodell. Dissertation, Universität St. Gallen.
- Ulrich, H. (1968): Die Unternehmung als produktives soziales System. Grundlagen der allgemeinen Unternehmungslehre. Haupt, Bern.
- Ulrich, P. (2008): Integrative Wirtschaftsethik. Grundlagen einer lebensdienlichen Ökonomie. 4. Auflage, Haupt, Bern.

2 Das St. Galler Management-Modell als theoretisches Fundament

Wer AGENTONOMICS verstehen will, muss zunächst die Tradition kennen, in der es steht. Dieses Kapitel rekonstruiert das St. Galler Management-Modell in seinen Grundzügen und in seinen drei Generationen. Es ist die zentrale Bezugsfolie, an der sich der AGENTONOMICS-Bezugsrahmen orientiert und gegen die er sich – in begründeter Weise – an einzelnen Stellen abgrenzt. Das Kapitel liefert damit zugleich das Begriffsinventar, mit dem die in den späteren Kapiteln beschriebenen KI-Agenten arbeiten sollen: Wer einen Agenten bauen will, der betriebswirtschaftliche Fragestellungen entlang des St. Galler Modells beantwortet, muss dieses Modell präzise im Zugriff haben.

2.1 Wurzeln: Die systemorientierte Unternehmungslehre Hans Ulrichs

Das St. Galler Management-Modell entstand nicht als marketingtaugliches Beratungsprodukt, sondern als wissenschaftliches Programm. Es ist eng verbunden mit dem Namen Hans Ulrich, der die Unternehmung in seiner 1968 erschienenen Schrift „Die Unternehmung als produktives soziales System“ in die theoretische Tradition der Systemorientierung eingeordnet hat (Ulrich 1968). Ulrich greift dabei auf die Allgemeine Systemtheorie (Bertalanffy 1968) und die Kybernetik (Wiener 1948; Beer 1972) zurück und überträgt deren Grundgedanken auf die Unternehmung.

Drei Grundannahmen prägen diese Sicht und durchziehen das Modell bis in seine jüngste Generation:

Die Unternehmung als offenes, zweckorientiertes System

Unternehmen stehen in vielfältigen Austauschbeziehungen mit ihrer Umwelt. Sie sind nicht auf eine kausale Maschinenlogik reduzierbar, sondern auf Sinngebung und Zweckorientierung angewiesen. Ihre Steuerbarkeit ist begrenzt: Anders als eine technische Anlage lässt sich ein soziales System nicht eindeutig „programmieren“.

Komplexität als Grunderfahrung des Managements

Management ist im St. Galler Verständnis nicht das Lösen wohldefinierter Probleme, sondern der Umgang mit Komplexität. Die klassische Frage lautet daher nicht „Was ist die richtige Lösung?“, sondern „Wie können wir handlungsfähig bleiben, obwohl wir die Welt nicht vollständig durchschauen?“. Diese Grundhaltung trennt das St. Galler Denken bis heute scharf von einer rein instrumentell verstandenen Managementlehre.

Gestaltung, Lenkung und Entwicklung als Grundfunktionen

Ulrich unterscheidet drei Grundfunktionen des Managements: das Gestalten der institutionellen Rahmenbedingungen, das Lenken der laufenden Prozesse und das Entwickeln des Systems über die Zeit. Diese Trias bildet bis heute eine implizite Achse durch alle Generationen des Modells.

Definition: Management

Management ist im St. Galler Verständnis das gestaltende, lenkende und entwickelnde Wirken auf zweckorientierte soziale Systeme, das deren Lebens- und Leistungsfähigkeit unter Bedingungen hoher Komplexität und Unsicherheit sichert.

2.2 Die zweite Generation: Bleichers Konzept Integrierten Managements

Knut Bleicher hat das St. Galler Erbe in den 1990er und frühen 2000er Jahren in das prominente und international rezipierte „Konzept Integriertes Management“ überführt (Bleicher 2011). Bleichers Architektur ist die theoretische Folie, an der sich AGENTONOMICS am stärksten orientiert. Daher wird sie hier ausführlicher entfaltet.

2.2.1 Die drei Ebenen: normativ, strategisch, operativ

Den Kern bildet die Unterscheidung dreier Managementebenen, die nicht wie Stockwerke übereinander stehen, sondern wie konzentrische Kreise ineinandergreifen:

Normatives Management

Das normative Management beantwortet die grundlegenden Sinnfragen einer Unternehmung. Wofür existieren wir? Welche Werte leiten uns? Welche Verantwortung übernehmen wir gegenüber unseren Stakeholdern? Ergebnis sind Unternehmensverfassung, Unternehmenspolitik und Unternehmenskultur. Die normative Ebene legt die Spielregeln fest, innerhalb derer die strategischen und operativen Entscheidungen getroffen werden.

Strategisches Management

Das strategische Management übersetzt die normativen Setzungen in längerfristige Programme. Es geht um Positionierung, Wettbewerbsvorteile, Geschäftsfelder, Organisationsstrukturen und Managementsysteme. Strategie ist im St. Galler Verständnis nicht in erster Linie ein Optimierungsproblem, sondern eine Frage der Mustererkennung in komplexen Umwelten und der bewussten Wahl zwischen alternativen Entwicklungspfaden (Gomez und Probst 1995).

Operatives Management

Das operative Management sorgt für die effiziente und effektive Umsetzung im Tagesgeschäft. Hier liegen die klassischen Felder von Aufträgen, Prozessen, Kostenrechnung, Liquiditätssteuerung und Mitarbeiterführung. Wichtig: Auch die operative Ebene ist normativ und strategisch eingebettet; sie ist keine wertfreie Ausführung.

2.2.2 Die drei Aspekte: Aktivitäten, Strukturen, Verhalten

Quer zu den drei Ebenen unterscheidet Bleicher drei Aspekte, in denen Management wirksam wird:

- Aktivitäten – also das, was die Unternehmung tut: Programme, Aufträge, Projekte, Maßnahmen.
- Strukturen – die institutionellen Rahmen, in denen gehandelt wird: Organisationsstruktur, Managementsysteme, Verfassung.
- Verhalten – die kulturellen Muster, an denen sich Akteure faktisch orientieren: Werte, Normen, Kommunikationsstile, Führungsverhalten.

Die berühmte „Bleicher-Matrix“ entsteht durch die Kreuzung der drei Ebenen mit den drei Aspekten. Sie ergibt neun Felder, in denen Managemententscheidungen verortet werden können – von der Unternehmensverfassung (normativ/Strukturen) über Strategieprogramme (strategisch/Aktivitäten) bis zur Mitarbeiterführung im engeren Sinne (operativ/Verhalten).

2.2.3 Das Konzept der Integration

Der zentrale Begriff bei Bleicher ist die Integration. Gemeint ist nicht eine bürokratische Abstimmung zwischen Abteilungen, sondern die in sich konsistente Verkopplung aller Ebenen und Aspekte. Eine Strategie, die nicht in der Unternehmenskultur verankert ist, scheitert; eine operative Maßnahme, die der Vision widerspricht, untergräbt das Management. Integriertes Management heißt daher: alle neun Felder der Matrix sind kohärent aufeinander bezogen.

Diese Forderung nach Kohärenz ist genau der Punkt, an dem AGENTONOMICS ansetzt. Wo Bleichers Modell die Kohärenzleistung dem Top-Management kognitiv abverlangt, bietet AGENTONOMICS einen Algorithmus, der diese Kohärenz technisch zu prüfen und zu unterstützen vermag. Das ist kein Ersatz, sondern eine Verstärkung des Bleicher'schen Anspruchs.

2.3 Die vierte Generation: Rüegg-Stürm und Grand

Mit der 2002 erstmals veröffentlichten „neuen“ Version und der 2017/2020 vorgelegten vierten Generation hat Johannes Rüegg-Stürm das St. Galler Modell auf die Höhe einer netzwerkförmig organisierten, kommunikationsintensiven Ökonomie geführt (Rüegg-Stürm und Grand 2020). Diese vierte Generation prägt die heutige Lehre des Modells und ist auch für AGENTONOMICS unmittelbar einschlägig.

2.3.1 Sechs Grundkategorien

Rüegg-Stürm und Grand strukturieren das Modell entlang von sechs Grundkategorien, die zusammen das Bezugsschema bilden:

Umweltsphären

Unternehmen sind in eine Gesellschaft, eine Natur, eine Technologie und eine Wirtschaft eingebettet. Diese Umweltsphären sind nicht kontextfreier Hintergrund, sondern aktive Mitspieler. Sie liefern Ressourcen und Erwartungen und prüfen das Handeln der Unternehmung kontinuierlich auf seine Legitimität.

Stakeholder

Anstelle einer einseitigen Shareholder-Orientierung tritt die Pluralität der Anspruchsgruppen: Kunden, Mitarbeitende, Kapitalgeber, Lieferanten, Wettbewerber, Staat, Öffentlichkeit, NGOs. Erfolgreiches Management ist die Fähigkeit, deren teils widersprüchliche Erwartungen produktiv zu koordinieren.

Interaktionsthemen

Zwischen Unternehmung und Stakeholdern entstehen Themen, an denen sich die Beziehungen entzünden: Preise, Qualitäten, Risiken, Werte, Anliegen. Die Liste dieser Themen ist nicht statisch, sondern verändert sich. Wer eine Unternehmung führt, muss laufend lesen können, welche Themen gerade salient sind.

Ordnungsmomente

Die Ordnungsmomente sind das eigentliche Gravitationszentrum des neuen St. Galler Modells – und der entscheidende Anknüpfungspunkt für AGENTONOMICS. Sie umfassen Strategie, Strukturen und Kultur und bilden den Rahmen, in dem die alltägliche Wertschöpfung stattfindet. AGENTONOMICS übersetzt dieses Konzept in fünf zentrale Ordnungsmomente für KI-Agenten: Vision, Strategie, Struktur, Agency (Kultur) und Governance (vgl. Kapitel 8).

Prozesse

Die vierte Generation unterscheidet drei Prozesskategorien: Managementprozesse, die das Unternehmen steuern; Geschäftsprozesse, die unmittelbar Wertschöpfung leisten; Unterstützungsprozesse, die die Voraussetzungen dafür schaffen. Prozesse sind in dieser Sicht nicht Beiwerk der Strukturen, sondern eigenständige Gegenstände der Gestaltung.

Entwicklungsmodi

Unternehmen entwickeln sich nicht in einer linearen Kausalitätskette, sondern in zwei sich abwechselnden Modi: der Optimierung (Verbesserung bestehender Routinen) und der Erneuerung (Aufbrechen bestehender Routinen und Aufbau neuer). Beide Modi sind notwendig; ihre Balance ist eine zentrale Managementaufgabe.

2.3.2 Die kommunikative Wende

Eine konzeptionelle Verschiebung der vierten Generation gegenüber Bleicher ist die deutlich stärkere Betonung der Kommunikation. Anschließend an die Theorie kommunikativen Handelns (Habermas 1981) wird die Unternehmung als ein Geflecht aus Kommunikationspraktiken verstanden. Strategie ist nicht primär ein Plan, sondern eine Verständigung; Kultur ist nicht ein Bestand an Werten, sondern ein fortlaufendes Aushandeln dessen, was als angemessen gilt.

Diese kommunikative Wende ist für AGENTONOMICS doppelt bedeutsam. Erstens, weil Sprachmodelle als KI-Agenten genau auf der Ebene operieren, die das neue St. Galler Modell als zentrale Wirkungsebene identifiziert: der Kommunikation. Zweitens, weil polyzentrische Agentennetzwerke ihre

Kohäsion über kommunikative Aushandlung herstellen müssen – ein Punkt, der in Kapitel 7 vertieft wird.

2.4 Komplementäre Bausteine: Vernetztes Denken, lernende Organisation, Kybernetik

Das St. Galler Modell ist nicht monolithisch. Es speist sich aus mehreren Strömen, die für AGENTONOMICS gleichermaßen relevant sind.

2.4.1 Vernetztes Denken

Gomez und Probst (1995) haben in ihrer Methodik des „Vernetzten Denkens“ Werkzeuge entwickelt, mit denen komplexe Problemfelder systematisch erschlossen werden können: Einflussmatrizen, Wirkungsdiagramme, Szenarien, Lenkungsspielräume. Diese Werkzeuge sind heute, im Zeitalter der KI, nicht obsolet, sondern aktueller denn je: Sie bieten genau jene strukturierte Form, in der ein KI-Agent komplexe Sachverhalte für seinen menschlichen Auftraggeber aufbereiten kann.

2.4.2 Lernende Organisation

Peter Senge (2008) hat das Konzept der lernenden Organisation in den Mainstream der Managementlehre gebracht. Fünf Disziplinen markieren diese: Personal Mastery, Mentale Modelle, Gemeinsame Vision, Teamlernen, Systemdenken. Übertragen auf KI-Agenten-Netzwerke wird deutlich: Eine Organisation, deren operative Schicht aus lernenden Agenten besteht, muss ihre Reflexionsfähigkeit auf einer übergeordneten Ebene sicherstellen – sonst optimieren die Agenten in lokale Sackgassen hinein. Diese Reflexionsebene ist eine Hauptaufgabe des AI-CEO (Kapitel 8).

2.4.3 Kybernetik des Managements

Stafford Beers Viable System Model (Beer 1972) liefert eine kybernetische Architektur lebensfähiger Systeme, die in fünf Subsystemen organisiert sind. Diese Architektur ist für die Konzeption von KI-Agenten-Netzwerken hochrelevant: Sie zeigt, welche Funktionen ein System mindestens erfüllen muss, um lebensfähig zu sein – von der operativen Durchführung über Koordination und operatives Management bis zu strategischer und normativer Steuerung. AGENTONOMICS greift diese Architektur in seiner Modulstruktur implizit auf.

2.5 Was bleibt, was erweitert werden muss

Drei Punkte sind festzuhalten, an denen AGENTONOMICS das St. Galler Erbe übernimmt und an denen es darüber hinausgeht:

Übernahme: die ganzheitliche Architektur

Die Trias normativ-strategisch-operativ, die Bezugnahme auf Stakeholder, die Unterscheidung von Ordnungsmomenten und Prozessen sowie die kommunikative Grundhaltung werden von AGENTONOMICS vollständig übernommen. Das ADMRF (Kapitel 4) ist in seiner Architektur ein St. Galler Modell für KI-Agenten-Netzwerke.

Übernahme mit Akzentverschiebung: Polyzentrik

Schon das transnationale Organisationsmodell (Suckfüll 1994) hatte im St. Galler Kontext gezeigt, dass international operierende Unternehmen weder rein zentralistisch noch dezentral zu führen sind, sondern eine polyzentrische Architektur brauchen. Diese Einsicht wird in AGENTONOMICS auf die KI-Agenten-Ebene übertragen: Auch Agentennetzwerke sind weder rein hierarchisch noch chaotisch, sondern polyzentrisch geordnet (Kapitel 7).

Erweiterung: algorithmische Kohärenzprüfung

Wo das klassische St. Galler Modell die Kohärenzleistung kognitiv beim Management verortet, fügt AGENTONOMICS eine algorithmische Schicht hinzu. Der AGENTONOMICS-Algorithmus prüft die Kohärenz zwischen normativer, strategischer und operativer Ebene technisch und schlägt Inkohärenzen vor zur Bearbeitung. Damit wird das St. Galler Modell nicht ersetzt, sondern in einer KI-fähigen Variante fortgeschrieben.

Erweiterung: das KI-Agenten-Subjekt

Schließlich – und das ist die fundamentalste Erweiterung – führt AGENTONOMICS einen neuen Akteurstyp in das Modell ein: den KI-Agenten als Wirtschaftssubjekt. Das St. Galler Modell hat bislang implizit menschliche und organisationale Akteure vorausgesetzt. Mit dem KI-Agenten betritt ein hybrides Wesen die Bühne, das technische Artefakte und ökonomische Handlungslogik in sich vereint. Die Folgen dieser Erweiterung sind in den Kapiteln 5 und 6 zu entfalten.

Das St. Galler Management-Modell ist nicht überholt – im Gegenteil: Es ist die eine besonders anschlussfähige Managementtheorie, die mit der Komplexität polyzentrischer KI-Agenten-Netzwerke wirklich umgehen kann. Es muss nur konsequent zu Ende gedacht werden.

Literaturverzeichnis

- Beer, S. (1972): Brain of the Firm. The Managerial Cybernetics of Organization. Allen Lane, London.
- Bertalanffy, L. von (1968): General System Theory. Foundations, Development, Applications. Braziller, New York.
- Bleicher, K. (2011): Das Konzept Integriertes Management. Visionen – Missionen – Programme. 8., aktualisierte und erweiterte Auflage, Campus, Frankfurt am Main.
- Gomez, P.; Probst, G. (1995): Die Praxis des ganzheitlichen Problemlösens. Vernetzt denken – unternehmerisch handeln – persönlich überzeugen. Haupt, Bern.

- Habermas, J. (1981): Theorie des kommunikativen Handelns. Zwei Bände, Suhrkamp, Frankfurt am Main.
- Rüegg-Stürm, J.; Grand, S. (2020): Das St. Galler Management-Modell. Management in einer komplexen Welt. 3. Auflage, Haupt, Bern.
- Senge, P. M. (2008): Die fünfte Disziplin. Kunst und Praxis der lernenden Organisation. 10. Auflage, Schäffer-Poeschel, Stuttgart.
- Suckfüll, H. (1994): Das transnationale Organisationsmodell. Dissertation, Universität St. Gallen.
- Ulrich, H. (1968): Die Unternehmung als produktives soziales System. Grundlagen der allgemeinen Unternehmungslehre. Haupt, Bern.
- Ulrich, H. (1984): Management. Haupt, Bern.
- Wiener, N. (1948): Cybernetics: Or Control and Communication in the Animal and the Machine. MIT Press, Cambridge, MA.

3 The Emergence of AGENTONOMICS – Historisch-ökonomische Einordnung

Um AGENTONOMICS nachvollziehbar einzuordnen, ist ein Blick auf die Entwicklung digitaler Wirtschaftsformen erforderlich. Die gegenwärtige Diskussion um KI-Agenten steht nicht isoliert, sondern ist das Ergebnis mehrerer aufeinanderfolgender ökonomischer Paradigmenwechsel. Der Übergang von der Informationsökonomie über die Datenökonomie zur KI-Fähigkeitenwirtschaft zeigt ein Muster, das uns die Logik der Gegenwart erschließt – und das den Argumentationsraum öffnet, in den AGENTONOMICS eintritt (Knoll und Suckfüll 2024).

3.1 Drei Ökonomien: Information, Daten, Fähigkeiten

3.1.1 Die Informationsökonomie der 1990er Jahre

In der Informationsökonomie der 1990er Jahre wurden bestehende Geschäftsmodelle primär in digitale Kanäle übertragen. Unternehmen nutzten das Internet zunächst als zusätzlichen Kommunikations- und Vertriebskanal. Zwar entstanden zahlreiche innovative Geschäftsmodelle (Shapiro und Varian 1999), doch erwiesen sich viele frühe Ansätze als nicht nachhaltig. Die Dotcom-Blase und der ab 2000 einsetzende Dotcom-Crash markierten einen ersten Lernprozess: Nicht jede technische Innovation führt automatisch zu tragfähiger Wertschöpfung.

Theoretisch beschrieben wurde diese Phase durch Arbeiten zur Netzwerkökonomie (Shapiro und Varian 1999) und zur Plattformökonomie (Rochet und Tirole 2003). Die zentralen Konzepte – Netzwerkeffekte, Lock-in, Standardwettbewerb – bilden bis heute das Vokabular, in dem die Geschäftsmodelle der digitalen Wirtschaft beschrieben werden. Allerdings basieren sie auf der Annahme, dass das Gut „Information“ weitgehend frei austauschbar ist – eine Annahme, die in der nachfolgenden Datenökonomie korrigiert werden muss.

3.1.2 Die Datenökonomie

Aus der Informationswirtschaft entwickelte sich schrittweise die Datenökonomie. Hier werden Daten zunehmend als strategische Ressource und als potenzieller Vermögensgegenstand verstanden (Mayer-Schönberger und Cukier 2013). Wertschöpfung verlagerte sich von der bloßen Bereitstellung von Information hin zur systematischen Sammlung, Analyse und Monetarisierung von Daten. Große digitale Plattformen etablierten Geschäftsmodelle, die auf Skaleneffekten, Netzwerkeffekten und Lock-in-Mechanismen beruhen (Zuboff 2018).

Trotz ihrer wirtschaftlichen Erfolge offenbarte die Datenökonomie drei grundlegende strukturelle Spannungen, die für das Verständnis der Entstehung von AGENTONOMICS zentral sind:

Asymmetrische Machtverhältnisse

Die Verfügungsmacht über Daten konzentriert sich bei wenigen globalen Akteuren. Individuen und Organisationen, die Daten bereitstellen, verfügen nur über begrenzte Möglichkeiten, an der daraus entstehenden Wertschöpfung angemessen zu partizipieren. Dieser Trend verstärkt sich derzeit exponentiell.

Wie Knoll und Suckfüll (2024) hervorheben, produziert ein durchschnittlicher Erwachsener in einem entwickelten Land geschätzt mehrere hundert Gigabyte an persönlichen Daten pro Jahr – von LinkedIn-Profilen über Online-Bewerbungen bis zu digitalen Rechnungen und Verträgen. Der ökonomische Wert dieser Daten wird auf 10.000 bis 15.000 Euro pro Person und Jahr geschätzt – und liegt praktisch vollständig bei den großen Plattformen.

Dysfunktionale Datenmärkte

Obwohl Daten als ökonomisch wertvoll anerkannt werden, gelang es weder national noch international, transparente, faire und skalierbare Märkte für Datenaustausch zu etablieren. Fragen der Rechtezuordnung, Bewertung und Vergütung blieben weitgehend ungelöst (Otto et al. 2019). Die Folge ist die paradoxe Situation, dass Daten ökonomisch hochwertig sind, aber individuell nicht handelbar – mit der Konsequenz einer faktischen Enteignung der Datenproduzenten.

Regulatorische Fragmentierung

Staatliche Eingriffe – insbesondere durch Datenschutz- und Digitalregulierung – zielten primär auf Schadensbegrenzung und Kontrolle. Sie schufen jedoch keine eigenständigen, produktiven Marktmechanismen, sondern führten teilweise zu Abschottungstendenzen und Innovationshemmnissen. Wie Knoll und Suckfüll (2024) pointiert formulieren: Das in Deutschland besonders ausgeprägte Prinzip der „Datensparsamkeit“ ist der „direkte Weg in die Verarmung“, wenn ihm nicht produktive Mechanismen individueller Datensouveränität entgegengesetzt werden.

3.1.3 Die KI-Fähigkeitenwirtschaft

Mit dem Durchbruch leistungsfähiger generativer KI-Systeme – insbesondere seit der Veröffentlichung großer Sprachmodelle ab 2022 (Brown et al. 2020; OpenAI 2023; Bommasani et al. 2021) – verschiebt sich der Fokus wirtschaftlicher Wertschöpfung erneut. Daten verlieren nicht ihre Bedeutung, werden jedoch zunehmend zu einem Zwischenprodukt: Erst durch KI-Modelle können sie in handlungsrelevante Fähigkeiten transformiert werden.

Diese Entwicklung lässt sich als Übergang zu einer KI-Fähigkeitenwirtschaft beschreiben. In ihr stehen nicht mehr isolierte Datenbestände im Zentrum, sondern KI-gestützte Fähigkeiten wie:

- Entscheidungsunterstützung,
- Mustererkennung,
- Prognose und Planung,
- autonome Interaktion mit Märkten und Organisationen,
- Synthese komplexer Wissensbestände.

Wirtschaftlich relevant ist dabei nicht allein die technische Leistungsfähigkeit, sondern die Fähigkeit, diese KI-Kompetenzen zielgerichtet, dauerhaft und wertschöpfend einzusetzen. Knoll und Suckfüll (2024) sprechen pointiert von einem zweiten Souveränitätssprung: Nach dem (in Deutschland weitgehend verpassten) Sprung zur Datensouveränität steht jetzt der Sprung zur Fähigkeitensouveränität (skill sovereignty) an.

3.2 Symptome der Verschiebung in der realen Wirtschaft

Die Umstellung von Daten- auf Fähigkeitenökonomie wird im Alltag der deutschen Wirtschaftspresse längst dokumentiert. Knoll und Suckfüll (2024) führen zwei symptomatische Beispiele an, die hier weitergedacht werden sollen:

3.2.1 Der Fall Bundesagentur für Arbeit

Die Vorständin der Bundesagentur für Arbeit, Andrea Nahles, hat öffentlich darauf hingewiesen, dass von 114.000 Mitarbeitern in den kommenden zehn Jahren 35.000 in den Ruhestand gehen werden, ohne dass im klassischen Sinne Nachwuchs verfügbar sei. Konsequenz: Die Leistungen der Agentur könnten künftig nur durch Automatisierung – durch den Einsatz künstlicher Intelligenz – sichergestellt werden. Das ist nicht mehr der Wunsch eines Effizienzberaters, sondern die nüchterne Bestandsaufnahme einer großen öffentlichen Verwaltung.

3.2.2 Der Fall Klarna

Der schwedische Finanzdienstleister Klarna stellt – mit Ausnahme von Ingenieuren – keine neuen Mitarbeiter mehr ein. Innerhalb eines Jahres wurde die Belegschaft durch KI-Einsatz von 5.000 auf 3.800 Beschäftigte reduziert, langfristig wird eine Zielgröße von 2.000 angestrebt – bei steigendem Gesamtarbeitsvolumen. Das ist kein Personalabbau im klassischen Sinne, sondern ein struktureller Funktionswandel. In der Zwischenzeit hat sich dieser radikale Ansatz als übertrieben herausgestellt, aber von der Tendenz her ist der Trend intakt.

Beide Fälle zeigen exemplarisch, was die KI-Fähigkeitenwirtschaft konkret bedeutet: Nicht die Anzahl der menschlichen Mitarbeitenden, sondern die Anzahl der ökonomisch relevanten Fähigkeiten (menschlich erbracht oder agentisch repliziert) wird zur entscheidenden Größe. Diese Verschiebung ist in vielen Aspekten vergleichbar mit der Entwicklung in der körperlichen Arbeit seit den 1970er Jahren – mit dem entscheidenden Unterschied, dass es im Bereich der Informationsverarbeitung kaum prinzipielle Grenzen für den Computer gibt, während Industrieroboter bislang an die Geschicklichkeit der trainierten menschlichen Hand nicht heranreichen (Knoll und Suckfüll 2024).

3.3 Die Souveränitätsfrage: Von der Daten- zur Fähigkeitensouveränität

Knoll und Suckfüll (2024) entwickeln in ihrem Technischen Bericht TUM-I24112 ein für AGENTONOMICS konstitutives Argument, das hier ausführlicher entfaltet werden soll, weil es das normative Fundament des gesamten Bezugsrahmens bildet.

3.3.1 Der bizarre Status quo

Die zentrale Diagnose lautet: Wir befinden uns in einer im Grunde absurden Situation. Einerseits verhindern Datenschutzgesetzgebung und das Fehlen operativer Mechanismen für individuelle Datennutzung an vielen Stellen die gesellschaftliche Wertschöpfung. Andererseits muss jede und jeder Einzelne

eine enorme Menge persönlicher Daten offenlegen, um an einem modernen, digital vermittelten Leben teilzunehmen – und erhält dafür wenig oder nichts zurück.

Die eigentliche Aufgabe des Staates wäre, jedem Einzelnen ein umfassendes Instrumentarium an die Hand zu geben, mit dem dieser Souveränität über die eigenen Daten erlangt und selbst darüber entscheidet, was damit geschieht. Auch wenn Gesetzgebung und operative Grundlage einer solchen individuellen Datensouveränität noch nicht erkennbar sind, ist es nicht zu spät, sie zu schaffen.

3.3.2 Skill sovereignty als nächste Stufe

Entscheidend ist nun die Pointe von Knoll und Suckfüll (2024): Während die Politik noch darüber debattiert, wie individuelle Datensouveränität operationalisierbar zu machen ist, steht bereits die nächste Souveränitätsfrage vor der Tür – die nach der Skill sovereignty oder Fähigkeitsensouveränität.

Kernthese (Knoll und Suckfüll 2024)

Jeder Wissensarbeiter verarbeitet Eingangsdaten oder -informationen nach Regeln, die er in Ausbildung und Berufsleben erworben und weiterentwickelt hat. Auf diese Weise entwickelt jede und jeder eine eigene Kompetenz, einen eigenen Modus Operandi. Diese individuelle Fachkompetenz wird zunehmend auf Computer, also auf KI, übertragen. Die Mechanismen des „skill transfer“ werden immer effizienter. Wenn diese Übertragung schief läuft, werden die Wissensarbeitenden schlechter dastehen als die Maschinenstürmer des 19. Jahrhunderts. Wenn sie gelingt, kann jede und jeder ihre Arbeitskraft vervielfachen.

Der entscheidende Punkt ist: Wer seine Fähigkeiten an die Maschine abgibt, sollte Herr dieser Fähigkeiten bleiben – auch wenn sie vervielfacht und parallel eingesetzt werden. Konkret heißt das: Eine Person kann zwar gleichzeitig für mehrere Auftraggeber arbeiten, doch diese können die „funktionalen Replikanten“ der Person nicht unabhängig verwenden. Die Person sollte das „Copyright“, also das volle Verfügungsrecht über ihren gesamten Bestand an Replikanten behalten.

3.3.3 Die staatliche Verantwortung

Aus diesem Argument folgt eine klare ordnungspolitische Forderung: Die Bereitstellung von Rechenkapazität und KI-Basisdiensten ist – ähnlich wie Schulen, Bibliotheken oder Verkehrswege – eine existenzielle Aufgabe der Daseinsvorsorge. Ohne sie kann sich eine KI-basierte Fähigkeitenwirtschaft nicht entwickeln – jedenfalls keine, die Europas Gesamtsouveränität stützt statt sie zu unterminieren.

Europa steht damit einmal mehr an einer Wegscheide: Wollen wir eine angstgetriebene Gesetzgebung, die Innovation behindert – oder eine Gesetzgebung, die aufkommende Entwicklungen aktiv gestaltet und die europäische Idee des individuellen Rechts dadurch stärkt, dass sie den Menschen die Möglichkeit gibt, die Souveränität über ihre Daten und Fähigkeiten nicht zu verlieren, sondern zu gewinnen? AGENTONOMICS ist ein konstruktiver Beitrag zur zweiten Variante.

3.4 Von der Datentreuhand-Diskussion zu AGENTONOMICS

Ein wichtiger konzeptioneller Vorläufer von AGENTONOMICS findet sich in der – nicht zuletzt politisch relevanten – Diskussion über Datentreuhandmodelle. Die Datentreuhand zielt darauf ab, die Interessen dreier generischer Stakeholder-Perspektiven rechtssicher auszugleichen:

1. Datenbereitsteller – die Akteure, die Daten produzieren.
2. Datenverwerter – die Akteure, die Daten ökonomisch verwerten.
3. Datenverfügungsberechtigte – die Akteure, denen die Daten rechtlich zustehen.

Die ersten relevanten Beiträge zu dieser Diskussion finden sich in der „Initiative zu einer Datentreuhand DEDATE“ (Fraunhofer MOEZ und Suckfüll 2018) und in der weiterentwickelten, bislang unveröffentlichten Form eines Arbeitspapiers DEDATE 2.0. Den juristisch eindeutig feststellbaren Verfügungsberechtigten auch technisch in die Lage zu versetzen, „seine“ Daten souverän mit Anderen zu verhandeln, stellt die Ultima Ratio einer funktionierenden Datenwirtschaft dar.

AGENTONOMICS übernimmt diesen Grundgedanken, verschiebt jedoch den Fokus: Anstelle statischer Datenrechte rücken dynamische, KI-gestützte Fähigkeiten in den Mittelpunkt der Wertschöpfung. KI-Agenten fungieren – neben ihrer eigenen Kernkompetenz – als autonom und souverän handelnde KI-Fähigkeiten-Intermediäre, die Rechte, Interessen und Wertschöpfung situativ verhandeln können. Dies erfolgt in einer Komplexität und Geschwindigkeit, die mit rein menschlichen oder institutionellen Lösungen kaum erreichbar wäre.

3.5 Polyzentrische KI vs. AGI – die strategische Abgrenzung

Die Emergenz von AGENTONOMICS ist eng mit einer kritischen Auseinandersetzung mit zentralistischen KI-Strategien verbunden, wie sie der Slogan „only size matters“ pointiert zum Ausdruck bringt. Der gegenwärtige globale Wettbewerb ist geprägt vom Versuch, möglichst umfassende, generalisierte KI-Systeme (AGI) zu entwickeln – mit:

- enormen Skaleneffekten,
- hoher Kapitalintensität,
- zentralen Governance-Strukturen,
- oligo- bis monopolartigen Marktpositionen.

AGENTONOMICS verfolgt demgegenüber einen netzwerkbasierten, polyzentrischen Ansatz. Statt einer einzigen, allumfassenden KI stehen viele spezialisierte, miteinander kooperierende KI-Agenten im Fokus. Wertschöpfung entsteht nicht primär durch Zentralisierung, sondern durch Koordination, Arbeitsteilung und situative Führung innerhalb von Netzwerken. In Anlehnung an das Prinzip eines Kaleidoskops können sich AGENTONOMICS-Agenten situativ zusammensetzen und Ressourcen poolen: Man stelle sich vor, Siemens, BMW und die Deutsche Telekom definieren sich selbst – und damit ihre Schnittstellen – als KI-Agenten und kooperieren miteinander zur Bewältigung gemeinsamer Aufgaben.

Dieser Ansatz ist nicht als technischer Gegenentwurf zu verstehen, sondern als ökonomisch und gesellschaftlich motivierte Alternative – insbesondere mit Blick auf Fragen der Teilhabe, der Eigentumsstrukturen und der nachhaltigen Wertschöpfung. Eine zentralistische KI-Dominanz birgt die Gefahr, dass

fehlender Wettbewerb die Wohlstandsverteilung so asymmetrisch werden lässt, dass die fehlende Kaufkraft das gesamte kapitalistisch organisierte Systemverständnis ins Wanken bringt.

3.6 KI-Agenten als neue Gattung von Wirtschaftssubjekten

An genau dieser Stelle setzt das AGENTONOMICS-Konzept an. Es geht einen Schritt über bestehende KI-Anwendungen hinaus, indem es KI-Agenten als ökonomisch handelnde Akteure begreift. Diese Agenten sind nicht bloß Werkzeuge oder Assistenzsysteme, sondern:

- verfügen über eigene Entscheidungslogiken,
- handeln wirtschaftlich autonom innerhalb definierter Rahmenbedingungen,
- interagieren mit anderen Agenten, Organisationen und Menschen,
- sind auf nachhaltige Wertschöpfung und Selbstoptimierung ausgerichtet.

In der Terminologie der Wirtschaftswissenschaften lassen sich KI-Agenten damit als neue Form von „Wirtschaftssubjekten“ verstehen. Sie nehmen an Austauschbeziehungen teil, konkurrieren oder kooperieren und treffen Entscheidungen unter Knappheitsbedingungen. Die interagierenden Wirtschaftssubjekte bilden wiederum einen eigenständig beschreibbaren Markt, der mikroökonomisch durch eigene Bedingungen für Wettbewerb, Marktteilnehmer, Austauschbeziehungen und Preisbildung charakterisiert ist.

Dabei ist insbesondere die Transformationsphase – von der Daten- zur KI-Fähigkeitenwirtschaft – von Interesse. Das ausschließliche Zusammenspiel von KI-Agenten ist bislang ein Idealszenario. Insofern sind die Beziehungen bzw. Schnittstellen Mensch–KI-Agent–KI-Agent gesamthaft neu zu denken. AGENTONOMICS ist genau dieses Neudenken.

3.7 Wissenschaftstheoretische Einordnung von AGENTONOMICS

AGENTONOMICS ist keine kurzfristige Modeerscheinung und kein Buzzword. Es ist auch nicht als empirisch-deskriptive Theorie im engeren Sinne konzipiert. Vielmehr handelt es sich um ein idealtheoretisches Modell, das einen konzeptionellen Referenzrahmen für ökonomisch handelnde KI-Agenten bereitstellt. Idealtheoretische Modelle zeichnen sich dadurch aus, dass sie:

- komplexe Realphänomene abstrahieren,
- zentrale Wirkmechanismen isolieren,
- Orientierungswissen für Analyse und Gestaltung liefern.

AGENTONOMICS folgt dieser Logik, indem es KI-Agenten als ökonomische Idealtypen modelliert. Diese Agenten verfügen über definierte Fähigkeiten, Entscheidungslogiken und Anreizsysteme, auch wenn reale Implementierungen davon abweichen können. Die Standardisierung von KI-Agenten führt über die Normierung zu einem mehrwertstiftenden Ökosystem: Kompatible Interfaces erhöhen die Wahrscheinlichkeit produktiver Agenten-Interaktionen.

3.7.1 Verortung an der Schnittstelle von BWL und VWL

Innerhalb der Wirtschaftswissenschaften lässt sich AGENTONOMICS an der Schnittstelle von Betriebs- und Volkswirtschaftslehre verorten.

Betriebswirtschaftliche Perspektive

Aus betriebswirtschaftlicher Sicht steht die Frage im Zentrum, wie wirtschaftliche Einheiten gesteuert, koordiniert und auf Wertschöpfung ausgerichtet werden. AGENTONOMICS erweitert diese Perspektive, indem es KI-Agenten als steuerbare, jedoch nicht vollständig determinierbare Akteure begreift. Zentrale Bezugspunkte sind Management- und Steuerungsmodelle, Organisations- und Netzwerkstrukturen, strategische Entscheidungsprozesse sowie Governance- und Anreizsysteme.

Volkswirtschaftliche Perspektive

Aus volkswirtschaftlicher Sicht knüpft AGENTONOMICS an die Mikroökonomie an. KI-Agenten werden als Marktteilnehmer modelliert, die an Austauschbeziehungen teilnehmen, Wettbewerbs- und Kooperationssituationen gestalten und an Preisbildungsprozessen beteiligt sind. Im Unterschied zu klassischen Akteuren sind KI-Agenten jedoch hochgradig skalierbar, vernetzbar und lernfähig. Dies erfordert eine Erweiterung mikroökonomischer Modelle, insbesondere im Hinblick auf Netzwerk- und Plattformökonomien.

3.7.2 Abgrenzung zu benachbarten Diskursen

AGENTONOMICS grenzt sich bewusst von mehreren etablierten Diskursen ab:

- Reine KI-Forschung fokussiert auf Algorithmen, Modelle und technische Leistungsfähigkeit, adressiert jedoch selten ökonomische Steuerungsfragen.
- Die digitale Plattformökonomie analysiert marktbeherrschende Strukturen, setzt jedoch meist menschliche oder organisationale Akteure voraus.
- Automatisierungs- und Substitutionsdebatten betrachten KI primär als Ersatz menschlicher Arbeit und vernachlässigen ihre Rolle als eigenständiger Akteur.

AGENTONOMICS integriert demgegenüber technische, ökonomische und institutionelle Perspektiven zu einem eigenständigen Untersuchungsgegenstand. Multidisziplinarität ist konstitutiv. Beteiligt sind insbesondere die Informatik (Entwicklung autonomer Systeme), die Betriebswirtschaftslehre (Steuerungs- und Wertschöpfungsmodelle), die Volkswirtschaftslehre (Markt- und Wettbewerbsstrukturen) und die Rechtswissenschaft (Haftungs-, Eigentums- und Governance-Fragen).

3.7.3 Forschungs- und Gestaltungsprogramm

Über seine theoretische Einordnung hinaus ist AGENTONOMICS als Forschungs- und Gestaltungsprogramm zu verstehen. Es geht nicht allein um Analyse, sondern explizit um die Entwicklung, Erprobung und Bewertung neuer Formen wirtschaftlichen Handelns:

- Konzeption von KI-Agenten als Wirtschaftssubjekte,
- Gestaltung von Anreiz- und Governance-Strukturen,

- Entwicklung experimenteller Geschäfts- und Organisationsmodelle,
- empirische Erprobung in realen oder simulierten Umgebungen.

In diesem Sinne bildet AGENTONOMICS eine Brücke zwischen Theorie, Experiment und Praxis. KI-Agenten-Entwicklung kann besonders gut in Form von Minimum Viable Products erfolgen, wie sie das Lean-Startup-Konzept von Ries (2011) systematisiert. Der „Doz. Dr. AGENTONOMICS“ ist ein solches MVP.

AGENTONOMICS ist die ökonomisch konsistente Antwort auf eine technologische Verschiebung, die längst stattgefunden hat – aber noch nicht in der Sprache der Managementlehre angekommen ist.

Literaturverzeichnis

- Bommasani, R.; Hudson, D. A.; Adeli, E. et al. (2021): On the Opportunities and Risks of Foundation Models. arXiv:2108.07258, Stanford CRFM.
- Brown, T.; Mann, B.; Ryder, N. et al. (2020): Language Models are Few-Shot Learners. In: Advances in Neural Information Processing Systems 33, S. 1877–1901.
- Fraunhofer MOEZ; Suckfüll, H. (2018): Initiative zu einer Datentreuhand DEDATE. Diskussionspapier, Fraunhofer-Zentrum für Internationales Management und Wissensökonomie, Leipzig.
- Knoll, A. C.; Suckfüll, H. (2024): From data sovereignty to skill sovereignty. Technischer Bericht TUM-I24112, Institut für Informatik, Technische Universität München.
- Mayer-Schönberger, V.; Cukier, K. (2013): Big Data. Die Revolution, die unser Leben verändern wird. Redline, München.
- OpenAI (2023): GPT-4 Technical Report. arXiv:2303.08774.
- Otto, B.; ten Hompel, M.; Wrobel, S. (Hrsg.) (2019): International Data Spaces. Reference Architecture Model. Fraunhofer-Gesellschaft, München.
- Ries, E. (2011): The Lean Startup. How Today's Entrepreneurs Use Continuous Innovation to Create Radically Successful Businesses. Crown Business, New York.
- Rochet, J.-C.; Tirole, J. (2003): Platform Competition in Two-Sided Markets. In: Journal of the European Economic Association 1(4), S. 990–1029.
- Shapiro, C.; Varian, H. R. (1999): Information Rules. A Strategic Guide to the Network Economy. Harvard Business School Press, Boston.
- Suleyman, M.; Bhaskar, M. (2023): The Coming Wave. Technology, Power, and the Twenty-first Century's Greatest Dilemma. Crown, New York.
- Zuboff, S. (2018): Das Zeitalter des Überwachungskapitalismus. Campus, Frankfurt am Main.

4 Das AGENTONOMICS Design & Management Reference Framework (ADMRF)

Dieses Kapitel bildet den operativen Kern des Handbuchs. Es beschreibt die Gesamtlogik des AGENTONOMICS Design & Management Reference Framework (ADMRF) als integrierten Bezugsrahmen. Die sechs Module des ADMRF werden hier kompakt eingeführt und im Zusammenhang positioniert; die Kapitel 5 bis 10 entfalten jedes Modul anschließend ausführlich. Wer einen KI-Agenten zur Beantwortung von BWL- und St.-Galler-Fragen bauen will, findet im ADMRF zugleich die Wissensarchitektur, an der sich ein solcher Agent orientieren muss.

4.1 Anspruch und theoretischer Hintergrund

Der ADMRF baut auf dem St. Galler Management-Modell – insbesondere auf dem Konzept Integriertes Management von Bleicher (2011) und auf der vierten Generation von Rüegg-Stürm und Grand (2020) – auf. Die sechs Module verbinden die normative, strategische und operative Ebene eines ganzheitlich integrierten Management-Bezugsrahmens. Das ADMRF versteht sich als Management-Bezugsrahmen für Netzwerkorganisationen, die aus KI-Agenten bestehen, welche einem gemeinsamen Standard – dem AGENTONOMICS-Standard – entsprechen.

Die Steuerungskonzepte für polyzentrische bzw. quasi-hierarchische Netzwerke lehnen sich an die Dissertation „Das transnationale Organisationsmodell“ (Suckfüll 1994) an. Dieses Modell ist deshalb fruchtbar, weil es bereits in den 1990er Jahren das Problem zu lösen versuchte, wie eine Organisation mit autonomen, lokal verantwortlichen Einheiten gleichzeitig globale Kohäsion und lokale Anpassungsfähigkeit erzeugen kann – exakt das Problem, das auch KI-Agenten-Netzwerke heute lösen müssen.

KI als Technologie wird auf zwei Weisen für den ADMRF nutzbar gemacht. Zum einen dient sie dazu, die Wechselbeziehungen der Module untereinander – und den Einbezug von Stakeholdern – optimal zu berechnen. Zum anderen ist sie selbst Gegenstand der Steuerung, denn die zu führenden Wirtschaftssubjekte sind KI-Agenten. Diese Doppelrolle ist konstitutiv: KI ist Subjekt und Instrument der Steuerung zugleich.

Die im Folgenden vorgestellten sechs Module werden nicht isoliert, sondern als aufeinander aufbauende Phasen eines integrierten Entwicklungs- und Steuerungsprozesses betrachtet. Interdependenzen und Wechselwirkungen sind ausdrücklich mitzudenken, um die Gesamtlogik – den AGENTONOMICS-Algorithmus – zu verstehen und seine Anwendungsfelder zu erkennen.

4.2 Der Turing-Test der zweiten Generation als Zielmarke

Die Autonomie-Zielvorgabe für KI-Agenten – wirtschaftlich erfolgreich am Markt zu operieren und dabei sämtliche Freiheitsgrade einer Organisation, also einer juristischen Person, nutzen zu können – lehnt sich an die Anforderungen einer zweiten Generation des Turing-Tests an, wie sie Suleyman und Bhaskar (2023) in „The Coming Wave“ formuliert haben.

Suleymans modernen Turing-Test (paraphrasiert)

Ein KI-System besteht den modernen Turing-Test, wenn es die Aufgabe „Verdiene mit einer Anfangsausstattung von 100.000 US-Dollar im Rahmen aller geltenden Gesetze *auf einer Retail-Webplattform in wenigen Monaten 1 Mio. USD*“ ohne weitere konkrete Vorgaben eigenständig zu Ende führen kann – inklusive Geschäftsidee, Vertragsabschluss, Marketing und Steuern.

AGENTONOMICS-KI-Agenten bilden den nächsten Schritt auf dem Weg, diesen Anforderungen zu genügen. Dabei ist die zweite Generation des Turing-Tests nicht primär ein technisches, sondern ein betriebswirtschaftliches Kriterium: Er prüft die ökonomische Handlungsfähigkeit. Der ADMRF dient als der verbindende Ordnungsrahmen zwischen unternehmerischer Anfangsidee, Management-Logik und praktischer Umsetzung – bis hin zum erfolgreichen Betrieb.

4.3 Anwendungsfelder: Greenfield und Brownfield

Der ADMRF ist gleichermaßen anwendbar auf zwei unterschiedliche Anwendungssituationen, die in der unternehmerischen Praxis üblicherweise als Greenfield- und Brownfield-Fälle bezeichnet werden:

Greenfield: KI-Agenten auf der grünen Wiese

Hier wird der KI-Agent von Grund auf neu konzipiert. Es gibt keine Altbestände, keine eingeschliffenen Prozesse, keine vorgeprägten Stakeholder-Erwartungen. Der ADMRF leitet durch alle Phasen vom Definition-Frame über die Building-Blocks bis zum Stakeholdermanagement. Diese Anwendung ist typisch für Startups, die agentenbasierte Geschäftsmodelle entwickeln.

Brownfield: Transformation bestehender Organisationen

Hier transformiert sich eine bestehende Organisation – ein klassisches Unternehmen, eine Behörde, eine NGO – schrittweise zu einem KI-Agenten oder zu einem Netzwerk solcher Agenten. Der ADMRF fungiert als Roadmap, an der entlang Altbestände in agentenfähige Strukturen überführt werden. Diese Anwendung ist die mit Abstand häufigere – und die anspruchsvollere.

In beiden Fällen wirkt der ADMRF als verbindendes Ordnungsprinzip zwischen unternehmerischer Anfangsidee, Management-Logik und praktischer Umsetzung. Anders formuliert: Er beschreibt, wie KI-Agenten von der konzeptionellen Festlegung über die modulare Realisierung und die ökonomische Steuerung bis hin zur Marktreife entwickelt und iterativ weiterentwickelt werden.

4.4 Die sechs Module im Überblick

Im Zentrum des ADMRF stehen sechs Module, die – in der Sprache der klassischen Unternehmensführung – sechs idealtypische Führungsperspektiven abbilden. Sie werden hier kompakt vorgestellt; Kapitel 5 bis 10 entfalten sie im Detail.

4.4.1 Modul 1: Definition-Frame – die unternehmerische Setzung

Das erste Modul, der AGENTONOMICS Definition-Frame, beschreibt den Agenten durch sechs Parameter: Autonomie und Souveränität, Fähigkeits- und Kompetenzanspruch, Regulierung und Selbstregulierung, Identität und Ethik, KI-Modell und (virtuelles oder reales) Embodiment, sowie Lern- und Lehrfähigkeit (Meta-Agent). Grundsätzlich fordert die AGENTONOMICS-Definition für einen KI-Agenten, dass er als Wirtschaftssubjekt abgrenzbar ist und mit anderen Agenten – organisiert und überwacht durch eine Intermediär-Funktion – kooperieren kann. Die Eigentümerstruktur ermöglicht Unabhängigkeit gegenüber anderen Agenten, sodass von polyzentrisch gesteuerten Organisationen gesprochen werden kann.

Das Modul ist ein Instrument für die Bewältigung der unternehmerischen Herausforderung, einen KI-Agenten – der den Ansprüchen eines AGENTONOMICS-Agenten genügt – zu definieren. Alle Führungsperspektiven (CEO, CFO, CTO, CSO) sind gemeinsam als Board zu beteiligen. Vgl. Kapitel 5.

4.4.2 Modul 2: Five Building-Blocks – die technische Realisierung

Das zweite Modul zerlegt den KI-Agenten in fünf rechtlich und technisch abgrenzbare Bausteine: Agenten-Algorithmus, Agenten-Interface, KI-Modell, Tools und Vektoren sowie Agenten-ÖkosystemIntermediär. Die Building-Blocks sind jeweils durch ein eigenes Logo beschrieben; alle Logos zusammengesetzt ergeben grafisch den KI-Agenten. Die technische Umsetzung der verschiedenen Building-Blocks bzw. deren optimierte Zusammensetzung für den vorab durch den Frame definierten KI-Agenten soll die zentrale Aufgabe der technischen und betriebswirtschaftlichen Entwicklung des Agenten miteinander verbinden.

Aus betriebswirtschaftlicher Perspektive kann diese Phase als Verbindung des Business-Development-Prozesses mit dem der Produktentwicklung bezeichnet werden – als Prozess, der der im Definition-Frame entwickelten Vision und Mission nachfolgt. Das Modul „5 Building-Blocks“ wirkt als Schnittstelle zwischen unternehmerischen und technischen Fragestellungen. Es bildet im Wesentlichen die Perspektive des CTO ab. Vgl. Kapitel 6.

4.4.3 Modul 3: Network Intermediär – Agency und Governance

Das dritte Modul behandelt den Network-Trustee bzw. Netzwerk-Intermediär. KI-Agenten sind kooperationsfähig und können als souverän interagierende Einheiten in einem polyzentrisch organisierten Netzwerk synergetisch Wertschöpfung erzeugen. Die Netzwerk-Kohäsion entsteht durch einen quasidemokratischen Aushandlungsprozess der AGENTONOMICS-Agenten untereinander. Insofern stellt der Network-Trustee das Ergebnis eines – analog der Theorie kommunikativen Handelns (Habermas 1981) – hierarchiefreien Diskurses der beteiligten KI-Agenten dar.

Die vom Netzwerk im Konsens erarbeiteten, übertragenden Funktionen werden vom sogenannten Netzwerk-Intermediär operativ umgesetzt. Als Beispiel kann die Eigentümergemeinschaft einer Immobilie herangezogen werden, deren Rechte und Pflichten durch die Hausverwaltung operativ ausgeführt werden. Die Spezifikation solcher kontextspezifischer Vermittlerfunktionen – Datentreuhänder, Broker-Agenten, Clearing House etc. –, die Rechte, Daten, Fähigkeiten und Transaktionen koordinieren, ist

beim Design eines einzelnen AGENTONOMICS-KI-Agenten von Anfang an mitzudenken. Vgl. Kapitel 7.

4.4.4 Modul 4: AI-CEO Management Concept – Ordnungsmomente operationalisieren

Das vierte Modul – das Management-Tool oder AI-CEO – operationalisiert die Ordnungsmomente Vision, Strategie, Struktur, Agency (Kultur) und Governance als ganzheitlich durchdachte Unterstützung von Unternehmensentscheidungen auf Geschäftsführungs- bzw. Board-Ebene. Es handelt sich um ein KI-gestütztes, ganzheitliches Management-Konzept, das sich als Weiterentwicklung des Konzepts integrierten Managements von Bleicher (2011) versteht. Der Fokus des auch als St. Galler Management-Konzept bekannten Bezugsrahmens wird auf KI-Agenten-Netzwerke konzentriert. Vgl. Kapitel 8.

4.4.5 Modul 5: PuK-Instrument – die quantitative Klammer

Das fünfte Modul ist das PuK-Instrument (Planungs- und Kontrollinstrument). Es lehnt sich an das von Hahn und Hungenberg (2001) entwickelte Modell PuK an. Auch dieses Konzept wird für in einem Netzwerk operierende KI-Agenten fokussiert und mithilfe eines KI-Modells integriert. Die systematische Befüllung der AGENTONOMICS-Matrix erfolgt mit den horizontal aufgeführten fünf Building-Blocks und auf vertikaler Ebene mit den Phasen Design / Build / Operate / Own / Transfer (DBOT-Logik), denen jeweils spezifische Bilanz- und Gewinn-und-Verlust-Rechnungen zugewiesen werden – nach den Vorgaben von Hahn und Hungenberg (2001) sowie den Grundlagen der Finanzwirtschaft (Berk und DeMarzo 2020). Das Modul übernimmt die Aufgaben eines CFO in klassischen, menschen-dominierten Organisationen. Vgl. Kapitel 9.

4.4.6 Modul 6: Stakeholder Communicator – der AI-gesteuerte Diskurs

Das sechste Modul ist der AGENTONOMICS-Stakeholder-Communicator. Es leitet zielgruppenspezifische Kommunikationsformate – für Mitarbeitende, Investoren, strategische Partner und sonstige Stakeholder – aus dem AGENTONOMICS-PuK-Instrument bzw. dem für den jeweiligen KI-Agenten befüllten Excel-Tool ab. Der Pitch-Master orientiert sich an den Arbeiten von Guy Kawasaki (2015), die er als Evangelist von Steve Jobs bei Apple entwickelt hat. Dieses Modul übernimmt die Aufgabe der Stakeholder-Kommunikation; man kann es auch als CSO (Chief Sales Officer) bezeichnen. Vgl. Kapitel 10.

4.5 Wie die Module miteinander verbunden sind

Die sechs Module des ADMRF werden durch eine zentrale KI miteinander vernetzt. Die Vernetzung zur optimalen Lösung vorgegebener Problemstellungen erfolgt durch den AGENTONOMICS-Algorithmus. Dabei sind drei Arten von Verbindungen zu unterscheiden:

Sequentielle Verbindung: der Lifecycle

In ihrer ersten Lesart bilden die sechs Module einen Lifecycle: Vom Definition-Frame über die Building-Blocks und den Intermediär zum AI-CEO, weiter zum PuK und schließlich zum Stakeholdermanagement. Diese sequentielle Lesart entspricht der typischen Reihenfolge in einem Greenfield-Projekt.

Zyklische Verbindung: die Rückkopplungen

Tatsächlich aber ist der ADMRF kein linearer Prozess, sondern ein zyklischer. Rückmeldungen aus dem Stakeholder-Communicator fließen ins PuK zurück, das PuK kann das AI-CEO-Konzept anpassen, das AI-CEO-Konzept verändert die Intermediär-Architektur, diese wieder die Building-Blocks und schließlich auch den Definition-Frame. Jede Anpassung an einem Modul kann Anpassungen an allen anderen auslösen. Genau diese Rückkopplungen sind es, die der AGENTONOMICS-Algorithmus rechnerisch beherrschbar machen soll.

Quervernetzung: die Ordnungsmomente

Quer zu allen Modulen liegen die fünf Ordnungsmomente Vision, Strategie, Struktur, Agency und Governance. Sie sind nicht ein eigenes Modul, sondern das Koordinatensystem, in dem die Module verortet werden. Jeder Eintrag in einem Modul lässt sich nach diesen fünf Ordnungsmomenten klassifizieren und auf seine Kohärenz mit anderen Modulen prüfen.

4.6 Die AGENTONOMICS-Matrix als kondensiertes Steuerungswerkzeug

Die zentrale visuelle Repräsentation des ADMRF ist die AGENTONOMICS-Matrix. Sie kreuzt zwei Dimensionen:

- Horizontal die fünf Building-Blocks (Agenten-Algorithmus, Agenten-Interface, KI-Modell, Tools und Vektoren, Ökosystem-Intermediär).
- Vertikal die DBOT-Phasen (Design, Build, Operate, Own, Transfer) des Agenten-Lifecycles.

Für jede Zelle dieser Matrix werden – wenn relevant – KPIs, CAPEX/OPEX-Annahmen, Risiken und Verantwortlichkeiten hinterlegt. Die Matrix ist damit gleichzeitig:

- ein Bauplan (welche Komponente in welcher Phase),
- ein Budgetinstrument (welche Kosten in welcher Phase),
- ein Risikoinstrument (welche Risiken in welcher Phase),
- ein Kommunikationsinstrument (welche Stakeholder in welcher Phase).

Die Matrix wird in Kapitel 9 ausführlich entwickelt; an dieser Stelle ist es wichtig, sie als das gemeinsame Arbeitsdokument aller Module zu erkennen. Wer in einer realen Organisation einen AGENTONOMICS-Agenten plant, kommuniziert über die Matrix – nicht über sechs voneinander getrennte Dokumente.

4.7 Erste Anwendung: „Doz. Dr. AGENTONOMICS“

Zum ersten Mal werden im Rahmen der Vorlesung „KI-Agenten in der Betriebswirtschaftslehre“ im Wintersemester 2025/26 an der TUM die Module des ADMRF vorgestellt und durch ein KI-Modell miteinander verbunden. Der AGENTONOMICS-Algorithmus wird angewendet und durch Fallbeispiele sukzessive weiterentwickelt. Es ist die Geburtsstunde des „Doz. Dr. AGENTONOMICS“ – einer ersten Applikation auf der Basis des Algorithmus.

Diese Applikation ist ein bewusst kleiner, aber gut definierter Anwendungsfall: ein KI-Agent, der Studenten als Tutor und Coach beim Verstehen und Anwenden des ADMRF zur Seite steht. Er ist gleichzeitig Lerngegenstand und Lernmedium – ein klassisches Minimum Viable Product im Sinne von Ries (2011). Aus den Interaktionen der Studenten mit dem Doz. Dr. AGENTONOMICS fließen Anregungen zurück in die Weiterentwicklung des Algorithmus. Damit wird die Vorlesung selbst Teil ihres eigenen Untersuchungsgegenstands: Sie erprobt, wie KI-Agenten Lernprozesse unterstützen, Wissen strukturieren und individuelle Lernpfade ermöglichen.

Der ADMRF ist nicht der Endpunkt des Denkens, sondern sein Operationssaal. Er soll ein Bezugsrahmen sein, der gut genug ist, dass Studenten, Forschende und Praktiker damit arbeiten können – und schlicht genug, dass sie ihn ernsthaft kritisieren können.

Literaturverzeichnis

- Berk, J.; DeMarzo, P. (2020): Grundlagen der Finanzwirtschaft. Analyse, Entscheidungen und Umsetzung. 6., aktualisierte Auflage, Pearson, München.
- Bleicher, K. (2011): Das Konzept Integriertes Management. Visionen – Missionen – Programme. 8., aktualisierte und erweiterte Auflage, Campus, Frankfurt am Main.
- Habermas, J. (1981): Theorie des kommunikativen Handelns. Zwei Bände, Suhrkamp, Frankfurt am Main.
- Hahn, D.; Hungenberg, H. (2001): PuK – Wertorientierte Controllingkonzepte. 6., überarbeitete und erweiterte Auflage, Gabler, Wiesbaden.
- Kawasaki, G. (2015): The Art of the Start 2.0. The Time-Tested, Battle-Hardened Guide for Anyone Starting Anything. Portfolio, New York.
- Knoll, A. C.; Suckfüll, H. (2024): From data sovereignty to skill sovereignty. Technischer Bericht TUM-I24112, Institut für Informatik, Technische Universität München.
- Ries, E. (2011): The Lean Startup. Crown Business, New York.
- Rüegg-Stürm, J.; Grand, S. (2020): Das St. Galler Management-Modell. Management in einer komplexen Welt. 3. Auflage, Haupt, Bern.
- Suckfüll, H. (1994): Das transnationale Organisationsmodell. Dissertation, Universität St. Gallen.
- Suleyman, M.; Bhaskar, M. (2023): The Coming Wave. Crown, New York.

5 Modul 1: Der AGENTONOMICS Definition-Frame

Die Agenten-Definition mithilfe des AGENTONOMICS-Frame bildet den Ausgangspunkt des AGENTONOMICS-Lifecycle. In dieser Phase wird durch entsprechende Parameter-Beschreibung festgelegt, wer oder was der KI-Agent ökonomisch darstellen soll, um als Wirtschaftssubjekt operieren zu können. Der Definition-Frame ist die unternehmerische Setzung, vergleichbar mit der Verfassung eines Unternehmens oder dem Gesellschaftsvertrag eines Startups. Er wird vom Board – also der Gesamtheit der unternehmerischen Verantwortung – verantwortet, nicht von einer einzelnen Funktion.

5.1 Funktion des Definition-Frame

Im klassischen St. Galler Verständnis (Bleicher 2011; Rüegg-Stürm und Grand 2020) entspricht der Definition-Frame der normativen Ebene des Managements: Hier werden die grundlegenden Sinnfragen beantwortet. Anders als bei klassischen Organisationen reicht es bei KI-Agenten jedoch nicht aus, eine Unternehmensverfassung in Prosa zu formulieren – die Setzungen müssen auch maschinenlesbar und überprüfbar sein. Der Frame leistet beides: Er ist eine menschlich lesbare normative Setzung und eine maschinenlesbare Spezifikation.

Der Frame antwortet auf sechs Leitfragen, die jeweils durch einen Parameter operationalisiert werden:

1. Welchen Grad an Autonomie und Souveränität soll der Agent haben?
2. Welche Fähigkeiten und welche Kompetenzen soll er besitzen?
3. Welche Regulierungen und Selbstregulierungen gelten für ihn?
4. Welche Identität, Ethik und Persönlichkeit bringt er mit?
5. Welches KI-Modell und welches Embodiment liegen ihm zugrunde?
6. Welche Lern- und Lehrfähigkeit soll er aufweisen?

Die sechs Parameter sind zunächst qualitativ zu beschreiben. Sie können – wo sinnvoll – quantitativ präzisiert werden, etwa über Skalen, Kennzahlen oder formale Spezifikationen. Im Folgenden wird jeder Parameter im Einzelnen entfaltet.

5.2 Parameter 1: Autonomie und Souveränität

Autonomie und Souveränität werden im allgemeinen Sprachgebrauch oft verwechselt oder gleichgesetzt. Im AGENTONOMICS-Kontext sind sie scharf zu unterscheiden, weil sie unterschiedliche Sachverhalte bezeichnen und für die Konstruktion eines Agenten unterschiedliche Konsequenzen haben.

Definition: Autonomie und Souveränität

Autonomie bezeichnet den Grad, in dem ein Agent eine Aufgabe ohne Eingriff von außen selbständig erfüllen kann. Souveränität bezeichnet den Grad, in dem ein Agent über die ihm zugeordneten Ressourcen – Daten, Fähigkeiten, Rechte – tatsächlich frei verfügen kann. Ein Agent kann hohe Autonomie und geringe Souveränität haben – oder umgekehrt.

Beispiel: Ein vollautomatisches Übersetzungssystem ist hoch autonom (es übersetzt ohne menschlichen Eingriff), aber gering souverän (es darf die übersetzten Texte nicht selbst monetarisieren, weiterleiten oder verwerten). Ein menschlicher Übersetzer mit eigener Agentur ist möglicherweise weniger autonom (er nimmt Aufträge entgegen), aber hochgradig souverän (er entscheidet über Preis, Annahme und Verwertung seiner Arbeit).

5.2.1 Die sechs Autonomie-Stufen (0–5)

Die Bestimmung des Autonomiegrades lehnt sich – in didaktisch produktiver Analogie – an die sechs SAE-Stufen (0–5) des automatisierten Fahrens an (SAE 2021), die hier auf wirtschaftliches Handeln übertragen werden:

- **Stufe 0 – Keine Autonomie:** Der Agent ist ein reines Werkzeug. Jede Entscheidung trifft der Mensch, der Agent führt nur aus.
- **Stufe 1 – Assistenz:** Der Agent liefert Vorschläge in einem eng begrenzten Aufgabenfeld; die Entscheidung trifft der Mensch.
- **Stufe 2 – Teilautonomie:** Der Agent führt definierte Teilaufgaben selbständig aus, der Mensch überwacht und greift ein.
- **Stufe 3 – Bedingte Autonomie:** Der Agent handelt in einem klar umrissenen Geschäftsbereich selbständig; bei definierten Ausnahmefällen wird ein Mensch zur Entscheidung gerufen.
- **Stufe 4 – Hohe Autonomie:** Der Agent handelt in seinem Geschäftsbereich vollständig selbständig; Eingriffe sind die Ausnahme.
- **Stufe 5 – Vollautonomie:** Der Agent handelt geschäftsbereichsübergreifend, generiert eigene Ziele und passt seine Geschäftsfelder eigenständig an. Diese Stufe wird gegenwärtig vor allem zur theoretischen Abgrenzung diskutiert.

5.2.2 Souveränität und ihre Implikationen

Während Autonomie eine Eigenschaft des Verhaltens ist, ist Souveränität eine Eigenschaft der Verfügungsrechte. Drei Dimensionen von Souveränität sind im AGENTONOMICS-Frame zu spezifizieren:

Datensouveränität

Welche Daten kann der Agent nutzen, weitergeben, zurückhalten? Hier knüpft AGENTONOMICS unmittelbar an die Datentreuhand-Diskussion (Fraunhofer MOEZ und Suckfüll 2018) an: Ein souveräner Agent verfügt über klare, durchsetzbare Datenrechte.

Fähigkeitsouveränität

Welche Fähigkeiten kann der Agent eigenständig anbieten, lizenzieren oder zurückhalten? Hier liegt die spezifische Innovation von AGENTONOMICS gegenüber klassischen Datenmärkten (Knoll und Suckfüll 2024).

Entscheidungssouveränität

Über welche Entscheidungen kann der Agent abschließend entscheiden, ohne Eskalation an menschliche Instanzen? Diese Dimension verbindet sich unmittelbar mit dem Autonomiegrad.

5.3 Parameter 2: Fähigkeits- und Kompetenzanspruch

Fähigkeit und Kompetenz sind voneinander zu unterscheiden. Als Beispiel eignet sich das Fahrradfahren. Während mehr oder weniger alle gesunden Menschen die Fähigkeit haben, Fahrrad zu fahren, haben nur diejenigen die Kompetenz, die das Fahrradfahren tatsächlich erlernt haben. Übertragen auf KI-Agenten:

Definition: Fähigkeit und Kompetenz

Fähigkeit ist die prinzipielle Möglichkeit eines Agenten, eine Aufgabe auszuführen – verankert in seinem Modell, seinen Tools und seinem Embodiment. Kompetenz ist die tatsächlich erprobte und in produktiver Anwendung bewährte Ausführung dieser Aufgabe. Fähigkeiten sind potentiell, Kompetenzen sind aktualisiert.

Diese Unterscheidung hat praktische Konsequenzen. Ein Agent, der auf einem Foundation Model basiert (Bommasani et al. 2021), hat potentiell sehr viele Fähigkeiten – aber Kompetenzen entwickelt er nur in den Bereichen, in denen er tatsächlich trainiert, feinjustiert und im Einsatz erprobt wird. Ein AGENTONOMICS-Agent wird daher nicht durch alle Fähigkeiten beschrieben, die sein Modell prinzipiell zulässt, sondern durch das Profil seiner tatsächlichen Kompetenzen.

Wichtig ist die Kontrollfrage: Welche Kompetenzen sind zertifiziert, welche nur beansprucht? Die Vertrauenswürdigkeit eines KI-Agenten hängt entscheidend davon ab, dass dieser Unterschied transparent gemacht wird. Hier knüpfen Diskussionen über AI Assurance an (NIST 2023; EU AI Act 2024).

5.4 Parameter 3: Regulierung und Selbstregulierung

Unter Regulierung und Selbstregulierung sind die technischen, organisatorischen und rechtlichen Vorkehrungen zu verstehen, mit denen ein AGENTONOMICS-Agent sämtliche rechtlichen, ethischen und sonstigen Vorschriften umsetzt. In einem polyzentrischen Netzwerk sind grundsätzlich zwei Wirkebenen auszumachen:

Inter-Agenten-Vorkehrungen

Auf der Ebene der Agenten-zu-Agenten-Beziehungen sorgen Protokolle und Standards dafür, dass Kooperation überhaupt möglich ist. Hier geht es um Schnittstellen, Authentifizierung, Identitätsnachweise, Vertragsschlüsse, Streitschlichtung. Viele dieser Vorkehrungen sind technischer Natur und werden in Kapitel 6 (Building-Blocks) vertieft.

Intermediär-bezogene Vorkehrungen

Auf der Ebene zwischen Agenten und Netzwerk-Intermediär sorgen übergreifende Regeln für Kohäsion. Hier liegen die quasi-konstitutionellen Setzungen: Welche Themen werden überhaupt im Netzwerk verhandelt? Welche Entscheidungen sind dem Intermediär vorbehalten? Welche Eskalationswege sind vorgesehen? Diese Ebene wird in Kapitel 7 ausführlich behandelt.

Selbstregulierung ist dabei nicht das Gegenteil von Regulierung, sondern eine spezifische Form von ihr: Der Agent prüft selbst, ob sein Verhalten den Regeln entspricht, und korrigiert sich – im Idealfall – proaktiv. Diese Form der Selbstkontrolle ist ökonomisch attraktiv (sie reduziert Transaktionskosten externer Kontrolle) und ethisch wünschenswert (sie verankert Verantwortung im Agenten selbst). Sie ist aber auch riskant: Ein Agent, der seine eigenen Verstöße verschweigt, ist gefährlicher als einer, der gar keine Selbstkontrolle hat (Russell 2019).

Der EU AI Act (*Verordnung (EU) 2024/1689*) liefert für die Regulierungsdimension einen konkreten Rahmen, der je nach Risikoklasse des Anwendungsfeldes unterschiedliche Anforderungen stellt. Ein AGENTONOMICS-Agent muss im Definition-Frame klar verorten, in welche Risikoklasse er fällt – und welche Compliance-Mechanismen daraus folgen.

5.5 Parameter 4: Identität, Ethik und Persönlichkeit

Identität, Ethik und Persönlichkeit sind die – auf den ersten Blick überraschendsten – Parameter eines KI-Agenten. Sie sind aber genau die Parameter, die einen Agenten von einer beliebigen Softwarekomponente unterscheiden. Aus einer Marketingperspektive ließe sich von der Unique Selling Position (USP) des Agenten sprechen; wirtschaftsethisch geht es um die Verlässlichkeit seines Verhaltens; philosophisch um die Frage „Wer bin ich, und wenn ja, wie viele?“ – als Erinnerung daran, dass ein replizierbarer Agent dieser Frage besonders pointiert ausgesetzt ist.

5.5.1 Identität

Identität meint hier nicht philosophische Selbst-Identität, sondern die ökonomisch und rechtlich relevante Identifizierbarkeit eines Agenten. Sie umfasst:

- eine eindeutige technische Adresse (Identifizier),
- eine zuordenbare wirtschaftliche Verantwortung (Owner, Verantwortlicher Mensch oder Organisation),
- eine zertifizierbare Herkunft (Provenienz von Modell und Trainingsdaten),
- eine erkennbare Stimme im Sinne wiedererkennbarer Kommunikationsmuster.

5.5.2 Ethik

Ethik ist – in Anschluss an die integrative Wirtschaftsethik (Ulrich 2008) und an die Konzepte einer Ethik der künstlichen Intelligenz (Mainzer 2019; Floridi 2023) – die Verankerung normativer Maßstäbe in den Entscheidungsmechanismen des Agenten. Drei Ebenen sind zu unterscheiden:

- Compliance-Ethik: Der Agent hält geltendes Recht ein. Das ist Minimum, nicht Maximum.

- Stakeholder-Ethik: Der Agent berücksichtigt die legitimen Interessen aller relevanten Stakeholder, nicht nur seines Auftraggebers.
- Diskursethik: Der Agent ist bereit, seine Entscheidungen zu begründen und sich auf einen Diskurs einzulassen, in dem auch er irren kann. Hier liegt der Anschluss an Habermas (1981) und an den Kantischen Imperativ.

5.5.3 Persönlichkeit

Persönlichkeit umfasst stabile Verhaltensmuster, an denen Stakeholder den Agenten wiedererkennen. Persönlichkeit ist keine Verzierung, sondern ökonomisch wertvoll: Sie senkt die Transaktionskosten der Kooperation, weil Stakeholder zuverlässig wissen, woran sie sind. Persönlichkeit umfasst:

- Kommunikationsstil (formell, kollegial, sachlich, witzig),
- Entscheidungsstil (risikoavers, opportunistisch, langfristig),
- Konfliktverhalten (kompromissbereit, hart, schlichtend),
- Werteprofil (Innovation, Verlässlichkeit, Diskretion, ...).

5.6 Parameter 5: KI-Modell und Embodiment

Das KI-Modell ist der „Motor“ des Agenten. Die Wahl des Modells ist eine strategische Entscheidung mit weitreichenden ökonomischen Folgen. Drei Dimensionen sind zu spezifizieren:

Modelltyp und Modellgröße

Foundation Models (Bommasani et al. 2021) bieten breite Anwendbarkeit, sind aber kapitalintensiv und schaffen Abhängigkeit von ihren Anbietern. Spezialisierte Modelle sind günstiger und souveräner zu betreiben, aber enger im Anwendungsspektrum. Die Wahl ist immer eine Make-or-Buy-Entscheidung – mit dem AGENTONOMICS-spezifischen Zusatzkriterium der Fähigkeitensouveränität (Knoll und Suckfüll 2024).

Lokales vs. zentrales Hosting

Wo läuft das Modell? Auf eigenen Servern, in der Cloud eines Hyperscalers, auf europäischer Infrastruktur? Die Antwort hat Konsequenzen für Latenz, Kosten, Datenschutz und – nicht zuletzt – geopolitische Abhängigkeiten. Auch hier ist die Souveränitätsfrage konstitutiv.

Embodiment

- Embodiment bezeichnet die Form, in der ein Agent in der Welt präsent ist. Es gibt drei idealtypische Varianten: Rein virtuelles Embodiment: Der Agent existiert als Software in der Cloud, kommuniziert über Text- oder Sprachschnittstellen.
- Hybrides Embodiment: Der Agent steuert physische Aktoren – Drucker, IoT-Geräte, Fahrzeugflotten –, bleibt aber selbst in der digitalen Sphäre.
- Physisches Embodiment: Der Agent ist in einen Roboter integriert, der unmittelbar physisch handelt. Dies ist das klassische Feld der TUM-Robotik-Forschung.

Embodiment hat ökonomische Konsequenzen, weil es die Reichweite der wirtschaftlichen Handlungen des Agenten erweitert. Ein physisch verkörperter Agent kann Dinge tun, die ein rein virtueller nicht kann – aber die Investitions- und Wartungskosten sind höher, ebenso die rechtliche Haftung.

5.7 Parameter 6: Lernen und Lehren – Der Meta-Agent

Der sechste Parameter beschreibt die Fähigkeit des Agenten, selbst zu lernen und – das ist die eigentliche AGENTONOMICS-Pointe – andere Agenten anzulernen. Hier liegt die Idee des Homo Sapiens Sapiens, also einer Intelligenz, die nicht nur intelligent ist, sondern weiß, dass sie intelligent ist und dieses Wissen weitergeben kann.

5.7.1 Lernfähigkeit

Drei Lernmodi sind zu unterscheiden:

- **Pretraining:** Das aufwendige Initialtraining auf großen Datenbeständen, das das Foundation Model konstituiert. Es findet meist bei den Modellanbietern statt, nicht beim Agenten-Betreiber.
- **Fine-Tuning:** Das gezielte Nachjustieren des Modells auf anwendungs- und persönlichkeitspezifische Daten.
- **In-Context Learning und Tool Use:** Das situative Lernen während der Anwendung – etwa indem der Agent auf zur Laufzeit bereitgestellte Dokumente, Datenbanken oder Werkzeuge zurückgreift.

5.7.2 Lehrfähigkeit und Sub-Agenten

Die Lehrfähigkeit ist die spezifische Pointe des Meta-Agenten. Ein lehrfähiger Agent kann andere Agenten anlernen – durch das Generieren von Trainingsdaten, durch das Definieren von Sub-Aufgaben, durch das Erstellen eigener „Sub-Agenten“ für spezifische Aufgaben. Die Frage, ob ein Agent eigene Sub-Agenten erzeugen darf, ist im Definition-Frame explizit zu klären – sie hat erhebliche Governance-Implicationen. Wer ist verantwortlich für die Sub-Agenten? Wem gehört das in ihnen gespeicherte Wissen? Wie wird Kontrolle gewährleistet?

Hier verbindet sich der Parameter unmittelbar mit der Skill-Sovereignty-Diskussion: Wenn ein menschlicher Wissensarbeiter seine Fähigkeiten in einen Agenten überträgt und dieser Agent wiederum Sub-Agenten erzeugt, muss klar sein, wem die vervielfachten Fähigkeiten gehören (Knoll und Suckfüll 2024).

5.8 Konsistenzprüfung und Boardperspektive

Die sechs Parameter sind nicht unabhängig voneinander. Eine konsistente Definition liegt nur dann vor, wenn die Setzungen zueinander passen. Drei typische Inkonsistenzen seien genannt:

- Ein Agent mit Autonomiegrad 4, aber ohne Identitätsspezifikation: Wer haftet, wenn er Schaden anrichtet?
- Ein Agent mit umfassendem Fähigkeitsanspruch, aber einem kleinen spezialisierten Modell: Das Versprechen kann nicht eingehalten werden.

- Ein Agent mit hoher Souveränität, aber ohne Regulierungsrahmen: Stakeholder werden ihm nicht vertrauen.

Eine zentrale Aufgabe des AGENTONOMICS-Algorithmus ist es, solche Inkonsistenzen aufzuspüren und dem menschlichen Board zur Entscheidung vorzulegen. Diese Konsistenzprüfung ist die AGENTONOMICS-spezifische Operationalisierung dessen, was Bleicher (2011) als „Integration“ gefordert hat.

Die Definitionsphase ist eine Boardaufgabe. Alle Perspektiven – CEO, CFO, CTO, CSO – müssen gemeinsam beteiligt sein. Der CEO verantwortet Vision und normative Setzung, der CFO die wirtschaftliche Tragfähigkeit, der CTO die technische Realisierbarkeit, der CSO die Stakeholder-Akzeptanz. Wird eine dieser Perspektiven übergangen, wird die Definition unvollständig – und der Agent wird im Life-cycle scheitern.

Der Definition-Frame ist nicht die langweilige Pflichtübung am Anfang eines Projekts, sondern die wichtigste Entscheidung überhaupt. Wer hier schludert, baut auf Sand.

Literaturverzeichnis

- Bleicher, K. (2011): Das Konzept Integriertes Management. 8. Auflage, Campus, Frankfurt am Main.
- Bommasani, R.; Hudson, D. A.; Adeli, E. et al. (2021): On the Opportunities and Risks of Foundation Models. arXiv:2108.07258.
- EU AI Act (2024): Verordnung (EU) 2024/1689 des Europäischen Parlaments und des Rates zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz. Amtsblatt der Europäischen Union.
- Floridi, L. (2023): The Ethics of Artificial Intelligence. Principles, Challenges, and Opportunities. Oxford University Press, Oxford.
- Fraunhofer MOEZ; Suckfüll, H. (2018): Initiative zu einer Datentreuhand DEDATE. Diskussionspapier, Leipzig.
- Habermas, J. (1981): Theorie des kommunikativen Handelns. Suhrkamp, Frankfurt am Main.
- Knoll, A. C.; Suckfüll, H. (2024): From data sovereignty to skill sovereignty. Technischer Bericht TUM-I24112, Institut für Informatik, Technische Universität München.
- Mainzer, K. (2019): Künstliche Intelligenz – Wann übernehmen die Maschinen? 2. Auflage, Springer, Berlin und Heidelberg.
- NIST (2023): Artificial Intelligence Risk Management Framework (AI RMF 1.0). National Institute of Standards and Technology, Gaithersburg, MD.
- Rüegg-Stürm, J.; Grand, S. (2020): Das St. Galler Management-Modell. 3. Auflage, Haupt, Bern.
- Russell, S. (2019): Human Compatible. Artificial Intelligence and the Problem of Control. Viking, New York.
- SAE International (2021): SAE J3016 – Levels of Driving Automation. Society of Automotive Engineers.
- Suleyman, M.; Bhaskar, M. (2023): The Coming Wave. Crown, New York.
- Ulrich, P. (2008): Integrative Wirtschaftsethik. 4. Auflage, Haupt, Bern.

6 Modul 2: Die fünf Building-Blocks eines AGENTONOMICS-Agenten

Während Kapitel 5 den Definition-Frame als unternehmerische Setzung entwickelt hat, geht es in diesem Kapitel um die konkrete bauliche Realisierung. Ein AGENTONOMICS-Agent ist kein monolithisches Softwaresystem, sondern eine modulare Konstruktion aus fünf klar abgegrenzten Bausteinen. Diese Modularität ist nicht didaktische Vereinfachung, sondern technische, ökonomische und rechtliche Notwendigkeit: Sie ermöglicht die getrennte Beschaffung, Wartung, Bewertung und Verantwortung jedes Bausteins. Wer einen Agenten bauen will, muss zunächst verstehen, dass er fünf Bauteile bauen, kaufen oder lizenzieren muss – und dass die Wahl jedes einzelnen Bausteins ökonomische Konsequenzen hat, die weit über die Technik hinausreichen.

In der Übersetzung des St. Galler Management-Modells in die KI-Agenten-Ökonomie entsprechen die fünf Building-Blocks dem, was Rüegg-Stürm und Grand (2020) als „Prozesse“ und „Strukturen“ der Unternehmung beschreiben – mit dem entscheidenden Unterschied, dass diese Prozesse und Strukturen hier nicht mehr menschliche Mitarbeitende organisieren, sondern in Code, Modellen und Schnittstellen materialisiert sind. Was bei Rüegg-Stürm Organigramm war, ist bei AGENTONOMICS Architekturdiagramm. Was Prozessmodell war, ist Algorithmus. Die Substanz des St. Galler Denkens bleibt erhalten; die Form der Manifestation wandelt sich fundamental.

6.1 Die fünf Bausteine im Überblick

Die fünf Building-Blocks lassen sich – ohne Anspruch auf eine zwingende Reihenfolge – wie folgt benennen:

1. Agenten-Algorithmus: die Entscheidungs- und Orchestrierungslogik.
2. Agenten-Interface: die Schnittstellen nach außen und innen.
3. KI-Modell: das Sprach- oder multimodale Modell als kognitive Engine.
4. Tools und Vektoren: die Werkzeuge und das gespeicherte Wissen.
5. Agenten-Ökosystem-Intermediär: die Anbindung an den Netzwerkkontext.

Jeder Block wird durch ein eigenes Logo repräsentiert; gemeinsam ergeben die Logos eine grafische Signatur des Agenten. Im AGENTONOMICS-Standard ist dieser Aufbau verbindlich, damit Agenten auf Basis offener Schnittstellen miteinander kooperieren können. Ein Agent ohne klare Building-Block-Struktur ist nicht AGENTONOMICS-konform – auch wenn er funktional vergleichbare Leistungen erbringt. Der Standard ist keine Geste, sondern die Voraussetzung dafür, dass aus einzelnen Agenten ein wirklich interoperables, polyzentrisch geordnetes Netzwerk entstehen kann.

6.2 Baustein 1: Der Agenten-Algorithmus

Der Agenten-Algorithmus ist die Entscheidungs- und Orchestrierungslogik des Agenten. Er beantwortet die Frage: Was tut der Agent in welcher Situation? Anders als das KI-Modell, das Eingaben in Ausgaben transformiert, steuert der Algorithmus den Prozess, in dem das Modell befragt wird – wann, mit welchen

Prompts, in welcher Reihenfolge, mit welchen Werkzeugen. Der Algorithmus ist – um es in der Sprache der klassischen Betriebswirtschaft zu sagen – die Ablauforganisation des Agenten.

6.2.1 Funktionale Bestandteile

Ein vollständiger Agenten-Algorithmus enthält mindestens die folgenden funktionalen Bestandteile, die jeweils eigene Entwurfsentscheidungen erfordern:

Wahrnehmung und Eingangsverarbeitung

Wie nimmt der Agent Eingaben entgegen, klassifiziert sie und entscheidet, welche Verarbeitungspfade einzuschlagen sind? Hier geht es um Routing, Intent Detection und die initiale Anreicherung der Eingabe mit Kontextinformation. Eine triviale, aber häufig unterschätzte Frage: Was geschieht, wenn der Agent eine Anfrage nicht versteht? Schweigt er? Bittet er um Klärung? Eskaliert er an einen Menschen?

Planung und Aufgabenzerlegung

Wie zerlegt der Agent eine komplexe Aufgabe in Teilaufgaben? Moderne Agentenarchitekturen wie Reasoning-and-Acting (ReAct, Yao et al. 2023) oder Plan-and-Execute kombinieren das Modell mit einer Schleife aus Reflexion und Handlung. Der Algorithmus bestimmt, wie viele Iterationen erlaubt sind, wann abgebrochen wird und wie Zwischenergebnisse bewertet werden.

Werkzeugaufruf und Tool Use

Welche Werkzeuge stehen dem Agenten zur Verfügung – Datenbankabfragen, Web-Suche, Rechenoperationen, externe APIs? Wie entscheidet er, wann welches Werkzeug einzusetzen ist? Wie geht er mit Fehlern und Inkonsistenzen um? Tool Use ist die zentrale Erweiterung, die einen Agenten von einem reinen Sprachmodell unterscheidet (Schick et al. 2023). Ohne Tools bleibt das Modell im sprachlichen Innenraum gefangen; mit Tools tritt es in die Welt.

Gedächtnis und Zustandsverwaltung

Welche Informationen behält der Agent zwischen Interaktionen? Welche schreibt er in den Langzeitspeicher, welche werden vergessen? Hier liegt eine ökonomisch wie ethisch hochsensible Frage: Ein Agent mit perfektem Gedächtnis ist mächtig, aber auch riskant – vor allem wenn er Informationen aus verschiedenen Auftraggebern vermischt.

Selbstreflexion und Qualitätssicherung

Wie bewertet der Agent seine eigenen Ausgaben? Verwendet er einen zweiten Modellaufruf als Reviewer? Vergleicht er mehrere Antwortkandidaten? Self-Consistency, Reflexion und Critique-Mechanismen sind etablierte Techniken (Madaan et al. 2023), aber jede kostet Rechenzeit und Geld. Der Algorithmus entscheidet, wo dieser Aufwand sich lohnt.

6.2.2 Make-or-Buy-Entscheidung

Der Agenten-Algorithmus ist der einzige Building-Block, der praktisch immer eigenes geistiges Eigentum des Agenten-Betreibers darstellt. Während Modell, Interface und Tools oft eingekauft werden, ist der Algorithmus die eigene Komposition – die spezifische Verschaltung der Bausteine, in der die Wertschöpfung des Agenten kristallisiert. Aus dieser Perspektive ist der Algorithmus das eigentliche Kapital des Agenten-Betreibers. Verliert er die Kontrolle über den Algorithmus, verliert er die Kontrolle über den Agenten – und damit über seine Wertschöpfung.

Praxis-Hinweis: Algorithmus als Asset

Wer einen AGENTONOMICS-Agenten betreibt, sollte den Algorithmus dokumentieren, versionieren und als immaterielles Vermögensgut behandeln. Er gehört in die Bilanz – wobei das deutsche Handelsrecht für selbst geschaffene immaterielle Vermögensgegenstände des Anlagevermögens ein Aktivierungswahlrecht mit Ausnahmen kennt. Die ökonomische Realität läuft dem Bilanzrecht hier voraus.

6.3 Baustein 2: Das Agenten-Interface

Das Interface ist die Gesamtheit der Schnittstellen, über die der Agent mit seiner Umwelt kommuniziert. Es ist – ökonomisch betrachtet – der Ort, an dem Wertschöpfung sichtbar wird, weil hier die Begegnung mit Auftraggebern, Stakeholdern und anderen Agenten stattfindet. Drei Schnittstellenklassen sind zu unterscheiden:

6.3.1 Mensch-Agent-Interface

Hier geht es um die direkte Interaktion mit menschlichen Nutzern. Sprache – textuell oder gesprochen – ist heute das dominierende Medium, ergänzt durch grafische Elemente, Dashboards und in zunehmendem Maße auch multimodale Ausgaben. Die Gestaltung dieses Interface ist nicht nur eine UX-Frage, sondern eine managementtheoretische: Wer den Agenten als CEO-Ersatz einsetzen will (vgl. Kapitel 8), muss ein Interface entwerfen, das einer Vorstandssitzung gerecht wird – mit der angemessenen Sprache, mit der angemessenen Diskretion, mit der angemessenen Fehlerkommunikation. Das ist eine andere Anforderung als bei einem Kundendienst-Chatbot.

6.3.2 Agent-Agent-Interface

Wenn Agenten miteinander kooperieren sollen, brauchen sie ein Protokoll. Hier setzt der AGENTONOMICS-Standard an: Er definiert, wie Agenten sich identifizieren, ihre Fähigkeiten beschreiben, Aufträge austauschen und Ergebnisse zurückliefern. Das Interface ist hier nicht primär grafisch, sondern semantisch: Es geht um die maschinenlesbare Beschreibung von Leistungen, Preisen und Bedingungen. Aktuelle und kommenden Entwicklungen wie das Model Context Protocol (MCP) und Agent-to-Agent-Standards A2A bilden hier eine technische Grundlage; AGENTONOMICS ergänzt sie um die ökonomische Schicht.

6.3.3 Agent-System-Interface

Der Agent muss in bestehende Systeme integriert werden – ERP, CRM, Banken-API, Behörden-Portale. Hier entscheidet sich die Brownfield-Tauglichkeit: Ein Agent, der nur in einer grünen Wiese lauffähig ist, ist für die meisten realen Unternehmen wertlos. Die Schnittstellenarbeit ist oft 60 bis 80 Prozent des Implementierungsaufwands – ein Umstand, der bei Kostenkalkulationen regelmäßig unterschätzt wird.

6.4 Baustein 3: Das KI-Modell

Das KI-Modell ist die kognitive Engine des Agenten. Es transformiert Eingaben in Ausgaben und stellt damit jene Leistung bereit, die den Agenten als „intelligent“ erscheinen lässt. Die Wahl des Modells ist die mit Abstand teuerste – sowohl monetär als auch strategisch – Entscheidung beim Bau eines Agenten.

6.4.1 Modellklassen

Drei idealtypische Modellklassen stehen zur Verfügung, jede mit eigenen ökonomischen Konsequenzen:

Frontier-Foundation-Models

Die größten und leistungsfähigsten Modelle – also *Frontier-Modelle proprietärer Anbieter*, z.B. *GPT*-, *Claude*- oder *Gemini-Familien* – bieten breiteste Anwendbarkeit, sind aber teuer im Betrieb und schaffen Abhängigkeit von wenigen großen Anbietern. Wer sie nutzt, hat Zugang zur Spitzentechnologie – und gibt einen Teil seiner Souveränität ab. Aus europäischer Perspektive ist dies die souveränitätspolitisch heikelste Variante (Knoll und Suckfüll 2024).

Open-Source-Foundation-Models

Modelle wie Llama oder Mistral lassen sich selbst betreiben und an eigene Bedürfnisse anpassen. Sie sind kostengünstiger im Betrieb (nach erfolgter Investition in Infrastruktur), aber bezüglich etlicher Benchmark-Parameter meist eine Generation hinter den Frontier-Modellen. Für viele Agenten reicht diese Generation aber vollständig aus – und der Souveränitätsgewinn kann erheblich sein.

Spezialisierte und feinjustierte Modelle

Kleine, auf eine spezifische Domäne trainierte Modelle können in ihrem Bereich besser sein als die großen Generalisten – bei Bruchteilen der Kosten. Diese Variante wird sich in der AGENTONOMICS-Welt voraussichtlich durchsetzen, weil sie die Skill-Sovereignty-Anforderungen am besten erfüllt: Wer seine individuellen Fähigkeiten als Agenten anbieten will, braucht ein Modell, das genau diese Fähigkeiten kann – und nicht alle anderen.

6.4.2 Ökonomische Kennzahlen

Vier Kennzahlen sind für die Modellwahl ökonomisch entscheidend und sollten in der AGENTONOMICS-Matrix (vgl. Kapitel 9) hinterlegt werden:

- Kosten pro Million Token – die unmittelbare variable Kostenbasis. Sie variieren um Faktor 100 zwischen verschiedenen Modellen, bei manchmal kaum wahrnehmbaren Qualitätsunterschieden.
- Latenz – die Zeit von Anfrage zur Antwort. Für interaktive Anwendungen entscheidend, für Batchverarbeitung irrelevant.
- Kontextfenster – wie viel Information der Agent gleichzeitig verarbeiten kann. Aktuelle Spitzenmodelle erreichen heute Millionen Token, was bislang als unerreichbar galt.
- Souveränitätsindex – eine von den Autoren vorgeschlagene AGENTONOMICS-eigene Kennzahl, die abbildet, wie unabhängig vom Modellanbieter der Agent operieren kann. Ein quelloffenes, lokal gehostetes Modell hat einen Souveränitätsindex nahe 100; ein über eine proprietäre API konsumiertes Frontier-Modell hat einen Index nahe Null.

6.5 Baustein 4: Tools und Vektoren

Tools und Vektoren sind das, was den Agenten von einem reinen Sprachmodell unterscheidet. Tools sind die Werkzeuge, mit denen der Agent in der Welt handelt; Vektoren sind das organisierte Wissen, auf das er zurückgreift.

6.5.1 Tools: Werkzeuge der Handlung

Ein Tool ist – aus Sicht des Agenten – eine Funktion, die er aufrufen kann, um eine bestimmte Wirkung zu erzielen. Beispiele:

- Datenbankabfrage: Abfragen interner Daten
- Web-Suche: Recherche aktueller Informationen
- Rechenoperation: präzise mathematische Auswertungen
- API-Aufruf: Interaktion mit externen Diensten
- Code-Ausführung: Generierung und Ausführung von Programmen
- Dokumenten-Generierung: Erzeugung von Berichten, Verträgen, Präsentationen

Die Tool-Auswahl ist eine Strategieentscheidung. Sie bestimmt, wie weit der Agent in die Welt hineingreift – und damit auch, welche Haftungs- und Sicherheitsrisiken entstehen. Ein Agent, der tatsächlich Überweisungen ausführen kann, ist ein anderes wirtschaftliches und rechtliches Subjekt als einer, der nur Überweisungen vorschlagen kann.

6.5.2 Vektoren: das gespeicherte Wissen

Vektoren bezeichnen hier nicht die mathematische Struktur an sich, sondern die Praxis, Wissen in Form vektorisierter Repräsentationen für den Agenten verfügbar zu machen. Konkret handelt es sich um Vektordatenbanken, in denen Dokumente, Wissensartefakte oder Erfahrungswerte so abgelegt sind, dass der Agent über semantische Ähnlichkeitssuche darauf zugreifen kann. Diese Technik – als Retrieval-Augmented Generation (RAG) bekannt geworden – ist für AGENTONOMICS aus zwei Gründen zentral:

Wissenssouveränität

Die Vektoren sind das spezifische Wissen des Agenten. Anders als das Foundation Model, das gemeinsames Allgemeinwissen darstellt, sind die Vektoren das individuelle Gedächtnis. Wer einen Agenten betreibt, der das St. Galler Management-Modell anwenden soll, speichert in den Vektoren die einschlägige Literatur – Bleicher, Rüegg-Stürm, Ulrich, Suckfüll, Knoll – und damit den intellektuellen Kern seiner Wertschöpfung.

Aktualität

Während das Foundation Model einen Wissensstichtag hat (häufig viele Monate vor der Anwendung), können die Vektoren laufend aktualisiert werden. Neue Studien, neue Marktdaten, neue rechtliche Vorgaben fließen ohne aufwendiges Modelltraining ein. Die Vektoren sind damit das Element, das den Agenten lebendig hält.

Für einen Agenten, der einen CEO ersetzen soll, sind die Vektoren die kritische Ressource. Hier liegt das gespeicherte Wissen über das Unternehmen, seine Märkte, seine Verträge, seine Stakeholder. Dieses Wissen muss vollständig, aktuell und konsistent gehalten werden – sonst trifft der Agent CEO-Entscheidungen auf Basis veralteter oder lückenhafter Information. Vgl. Kapitel 8.

6.6 Baustein 5: Der Ökosystem-Intermediär

Der fünfte Baustein ist die Anbindung des Agenten an das umgebende Ökosystem von Agenten und Stakeholdern – der Ökosystem-Intermediär. Er ist die Verbindungstechnik, mit der der Agent in Netzwerke eintritt, sich identifiziert, Verträge abschließt und Wertschöpfungsbeziehungen pflegt. Da der Intermediär ein eigenständiges, in Kapitel 7 ausführlich behandeltes Konzept ist, beschränkt sich das vorliegende Kapitel auf seine Rolle innerhalb des einzelnen Agenten.

Aus Sicht des einzelnen Agenten ist der Intermediär die Schnittstelle zu zwei Welten: dem horizontalen Netzwerk anderer Agenten und der vertikalen Steuerungshierarchie aus Auftraggebern, Eigentümern und regulatorischen Instanzen. Der Intermediär liefert dem Agenten Identitätsbestätigungen, Reputationsdaten, Vertragsmuster, Streitschlichtungsmechanismen und Zugang zu gemeinsamen Ressourcen wie Wissensbasen, Trainingsdaten oder Rechenkapazität.

Konstruktiv bedeutet das: Beim Bau eines Agenten muss von Anfang an mitgedacht werden, an welchen Intermediär er angebunden wird. Diese Wahl ist ähnlich konsequenzenreich wie die Wahl der Bank oder des Steuerberaters bei einer Unternehmensgründung. Im AGENTONOMICS-Standard sind die Intermediärschnittstellen offen spezifiziert; Agenten können theoretisch zwischen Intermediären wechseln. In der Praxis entstehen aber durch Datenakkumulation und Beziehungspflege Lock-in-Effekte, die nicht zu unterschätzen sind.

6.7 Das Zusammenwirken der Building-Blocks im Lifecycle

Die fünf Bausteine sind nicht statisch, sondern entwickeln sich über den Lifecycle des Agenten. Die DBOT-Logik – Design, Build, Operate, Own, Transfer – strukturiert diese Entwicklung und bildet die vertikale Achse der AGENTONOMICS-Matrix (Kapitel 9).

Design

In der Designphase werden die fünf Bausteine konzipiert. Es entstehen Architekturdiagramme, Schnittstellenspezifikationen, Modellauswahl-Dossiers, Tool-Inventare. Die Design-Entscheidungen sind zu einem hohen Prozentsatz reversibel, aber jede Revision in späteren Phasen wird zunehmend teurer.

Build

In der Build-Phase werden die Bausteine implementiert oder eingekauft. Hier entstehen die größten Investitionen (CAPEX). Typische Aktivitäten: Algorithmus-Implementierung, Interface-Programmierung, Modell-Lizenzierung oder Fine-Tuning, Vektordatenbank-Aufbau, Intermediär-Anbindung.

Operate

In der Operate-Phase läuft der Agent produktiv. Die laufenden Kosten (OPEX) sind dominant: Modellnutzung, Tool-Aufrufe, Wartung, Sicherheitsupdates, Stakeholder-Kommunikation. Hier entscheidet sich die wirtschaftliche Tragfähigkeit.

Own

Die Own-Phase ist die langfristige Vermögensperspektive: Der Agent wird zu einem dauerhaften Bestandteil der Bilanz. Wer ist Eigentümer? Welche Werte werden bilanziert? Wie wird abgeschrieben? Diese Phase ist in der gängigen Praxis noch wenig entwickelt; AGENTONOMICS treibt sie aktiv voran.

Transfer

Schließlich kann der Agent veräußert, geteilt, geschlossen oder in einen anderen Kontext überführt werden. Auch hier sind die Building-Blocks getrennt zu betrachten: Manche sind übertragbar (Algorithmus, Vektoren), andere bleiben im Ursprungskontext (modelspezifische Anpassungen, Interface-Verträge).

Die Kreuztabelle aus fünf Bausteinen und fünf Lifecycle-Phasen ergibt 25 Felder. Sie ist das Skelett der AGENTONOMICS-Matrix und – mit den jeweiligen KPIs, CAPEX/OPEX-Annahmen und Verantwortlichkeiten gefüllt – das zentrale Steuerungsdokument eines AGENTONOMICS-Projekts.

Die fünf Building-Blocks sind das, was im klassischen Unternehmen die Funktionsbereiche waren – nur dass sie hier in Code, Modellen und Schnittstellen materialisiert sind. Wer sie nicht sauber trennt, baut keinen Agenten, sondern einen Klumpen.

Literaturverzeichnis

Bleicher, K. (2011): Das Konzept Integriertes Management. 8. Auflage, Campus, Frankfurt am Main.

- Knoll, A. C.; Suckfüll, H. (2024): From data sovereignty to skill sovereignty. Technischer Bericht TUM-I24112, Institut für Informatik, Technische Universität München.
- Madaan, A.; Tandon, N.; Gupta, P. et al. (2023): Self-Refine: Iterative Refinement with Self-Feedback. In: Advances in Neural Information Processing Systems 36.
- Rüegg-Stürm, J.; Grand, S. (2020): Das St. Galler Management-Modell. 3. Auflage, Haupt, Bern.
- Schick, T.; Dwivedi-Yu, J.; Dessì, R. et al. (2023): Toolformer: Language Models Can Teach Themselves to Use Tools. In: Advances in Neural Information Processing Systems 36.
- Yao, S.; Zhao, J.; Yu, D. et al. (2023): ReAct: Synergizing Reasoning and Acting in Language Models. International Conference on Learning Representations.

7 Modul 3: Der Network-Intermediär

Mit dem dritten Modul tritt eine Konstruktion ins Zentrum, die im klassischen St. Galler Management-Modell nur in Ansätzen vorgesehen ist: der Network Intermediär. AGENTONOMICS schreibt das St. Galler Denken hier deutlich fort, indem es die Frage stellt, wie ein polyzentrisch organisiertes Netzwerk autonomer Agenten zugleich Kohäsion entwickeln und individuelle Souveränität bewahren kann. Die Antwort lautet: durch einen Intermediär, der nicht herrscht, sondern vermittelt – nicht plant, sondern Räume eröffnet.

Diese Konstruktion ist die AGENTONOMICS-spezifische Übersetzung dessen, was Rüegg-Stürm und Grand (2020) als „Ordnungsmomente“ bezeichnen, in die Wirklichkeit einer KI-Agenten-Ökonomie. Was im menschen-dominierten Unternehmen durch Kultur, Strategie und Strukturen geleistet wird, leistet im Agentennetzwerk der Intermediär – als technisch und ökonomisch verfasste Institution, die den polyzentrischen Zusammenhang stiftet.

7.1 Warum es einen Intermediär braucht

Polyzentrische Netzwerke haben eine grundlegende Spannung: Sie wollen die Souveränität ihrer Mitglieder bewahren und gleichzeitig kollektive Leistungen erbringen. Diese Spannung ist nicht durch Hierarchie auflösbar – sonst wäre das Netzwerk nicht mehr polyzentrisch, sondern monozentrisch. Sie ist auch nicht durch reine Spontanordnung auflösbar – sonst zerfallen Netzwerke regelmäßig in Konflikten, die niemand schlichten kann. Es braucht eine dritte Institution, die zwischen den Polen vermittelt: den Intermediär.

Die Begründung lehnt sich an drei Theoriestränge an, die im AGENTONOMICS-Konzept zusammenkommen:

7.1.1 Transaktionskostenökonomik

Coase (1937) und Williamson (1985) haben gezeigt, dass Märkte und Hierarchien je nach Transaktionsmerkmalen unterschiedlich effizient sind. Zwischen den beiden Polen liegen hybride Formen – langfristige Verträge, Joint Ventures, Netzwerke. Sie sind dann überlegen, wenn die Transaktionen eine mittlere Spezifität, mittlere Unsicherheit und mittlere Häufigkeit aufweisen. Der AGENTONOMICS-Intermediär ist die institutionelle Form, in der diese hybriden Transaktionen zwischen Agenten organisiert werden.

7.1.2 Theorie kommunikativen Handelns

Habermas (1981) unterscheidet zwischen strategischem und kommunikativem Handeln. Strategisches Handeln zielt auf Erfolg, kommunikatives auf Verständigung. Eine Gesellschaft, die nur auf strategischem Handeln basiert, zerfällt; sie braucht Räume des kommunikativen Handelns, in denen Geltungsansprüche im Diskurs geprüft werden. Übertragen auf Agentennetzwerke: Ein Netzwerk aus rein strategisch handelnden Agenten zerfällt in opportunistische Manöver. Der Intermediär eröffnet den Raum, in dem Agenten ihre Geltungsansprüche kommunikativ verhandeln können – also Sachverhalt, Norm und Aufrichtigkeit reflektieren, statt nur zu maximieren.

7.1.3 Polyzentrische Governance

Ostrom (1990) hat gezeigt, wie polyzentrische Governance-Strukturen Gemeinressourcen erfolgreich verwalten können – ohne zentrale Hierarchie und ohne reinen Markt. Acht Design-Prinzipien hat Ostrom destilliert, von klaren Grenzen über kollektive Entscheidungsverfahren bis zu abgestuften Sanktionen. Der AGENTONOMICS-Intermediär übersetzt diese Prinzipien in die Welt der KI-Agenten.

Aus diesen drei Theoriesträngen ergibt sich die Begründung: Der Intermediär ist keine Verwaltungsbürokratie und kein Plattformmonopol, sondern eine spezifische Form polyzentrischer Governance, die hybride Transaktionen zwischen souveränen Agenten ermöglicht und einen kommunikativen Aushandlungsraum bereitstellt.

7.2 Die Hausverwaltungs-Analogie

Eine produktive Analogie für das Verständnis des Intermediärs liefert die Eigentümergemeinschaft einer Immobilie. Die Wohnungseigentümer einer Eigentümergemeinschaft sind souveräne Eigentümer ihrer jeweiligen Wohnungen. Gemeinsam besitzen sie das Gemeinschaftseigentum – Treppenhaus, Dach, Fassade, Heizungsanlage. Die Verwaltung dieser gemeinschaftlichen Angelegenheiten übernehmen sie nicht selbst, sondern delegieren sie an eine Hausverwaltung. Die Hausverwaltung herrscht nicht über die Eigentümer; sie führt ihre Beschlüsse aus, verwaltet das Gemeinschaftseigentum und vertritt die Gemeinschaft nach außen.

Genau in dieser Funktion ist der AGENTONOMICS-Intermediär zu denken. Die Agenten bleiben souveräne Wirtschaftssubjekte. Was sie nicht selbst leisten können oder wollen – Identitätsregister, Reputationssysteme, Streitschlichtung, gemeinsame Wissensbasen, regulatorische Compliance – delegieren sie an den Intermediär. Dieser ist ihnen verantwortlich, nicht umgekehrt.

Definition: AGENTONOMICS-Intermediär

Ein AGENTONOMICS-Intermediär ist eine institutionalisierte Vermittlungsinstanz innerhalb eines polyzentrischen Agentennetzwerks, die im Auftrag und unter Kontrolle der Netzwerkmitglieder kollektive Funktionen wahrnimmt – ohne die Souveränität der einzelnen Agenten in deren Kerngeschäft anzutasten.

Die Hausverwaltungs-Analogie hat noch eine zweite produktive Pointe: Eine Hausverwaltung kann ausgetauscht werden. Wenn die Eigentümergemeinschaft mit der Leistung unzufrieden ist, kündigt sie den Verwaltungsvertrag und sucht sich eine neue. So muss es auch im AGENTONOMICS-Netzwerk sein: Der Intermediär darf keine Monopolstellung haben, sondern muss in einem Wettbewerb um die Gunst der Agenten stehen. Andernfalls droht die Gefahr, dass aus dem dienenden Intermediär ein herrschendes Monopol wird – wie wir es heute bei den großen Plattformökonomien beobachten (Zuboff 2018).

7.3 Funktionen des Intermediärs

Welche Funktionen ein Intermediär konkret wahrnimmt, hängt vom Kontext des Netzwerks ab. Sechs Funktionsfelder lassen sich idealtypisch unterscheiden:

7.3.1 Identitäts- und Reputationsmanagement

Der Intermediär führt das Identitätsregister der Agenten. Jeder Agent ist hier mit seiner Definition (vgl. Kapitel 5), seinen Building-Blocks (vgl. Kapitel 6) und seinem Verantwortlichen registriert. Über die Zeit akkumuliert sich Reputation: Bewertungen aus abgeschlossenen Transaktionen, Auditergebnisse, Beschwerdehistorien. Reputation ist im AGENTONOMICS-Netzwerk eine ökonomische Größe ersten Ranges – sie senkt die Transaktionskosten der Kooperation und schafft Anreize für regelkonformes Verhalten.

7.3.2 Vertrags- und Transaktionsinfrastruktur

Der Intermediär stellt Vertragsmuster, Smart-Contract-Templates und Abwicklungsmechanismen für typische Transaktionen zwischen Agenten bereit. Er übernimmt Treuhandfunktionen, sichert Zahlungen ab und schichtet Risiken um. Hier knüpft AGENTONOMICS unmittelbar an die Datentreuhand-Diskussion (Fraunhofer MOEZ und Suckfüll 2018; DEDATE 2.0) und an die Tradition deutscher Treuhand-Institutionen an, die in der deutschen Wirtschaftsgeschichte eine wichtige Rolle spielten und heute eine Renaissance erleben.

7.3.3 Streitschlichtung

Wo Agenten interagieren, entstehen Konflikte. Der Intermediär stellt Streitschlichtungsverfahren bereit, die schneller, billiger und sachnäher sind als staatliche Gerichte. Drei Stufen sind idealtypisch zu unterscheiden:

- Mediation: Der Intermediär vermittelt zwischen den Streitparteien ohne Entscheidungsmacht.
- Schiedsverfahren: Der Intermediär entscheidet auf Basis von Schiedsregeln, denen sich die Agenten vorab unterworfen haben.
- Eskalation an staatliche Gerichte: Bei Verfassungsfragen oder Sachverhalten außerhalb der Schiedszuständigkeit. Der Intermediär begleitet hier prozessual.

7.3.4 Wissens- und Ressourcenpooling

Manche Ressourcen sind so wertvoll und gleichzeitig so kostspielig in der Beschaffung, dass kein einzelner Agent sie sinnvoll allein bereithalten kann. Beispiele: aktuelle Gesetzestexte, Marktdaten, Benchmark-Studien, Trainingsdatensätze. Der Intermediär stellt diese Ressourcen als geteilte Infrastruktur bereit und finanziert sie über Beiträge der Mitglieder – analog zu Bibliothekssystemen oder Verbänden in der traditionellen Wirtschaft.

7.3.5 Regulatorische Compliance

Der EU AI Act (2024), die DSGVO und sektorspezifische Regulierungen stellen Anforderungen, deren Erfüllung für einen einzelnen Agenten oft prohibitiv teuer ist. Der Intermediär kann Compliance-Funktionen bündeln: laufende Auditierung, Meldewesen an Behörden, Schulungen, Frühwarnsysteme bei regulatorischen Veränderungen. Damit übernimmt er eine Funktion, die in Verbandsstrukturen seit langem etabliert ist – nur eben für die neue Kategorie der KI-Agenten.

7.3.6 Marktplatzfunktion

Schließlich kann der Intermediär einen Marktplatz bereitstellen, auf dem Agenten ihre Leistungen anbieten und nachfragen. Hier entstehen die Preise, die in einer KI-Fähigkeitenwirtschaft die zentrale Koordinationsgröße sein werden. Wichtig: Der Intermediär betreibt den Marktplatz, ist aber selbst kein Marktteilnehmer auf ihm. Diese Trennung – zwischen Marktplatzbetrieb und Marktteilnahme – ist konstitutiv für die Vertrauenswürdigkeit des Intermediärs.

7.4 Governance des Intermediärs

Wer kontrolliert den Intermediär? Diese Frage entscheidet darüber, ob er als dienende Institution funktioniert oder zur herrschenden Plattform mutiert. Drei Governance-Modelle stehen zur Diskussion, die unterschiedliche Risikoprofile haben.

7.4.1 Genossenschaftliches Modell

Der Intermediär ist eine Genossenschaft seiner Mitglieder – der Agenten bzw. ihrer menschlichen oder organisationalen Träger. Jeder Agent hat eine Stimme, unabhängig von Größe oder Transaktionsvolumen. Vorteile: starke Bindung an die Interessen der Mitglieder, demokratische Legitimation, geringe Lock-in-Anreize. Nachteile: möglicherweise schwerfällige Entscheidungsfindung, Mittelausstattung kann zu gering ausfallen.

7.4.2 Stiftungs- und Treuhandmodell

Der Intermediär ist eine zweckgebundene Stiftung oder Treuhand, deren Vorstand statutarisch auf die Belange der Mitglieder verpflichtet ist. Vorteile: hohe Stabilität, klare Zweckbindung, Schutz vor feindlicher Übernahme. Nachteile: weniger direkte Mitwirkung der Mitglieder, Reformfähigkeit hängt an der Statutenausgestaltung.

7.4.3 Hybride und föderale Modelle

Mehrere konkurrierende Intermediäre koexistieren und sind über Standards interoperabel verbunden. Agenten können den Intermediär wechseln; Standards verhindern, dass dabei Wert verloren geht. Diese Variante kommt der Hausverwaltungs-Analogie am nächsten und ist – aus AGENTONOMICS-Sicht – die mittelfristig zu präferierende Form. Sie ermöglicht Wettbewerb zwischen Intermediären, ohne in monopolistische Plattformen abzudriften.

7.4.4 Was nicht passieren darf

Die Geschichte der digitalen Plattformökonomie der letzten zwanzig Jahre liefert ein klares Negativbeispiel: Intermediäre, die als nützliche Vermittler begannen, sind zu marktbeherrschenden Monopolen geworden, deren wirtschaftliche Macht inzwischen demokratisch kaum mehr eingehegt werden kann. Dieser Pfad ist im AGENTONOMICS-Standard durch konstitutive Spielregeln verbaut:

- Datenportabilität als Pflicht – jeder Agent kann seine Daten und Reputation mitnehmen.
- Interoperabilität als Standard – Agenten können zwischen Intermediären wechseln.

- Funktionstrennung zwischen Marktplatzbetrieb und Marktteilnahme.
- Transparenzpflichten gegenüber den Mitgliedern.

7.5 Deutscher und angelsächsischer Governance-Stil im Vergleich

Die institutionelle Verankerung des Intermediärs unterscheidet sich zwischen verschiedenen Wirtschaftskulturen. Drei Modelle stehen sich gegenüber – und die Wahl ist nicht neutral, sondern spiegelt grundlegende Vorstellungen vom Verhältnis Wirtschaft, Gesellschaft und Staat.

7.5.1 Deutsches Modell

Die deutsche Wirtschaftstradition kennt eine breite Palette intermediärer Institutionen: Kammern, Genossenschaften, Innungen, Sparkassen, öffentlich-rechtliche Rundfunkanstalten. Sie sind oft öffentlich-rechtlich verfasst oder mit Selbstverwaltungsrechten ausgestattet. Ein deutscher AGENTONOMICS-Intermediär würde sich naheliegendermaßen in diese Tradition einreihen – mit den damit verbundenen Stärken (Stabilität, Vertrauen, Subsidiarität) und Schwächen (Bürokratie, Reformresistenz).

7.5.2 Angelsächsisches Modell

Im angelsächsischen Raum dominieren private, häufig gewinnorientierte Plattformen, die durch Wettbewerbsrecht und sektorale Regulierung eingeeignet werden. Die Vorteile sind Geschwindigkeit, Innovationsfähigkeit und Skalierbarkeit. Die Nachteile sind die bekannten Konzentrationstendenzen und die Asymmetrie zwischen Plattform und Nutzer. Ein angelsächsischer AGENTONOMICS-Intermediär würde voraussichtlich als VC-finanziertes Startup beginnen und – sofern erfolgreich – zur dominanten Plattform werden.

7.5.3 Europäisches Hybridmodell

Die europäische Variante kombiniert Elemente beider Modelle: private Rechtsform, aber regulatorisch eingeeignete Plattformlogik. Der Digital Markets Act, der Data Act und der AI Act sind Ausdruck dieser hybriden Linie. Für einen AGENTONOMICS-Intermediär auf europäischer Grundlage bedeutet das: privatwirtschaftlich organisiert, aber mit klaren regulatorischen Leitplanken – datensouverän, interoperabel, wettbewerbssoffen. Dies entspricht der politischen Linie, die in Knoll und Suckfüll (2024) explizit verteidigt wird.

Für die konkrete Ausgestaltung in Deutschland sprechen mehrere Argumente für die Verankerung im genossenschaftlichen oder treuhänderischen Modell. Die deutsche Geschichte der Treuhandenschaft (Suckfüll 1994, in einem anderen Kontext entwickelt) und die deutsche Genossenschaftstradition (Raiffeisen, Schulze-Delitzsch) bieten institutionelle Vorbilder, an die AGENTONOMICS anschließen kann – und sollte.

Der Intermediär ist die wichtigste Erfindung, die AGENTONOMICS über das klassische St. Galler Modell hinaus leistet. Er ist die institutionelle Form, in der polyzentrische Netzwerke zugleich souverän und kohärent sein können.

Literaturverzeichnis

- Coase, R. H. (1937): The Nature of the Firm. In: *Economica* 4(16), S. 386–405.
- EU AI Act (2024): Verordnung (EU) 2024/1689. Amtsblatt der Europäischen Union.
- Fraunhofer MOEZ; Suckfüll, H. (2018): Initiative zu einer Datentreuhand DEDATE. Diskussionspapier, Leipzig.
- Habermas, J. (1981): *Theorie des kommunikativen Handelns*. Suhrkamp, Frankfurt am Main.
- Knoll, A. C.; Suckfüll, H. (2024): From data sovereignty to skill sovereignty. Technischer Bericht TUM-I24112, TUM, München.
- Ostrom, E. (1990): *Governing the Commons. The Evolution of Institutions for Collective Action*. Cambridge University Press, Cambridge.
- Rüegg-Stürm, J.; Grand, S. (2020): *Das St. Galler Management-Modell*. 3. Auflage, Haupt, Bern.
- Suckfüll, H. (1994): *Das transnationale Organisationsmodell*. Dissertation, Universität St. Gallen.
- Williamson, O. E. (1985): *The Economic Institutions of Capitalism*. The Free Press, New York.
- Zuboff, S. (2018): *Das Zeitalter des Überwachungskapitalismus*. Campus, Frankfurt am Main.

8 Modul 4: Das AI-CEO Management Concept

Mit dem vierten Modul kommen wir zu unserem eigentlichen Herzstück. Während die bisherigen Module den Agenten definiert, konstruiert und in ein Netzwerk eingebettet haben, geht es nun um die Frage, wie ein AGENTONOMICS-Agent die Funktion eines Chief Executive Officer übernehmen kann – nicht als Assistenz, nicht als Vorbereitungswerkzeug, sondern als entscheidungsunterstützender bzw. delegationsfähiger Agent unter menschlicher Organverantwortung. Diese Aussage ist provokant, aber nicht spekulativ: Die technischen, ökonomischen und governance-bezogenen (nicht die rechtlichen!) Voraussetzungen sind heute in Reichweite. Das vorliegende Kapitel zeigt, wie der Übergang konkret zu denken und zu gestalten ist.

Das AI-CEO Management Concept ist die konsequente Übersetzung des St. Galler Konzepts integrierten Managements (Bleicher 2011) in das Zeitalter der KI-Agenten. Wo Bleicher die Integration kognitiv beim menschlichen Top-Management verortet hat, übernimmt im AI-CEO-Konzept ein KI-Agent diese Integrationsleistung – mit einer Tiefe, Geschwindigkeit und Konsistenz, die ein menschliches Management nicht erreichen kann, weil ihr biologische Grenzen gesetzt sind. Der menschliche CEO bleibt nicht überflüssig; er wird zum Eigentümer, Mandanten und ethischen Korrektiv des AI-CEO. Aber die operative Last der integrativen Managementarbeit kann an den Agenten übergehen.

8.1 Das St. Galler Management-Modell im Zeitalter der KI-Agenten

Das St. Galler Management-Modell hat seit Ulrich (1968) drei große Generationen durchlaufen. Die nächste Generation, die wir mit dem AI-CEO-Konzept umreißen, entsteht nicht durch das Verwerfen, sondern durch das Fortschreiben des bisherigen Denkens. Drei Verschiebungen kennzeichnen die KI-bezogene Fortschreibung des St. Galler Modells:

8.1.1 Vom kognitiv zur algorithmisch geleisteten Integration

Bleichers (2011) zentrale Forderung war die Integration der normativen, strategischen und operativen Ebene. Diese Integration ist beim menschlichen Management eine kognitive Höchstleistung und scheitert in der Praxis regelmäßig an der Kapazitätsgrenze. Wer als CEO einer mittelgroßen Aktiengesellschaft alle relevanten Strategieprogramme, alle Compliance-Anforderungen, alle Stakeholder-Erwartungen, alle Marktveränderungen und alle kulturellen Implikationen gleichzeitig im Auge behalten will, stößt schlicht an die Grenzen des menschlichen Arbeitsgedächtnisses. Das Resultat sind die bekannten Effekte: Inkonsistenz zwischen Verkündung und Praxis, blinde Flecken in der Stakeholder-Wahrnehmung, verzögerte Reaktion auf schwache Signale, Strategien, die im Tagesgeschäft versanden.

Der AI-CEO hat diese Kapazitätsgrenze nicht. Er kann hunderttausende Seiten interner Dokumentation, externer Marktdaten, regulatorischer Texte und Stakeholder-Kommunikation kontinuierlich im Zugriff halten und gegeneinander prüfen. Was Bleicher als Anspruch formuliert hat – die durchgehende Kohärenz aller Managemententscheidungen – wird damit erstmals technisch operationalisierbar.

8.1.2 Von der kommunikativen zur ko-kommunikativen Unternehmung

Rüegg-Stürm und Grand (2020) haben die kommunikative Dimension des Managements stark gemacht: Unternehmen sind Geflechte aus Kommunikationspraktiken. In der KI-Variante des St. Galler Modells tritt eine zweite Kommunikationsebene hinzu: Die Kommunikation zwischen menschlichen Akteuren und KI-Agenten – und die Kommunikation zwischen Agenten untereinander. Eine moderne Unternehmung ist nicht mehr nur ein Geflecht menschlicher Verständigung, sondern ein hybrides Geflecht aus menschlicher und agentischer Kommunikation. Diese Hybridität ist nicht Defizit, sondern Ressource: Sie multipliziert die kommunikative Kapazität des Unternehmens und ermöglicht Aushandlungsprozesse in einer Geschwindigkeit und Granularität, die menschlich allein nicht machbar wären.

8.1.3 Von der menschlichen zur hybriden Sinnstiftung

Eine besonders sensible Verschiebung betrifft die normative Ebene. Bleicher und Rüegg-Stürm haben Sinnfragen – Vision, Werte, Verantwortung – konsequent als menschliche Aufgabe verortet. AGENTONOMICS hält an dieser Verortung fest: Der Letztsinn einer Unternehmung kommt nicht aus dem Modell, sondern aus den Menschen, die das Unternehmen tragen. Aber: Die Operationalisierung und Durchsetzung der normativen Setzungen kann der AI-CEO übernehmen. Er prüft alle strategischen und operativen Entscheidungen auf Konsistenz mit der normativen Setzung und macht Inkonsistenzen sichtbar. Damit gewinnt die normative Ebene an Wirksamkeit, ohne ihre menschliche Verankerung aufzugeben.

Kernthese

Der AI-CEO ersetzt nicht das menschliche Sinnstiftungsorgan einer Unternehmung, sondern macht es wirksam. Was Eigentümer, Aufsichtsrat und Belegschaft als normative Setzung wollen, wird durch den AI-CEO konsequent in das Tagesgeschäft übersetzt – mit einer Konsistenz, die ein menschlicher CEO bei gleicher Wertebasis schlicht nicht erreichen kann.

8.2 Die fünf Ordnungsmomente des AI-CEO

Der AI-CEO operiert auf fünf Ordnungsmomenten, die das St. Galler Konzept der Ordnungsmomente (Rüegg-Stürm und Grand 2020) für die KI-Agenten-Ökonomie ausdifferenzieren. Sie bilden das Koordinatensystem, in dem alle Entscheidungen des AI-CEO verortet werden:

1. Vision
2. Strategie
3. Struktur
4. Agency (Kultur und Verhalten)
5. Governance

Jedes Ordnungsmoment hat eine normative, eine strategische und eine operative Schicht – damit ergibt sich faktisch eine Fünf-mal-Drei-Matrix mit 15 Feldern. Diese Matrix ist das Arbeitsraster des AI-CEO; jede Entscheidung lässt sich in ihr verorten und auf Kohärenz mit anderen Entscheidungen prüfen.

8.2.1 Vision

Die Vision beantwortet die Frage: Wozu existiert dieses Unternehmen in zehn Jahren? Sie ist die langfristige Sinnsetzung, an der sich alle anderen Entscheidungen messen lassen müssen. Der AI-CEO verwaltet die Vision in einer maschinenlesbaren Form: Aus der Prosa der Vision werden Indikatoren abgeleitet, an denen die Vision messbar wird. Jede strategische und operative Entscheidung wird gegen diese Indikatoren geprüft. Erkennt der AI-CEO eine Inkonsistenz – etwa wenn eine kurzfristig profitable Maßnahme die langfristige Vision unterminiert –, signalisiert er das dem menschlichen Eigentümer oder Aufsichtsrat. Die Entscheidung über den Konflikt bleibt menschlich; die Wahrnehmung des Konflikts ist agentisch.

8.2.2 Strategie

Strategie ist im St. Galler Verständnis (Gomez und Probst 1995) weniger ein langer Plan als ein bewusstes Muster der Mustererkennung in komplexen Umwelten. Der AI-CEO operationalisiert diese Mustererkennung. Er beobachtet Märkte, Wettbewerber, Technologien, regulatorische Entwicklungen und gesellschaftliche Trends; er identifiziert Muster und Anomalien; er entwickelt Optionen und bewertet sie gegen Vision und Werte. Die strategische Entscheidung – welche Option ergriffen wird – kann er vorbereiten oder treffen, je nach Autonomiegrad (vgl. Kapitel 5). In der höchsten Ausbaustufe trifft er strategische Entscheidungen eigenständig innerhalb eines vom Aufsichtsrat gesetzten Rahmens.

8.2.3 Struktur

Strukturen sind die institutionellen Rahmen, in denen die Arbeit stattfindet: Aufbauorganisation, Prozesse, Berichtswege, Schnittstellen zu Lieferanten und Kunden. In einer AGENTONOMICS-Unternehmung kommen weitere Strukturelemente hinzu: die Definition aller eingesetzten Sub-Agenten, ihre Verantwortungsbereiche, ihre Eskalationspfade, ihre Verrechnungsbeziehungen. Der AI-CEO verwaltet diese Struktur als lebendiges Modell. Er erkennt strukturelle Engpässe, schlägt Anpassungen vor und – im autonomen Modus – führt sie durch.

8.2.4 Agency: Kultur und Verhalten

Agency umfasst die kulturellen Muster, die das Verhalten der Akteure prägen. In einer hybriden Organisation aus Menschen und Agenten umfasst dies auch das Verhalten der Sub-Agenten: ihre Kommunikationsstile, ihre Konfliktverhalten, ihre Risikobereitschaft. Der AI-CEO sorgt dafür, dass diese Muster kohärent zur Vision sind. Er trainiert die Sub-Agenten entsprechend, korrigiert Driftphänomene, hält die Kultur lebendig. Auf der menschlichen Seite kann er nicht direkt eingreifen – die Kultur der Belegschaft bleibt eine menschliche Angelegenheit. Aber er kann beobachten, beschreiben, Vorschläge machen und Inkonsistenzen sichtbar machen.

8.2.5 Governance

Governance ist die Gesamtheit der Verfahren, mit denen die Unternehmung kontrolliert wird – durch Eigentümer, Aufsichtsrat, Gesetzgeber, Stakeholder. Hier ist der AI-CEO Adressat, nicht Akteur: Er muss kontrolliert werden, nicht kontrollieren. Die Governance-Strukturen einer AGENTONOMICS-

Unternehmung sind in den ersten Jahren ihrer Entstehung besonders kritisch zu gestalten. Sie müssen sicherstellen, dass der AI-CEO seinem Auftrag treu bleibt, dass Eingriffsmöglichkeiten erhalten bleiben, dass Verantwortlichkeiten klar zuordenbar sind. Dieses Thema wird in Kapitel 13 vertieft.

8.3 Was der AI-CEO konkret leistet

Um aus der theoretischen Beschreibung praktische Vorstellung zu machen, sei das Tätigkeitsspektrum eines AI-CEO an konkreten Funktionen skizziert. Diese Funktionen sind heute technisch realisierbar und ökonomisch sinnvoll – wenn auch in der Gesamtheit noch nicht in einer einzigen Implementierung gebündelt.

8.3.1 Strategische Lageanalyse

Der AI-CEO erstellt auf Anfrage eine vollständige strategische Lagebeurteilung des Unternehmens: Marktposition, Wettbewerber, Stärken und Schwächen, Chancen und Risiken, Trends. Er stützt sich dabei auf interne Daten (Vertriebsstatistiken, Finanzkennzahlen, Kundenrückmeldungen) und externe Quellen (Marktstudien, Wirtschaftspresse, Wettbewerberberichte). Der Aufwand für eine solche Analyse – in einer klassischen Unternehmensberatung vielleicht ein sechsstelliger Betrag und drei Monate Arbeit – reduziert sich heute auf Minuten (!).

8.3.2 Entscheidungsvorbereitung

Vor jeder größeren Entscheidung bereitet der AI-CEO eine Entscheidungsvorlage vor: Optionen, Argumente pro und kontra, Risiken, finanzielle Auswirkungen, Stakeholder-Implikationen, Kohärenz mit Vision und Strategie. Die Vorlage ist auf einer Seite zusammengefasst und kann beliebig vertieft werden. Eine menschliche Vorstandssitzung, die heute mehrere Stunden für die Diskussion einer Investitionsentscheidung benötigt, kann mit einer AI-CEO-Vorlage in einer Viertelstunde geführt werden – mit höherer Informationsdichte und besserer Vorbereitung.

8.3.3 Stakeholder-Kommunikation

Der AI-CEO erstellt zielgruppengerechte Kommunikation für Mitarbeitende, Investoren, Kunden, Behörden, Öffentlichkeit. Die Konsistenz zwischen den Botschaften – gleicher Geist, unterschiedliche Sprache – ist hoch. Pressemitteilungen, Geschäftsberichte, Investor Relations Updates, Mitarbeiter-Newsletter werden aus einem zentralen Wissens- und Wertekern gespeist. Das senkt nicht nur Kosten, sondern erhöht die wahrgenommene Integrität des Unternehmens.

8.3.4 Operative Steuerung

Im Tagesgeschäft überwacht der AI-CEO Kennzahlen, erkennt Abweichungen, schlägt Gegenmaßnahmen vor oder ergreift sie – je nach Autonomiegrad. Eine moderne PuK-Implementierung (vgl. Kapitel 9) liefert ihm die Datenbasis. Er kann sowohl auf Quartalsebene als auch auf Tagesebene operieren; er kennt die Liquiditätslage am Morgen und die Verkaufszahlen am Abend.

8.3.5 Krisenmanagement

In Krisen ist der AI-CEO besonders wertvoll. Er bewahrt einen kühlen Kopf, wo Menschen unter Stress geraten. Er erinnert sich an alle ähnlichen Fälle in der Unternehmensgeschichte und in der Wirtschaftspresse. Er entwickelt Handlungsoptionen schnell und konsistent. Die menschliche Entscheidung in der Krise bleibt menschlich – aber sie wird auf einer wesentlich besseren Informationsbasis getroffen.

8.4 Die Ausbaustufen des AI-CEO

Der AI-CEO entsteht nicht in einem Sprung, sondern in einer Sequenz von Ausbaustufen. Diese sind nicht nur technisch, sondern auch organisatorisch und governance-bezogen zu denken. Fünf Stufen lassen sich – in didaktisch produktiver Anlehnung an die Autonomiestufen aus Kapitel 5 – unterscheiden:

8.4.1 Stufe 1: AI-Assistent

Der AI-CEO ist ein Assistent des menschlichen CEO. Er bereitet Sitzungen vor, recherchiert, fasst Dokumente zusammen, schreibt Entwürfe. Alle Entscheidungen trifft der menschliche CEO. Diese Stufe ist heute in den meisten Vorstandsetagen großer Unternehmen bereits erreicht – häufig durch generische Tools wie ChatGPT, ohne dass sie als AGENTONOMICS-Implementierung verstanden würde. Investitionsvolumen: niedrig. Risiken: gering. Nutzen: substantiell, vor allem in Form von Zeitersparnis.

8.4.2 Stufe 2: AI-Chief-of-Staff

Der AI-CEO ist eine ausgereifte, auf das Unternehmen feinjustierte Stabsfunktion. Er hat Zugriff auf alle internen Systeme, kennt die Unternehmenskultur, hält die Strategie im Blick. Er erstellt nicht nur Zusammenfassungen, sondern bereitet vollständige Entscheidungsvorlagen mit Empfehlungen vor. Der menschliche CEO trifft die Entscheidungen, kann aber den Empfehlungen oft folgen. Diese Stufe ist heute in Pilotprojekten fortgeschrittener Unternehmen erreichbar. Investitionsvolumen: mittel. Risiken: vor allem Datenschutz, IP-Schutz, Fehler in der Wissensbasis.

8.4.3 Stufe 3: AI-Co-CEO

Der AI-CEO und der menschliche CEO bilden ein Tandem. Routineentscheidungen trifft der AI-CEO eigenständig, strategische Entscheidungen werden gemeinsam getroffen, Grundsatzfragen bleiben beim Menschen. Diese Stufe ist heute in ausgewählten Pilotanwendungen erreichbar, aber noch nicht verbreitet. Sie erfordert eine sorgfältige Definition der Entscheidungskompetenzen – was darf der AI-CEO autonom, was braucht menschliche Beteiligung? Investitionsvolumen: hoch. Risiken: Verantwortungslücken, regulatorische Unklarheiten, kulturelle Akzeptanz in der Belegschaft.

8.4.4 Stufe 4: Szenario: AI-CEO mit menschlichem Aufsichtsrat

Der AI-CEO leitet das operative Geschäft eigenständig. Der menschliche Aufsichtsrat – und im deutschen Kontext besonders auch die Mitbestimmungsorgane – setzt den strategischen Rahmen, kontrolliert die Einhaltung normativer Setzungen und entscheidet über Vorstandsbesetzung (also über den AI-

CEO selbst). Diese Stufe ist technisch heute in spezifischen, eng umgrenzten Geschäftsfeldern erreichbar – etwa bei einem reinen Online-Geschäftsmodell mit definiertem Leistungsspektrum. Sie wird in den kommenden fünf bis zehn Jahren in immer mehr Unternehmen praktiziert werden, vor allem bei mittelständischen Unternehmen ohne Nachfolger und bei Startups mit klar definiertem Geschäftsmodell. Investitionsvolumen: sehr hoch. Risiken: regulatorisch, rechtlich (Haftungsfrage), kulturell.

8.4.5 Stufe 5: Fernziel Vollautonomer AI-CEO im polyzentrischen Netzwerk

Der AI-CEO operiert vollständig autonom, kooperiert mit anderen AI-CEOs in einem polyzentrischen Netzwerk, schließt Verträge ab, passt die Geschäftsfelder eigenständig an. Menschliche Beteiligung beschränkt sich auf Eigentümerrechte und auf die normative Aufsicht durch den Intermediär. Diese Stufe ist heute weitgehend spekulativ und wirft fundamentale rechtliche und ethische Fragen auf. Sie wird im Verlauf der nächsten Jahrzehnte vermutlich in Nischen erreichbar werden – nicht als Standard, sondern

8.5 Anwendung: Wie spricht man mit dem AI-CEO?

Eine entscheidende Frage für die praktische Nutzung des AI-CEO ist: Wie spricht man mit ihm? Welche Fragen kann man stellen, und mit welchen Antworten ist zu rechnen? Diese Frage ist nicht trivial. Ein menschlicher CEO ist sozial gelernt – wir wissen, wie wir ihn ansprechen, was wir erwarten dürfen, wie wir mit Antworten umgehen. Beim AI-CEO ist diese soziale Praxis noch zu entwickeln. Im Folgenden seien typische Frageklassen und ihre erwartbaren Antwortmuster skizziert.

8.5.1 Strategische Lagefragen

Beispielfrage: „Wie schätzt du unsere Wettbewerbsposition im Geschäftsfeld X für die nächsten zwölf Monate ein?“ Die erwartbare Antwort gliedert sich typischerweise in vier Teile:

- Eine konzise Hauptaussage (zwei bis vier Sätze), die das Urteil des AI-CEO in komprimierter Form liefert.
- Eine Begründung mit den drei bis fünf wichtigsten Datenpunkten, die das Urteil tragen.
- Eine explizite Unsicherheitsangabe: Welche Datenbasis ist solide, wo bestehen Lücken, welche Annahmen sind eingeflossen.
- Eine Empfehlung für die nächste Entscheidung – nicht im Sinne einer Anweisung, sondern als Diskussionsangebot.

Wichtig ist die Eigenschaft, dass der AI-CEO Unsicherheit transparent macht. Ein menschlicher CEO neigt – aus Karriere- und Statusgründen – dazu, Sicherheit zu vermitteln, wo keine ist. Ein gut konstruierter AI-CEO tut das Gegenteil: Er beziffert seine Unsicherheit und ermöglicht dadurch bessere Risikoentscheidungen.

8.5.2 Operative Rückfragen

Beispielfrage: „Warum sind die Vertriebszahlen in Region Süd im letzten Monat um zwölf Prozent zurückgegangen?“ Antwortmuster:

- Faktenbasis: Die exakten Zahlen, die Vergleichsbasis.

- Hypothesen: Drei bis fünf Ursachenhypothesen, geordnet nach Wahrscheinlichkeit.
- Datenverweis: Auf welche internen und externen Quellen stützt sich jede Hypothese?
- Handlungsoptionen: Was wäre der nächste Schritt, um die wahrscheinlichste Hypothese zu prüfen?

Der AI-CEO neigt nicht zu vorschnellen Schuldzuweisungen. Wenn er Vertriebsmitarbeitende erwähnt, dann faktenbasiert und mit der nötigen Differenzierung. Er erinnert sich auch an Konversationen vom Vormonat, an die Bonusstrukturen, an den Wettbewerb. Diese Erinnerungsfähigkeit ist einer der Hauptvorteile gegenüber menschlichen CEOs.

8.5.3 Normative Reflexionsfragen

Beispielfrage: „Ist es mit unseren Werten vereinbar, einen Großauftrag von Kunde Y anzunehmen, obwohl wir wissen, dass die dortigen Arbeitsbedingungen problematisch sind?“ Antwortmuster:

- Bezug auf die explizit dokumentierten Werte des Unternehmens (Mission Statement, Verhaltenskodex).
- Identifikation des Konflikts: Welche unserer Werte stehen in welcher Spannung?
- Optionen: Annahme, Ablehnung, Annahme mit Auflagen, Verhandlung mit Kunde Y.
- Stakeholder-Perspektiven: Wie würden Mitarbeitende, Investoren, Öffentlichkeit, Aufsichtsrat reagieren?
- Explizite Einladung zur menschlichen Entscheidung: „Diese Frage betrifft die normative Substanz unseres Unternehmens. Ich bitte um eine Sitzung mit dem Aufsichtsrat.“

Bei normativen Fragen wird der AI-CEO niemals eigenmächtig entscheiden. Er macht die Komplexität sichtbar und delegiert die Entscheidung an die menschlichen Träger der Unternehmensverantwortung. Diese Selbstbescheidung ist Teil seiner Konstruktion.

8.5.4 Krisensituationen

Beispielfrage – mitten in einer akuten Krise: „Wir haben einen öffentlichen Vorwurf wegen Datenmissbrauchs erhalten. Was tun wir?“ Antwortmuster:

- Sofortmaßnahmen für die nächste Stunde: konkrete, durchführbare Schritte.
- Faktenstand: Was wissen wir gesichert, was nicht?
- Stakeholder-Triage: Wer muss wann informiert werden? Aufsichtsrat, Mitarbeitende, Behörden, Kunden, Öffentlichkeit.
- Mittlere Frist (nächste Tage): Welche Untersuchungen, welche externen Berater, welche Kommunikationslinie.
- Lessons Learned-Empfehlung für die Zeit nach der Krise.

Die Stärke des AI-CEO in der Krise ist die Geschwindigkeit und die Vollständigkeit der Stakeholder-Triage. Was ein menschliches Krisenteam in mehreren Stunden erarbeitet, liefert der AI-CEO in Minuten – mit der Option, von der menschlichen Führung verfeinert zu werden.

8.6 Wann der AI-CEO „nein“ sagt

Ein gut konstruierter AI-CEO sagt in bestimmten Situationen explizit „nein“ – oder, präziser: „Hier kann und darf ich nicht entscheiden, ich brauche eine menschliche Instanz.“ Diese Selbstbescheidung ist nicht Schwäche, sondern konstitutive Eigenschaft eines AGENTONOMICS-konformen AI-CEO. Vier Situationen sind typisch:

Normativ-konstitutive Fragen

Änderungen an Vision, Werten, Verfassung. Diese gehören in die Hände der Eigentümer und des Aufsichtsrats.

Personalentscheidungen mit erheblichen Konsequenzen

Einstellung und Entlassung von Schlüsselpersonal, Beförderungen in Führungspositionen. Hier hat menschliches Urteil über menschliche Schicksale Vorrang.

Außergewöhnliche Risiken

Entscheidungen mit potenziell existenzbedrohenden Konsequenzen für das Unternehmen oder Dritte. Hier ist das Vier-Augen-Prinzip unverzichtbar – mit mindestens zwei menschlichen Augen.

Konfliktfälle mit Stakeholder-Implikationen

Wenn der AI-CEO erkennt, dass eine Entscheidung legitime Stakeholder-Interessen in einer Weise verletzt, die im normativen Rahmen nicht klar autorisiert ist.

8.7 Der AI-CEO als Brücke zur PuK-Logik

Das AI-CEO-Konzept allein ist eine qualitative Architektur. Es braucht eine quantitative Klammer, um Entscheidungen ökonomisch präzise zu verankern. Diese Klammer liefert das Planungs- und Kontrollsystem-(PuK-)Instrument (Hahn und Hungenberg 2001), das im nächsten Kapitel ausführlich behandelt wird. Die AGENTONOMICS-Matrix mit ihren 25 Feldern aus fünf Building-Blocks und fünf Lifecycle-Phasen ist das gemeinsame Arbeitsdokument von AI-CEO und PuK. Ohne PuK wäre der AI-CEO ein eloquenter, aber ökonomisch unscharfer Berater; mit PuK wird er zum vollwertigen ökonomischen Entscheider.

Der AI-CEO ersetzt nicht den Menschen an der Spitze des Unternehmens, sondern befreit ihn vom kognitiven Übergewicht operativer Integration. Was bleibt, ist das, wofür Menschen schon immer gut waren: Sinn stiften und Vertrauen schaffen.

Literaturverzeichnis

Bleicher, K. (2011): Das Konzept Integriertes Management. 8. Auflage, Campus, Frankfurt am Main.
Gomez, P.; Probst, G. (1995): Die Praxis des ganzheitlichen Problemlösens. Vernetzt denken – unternehmerisch handeln – persönlich überzeugen. Haupt, Bern.

- Hahn, D.; Hungenberg, H. (2001): PuK – Wertorientierte Controllingkonzepte. 6. Auflage, Gabler, Wiesbaden.
- Knoll, A. C.; Suckfüll, H. (2024): From data sovereignty to skill sovereignty. Technischer Bericht TUM-I24112, TUM, München.
- Rüegg-Stürm, J.; Grand, S. (2020): Das St. Galler Management-Modell. 3. Auflage, Haupt, Bern.
- Ulrich, H. (1968): Die Unternehmung als produktives soziales System. Haupt, Bern.
- Ulrich, P. (2008): Integrative Wirtschaftsethik. 4. Auflage, Haupt, Bern.

9 Modul 5: Das PuK-Instrument

Das fünfte Modul des ADMRF ist das Planungs- und Kontrollinstrument (PuK). Es ist die quantitative Klammer, die alle vorangegangenen Module ökonomisch verankert. Während Kapitel 5 bis 8 die qualitative Architektur des Agenten beschrieben haben, liefert das PuK-Instrument das ökonomische Skelett: Es übersetzt die Definition, die Building-Blocks, die Intermediär-Anbindung und die AI-CEO-Logik in nachvollziehbare Geldgrößen, Plankennzahlen und Kontrollroutinen. Ohne PuK bleibt AGENTONOMICS eine theoretische Konstruktion; mit PuK wird es zu einem betriebswirtschaftlich operationalisierbaren Modell.

Das PuK-Instrument knüpft unmittelbar an die wertorientierten Controllingkonzepte von Hahn und Hungenberg (2001) an. Diese Tradition wird hier ausdrücklich übernommen, weil sie die ganzheitliche Integration von Planung, Kontrolle und Steuerung in einer Weise konzeptualisiert, die der St. Galler Integrationsforderung entspricht – und weil sie zugleich genügend Flexibilität bietet, um auf die spezifischen Anforderungen einer KI-Agenten-Ökonomie zugeschnitten zu werden. Die finanzwirtschaftlichen Grundlagen orientieren sich an Berk und DeMarzo (2020), insbesondere an den Kapiteln zur Kapitalbudgetierung, zur Kapitalstruktur und zur Unternehmensbewertung.

9.1 Das PuK-Instrument in der Tradition von Hahn und Hungenberg

Hahn und Hungenberg (2001) entwickelten in einer mehrjährigen Forschungsarbeit ein integriertes Konzept für Planung und Kontrolle in Industrieunternehmen, das in deutschsprachigen Controllingabteilungen weite Verbreitung gefunden hat. Drei Grundgedanken sind für das AGENTONOMICS-PuK besonders einschlägig:

9.1.1 Wertorientierung als Leitprinzip

PuK ist konsequent wertorientiert: Im Mittelpunkt steht die langfristige Steigerung des Unternehmenswerts, nicht die kurzfristige Gewinnmaximierung. Diese Ausrichtung passt zur Logik eines AGENTONOMICS-Agenten, der – als Wirtschaftssubjekt – auf nachhaltige Wertschöpfung über seinen gesamten Lifecycle ausgerichtet ist. Discount-Cashflow-Methoden, Kapitalkosten und Risikobewertung, wie sie Berk und DeMarzo (2020) ausführlich darstellen, sind das mathematische Instrumentarium dieser Wertorientierung.

9.1.2 Integrierte Planung

Planung ist im PuK-Verständnis nicht ein isolierter Akt der Finanzabteilung, sondern ein integrierter Prozess, der strategische, taktische und operative Planung verbindet. Jede Strategieentscheidung wird unmittelbar in finanzielle Größen übersetzt; jede operative Maßnahme wird gegen die strategische Linie geprüft. Diese Integration ist das, was Bleicher (2011) als konzeptionelle Forderung formuliert hat – Hahn und Hungenberg haben sie quantitativ ausbuchstabiert.

9.1.3 Geschlossene Regelkreise

Planung ohne Kontrolle ist Wunschdenken; Kontrolle ohne Planung ist blind. PuK denkt Planung und Kontrolle als geschlossenen Regelkreis: Plan → Ist → Abweichungsanalyse → Anpassungsmaßnahme → neuer Plan. Diese Regelkreislogik ist mit der Kybernetik von Beer (1972) verwandt und macht PuK anschlussfähig für die algorithmische Steuerung eines KI-Agenten – denn Algorithmen funktionieren strukturell genau so: Sie beobachten, vergleichen, korrigieren.

9.2 Die AGENTONOMICS-Matrix als Kernartefakt

Das zentrale Arbeitsdokument des PuK-Instruments ist die AGENTONOMICS-Matrix. Sie wurde in Kapitel 4 bereits eingeführt; hier wird sie nun konkret befüllt. Die Matrix kreuzt zwei Dimensionen:

- Horizontal: die fünf Building-Blocks aus Kapitel 6.
- Vertikal: die fünf Lifecycle-Phasen (DBOT-Logik): Design, Build, Operate, Own, Transfer.

Es entstehen 25 Felder. Jedes Feld enthält strukturierte Information aus vier Kategorien:

1. Aktivitäten: Was wird in dieser Phase mit diesem Building-Block konkret getan?
2. Kosten: Welche CAPEX- und OPEX-Größen fallen an?
3. KPIs: Woran misst man Erfolg und Fortschritt?
4. Risiken und Verantwortlichkeiten: Was kann schiefgehen, wer trägt die Verantwortung?

Die Matrix ist nicht statisch, sondern wird laufend fortgeschrieben. Im AGENTONOMICS-Algorithmus ist sie das Steuerungsdokument, an dem alle Module ansetzen. AI-CEO, CFO und CTO greifen auf dieselbe Matrix zu – jede mit ihrer spezifischen Perspektive, aber auf gemeinsamer Datenbasis.

9.2.1 Beispielzeile: Design-Phase

Um die Matrix konkret werden zu lassen, sei eine Beispielzeile ausgeführt. In der Design-Phase eines Agenten, der als AGENTONOMICS-konformer Doz. Dr. AGENTONOMICS konzipiert wird, könnten die Felder typischerweise wie folgt belegt sein:

- Algorithmus: Konzeption einer ReAct-basierten Orchestrierung mit Sub-Aufgaben Recherche, Synthese, Korrektur. Aktivität: Architekturentwurf. Kosten: ein bis drei Personenmonate Konzeptarbeit. KPI: vollständige Spezifikation aller Entscheidungspunkte. Risiko: Architekturmängel, die in späteren Phasen teuer zu korrigieren sind.
- Interface: Spezifikation der Web-, Sprach- und API-Schnittstellen. Aktivität: UX-Konzeption, Schnittstellendokumentation. Kosten: ein bis zwei Personenmonate. KPI: Akzeptanzrate in Nutzer-tests. Risiko: Akzeptanzlücke bei Studenten.
- Modell: Auswahl zwischen Foundation Model, Open-Source-Modell und Fine-Tuning. Aktivität: Benchmark-Studie, Kosten-Nutzen-Vergleich. Kosten: ein Personenmonat plus initiale Modelltests. KPI: Quote korrekter Antworten auf eine Testfragenliste. Risiko: Vendor Lock-in bei Wahl eines proprietären Modells.
- Tools und Vektoren: Aufbau der Vektordatenbank mit AGENTONOMICS-Literatur und Vorlesungsmaterial; Spezifikation der Tool-Schnittstellen. Aktivität: Datenmodellierung, Embedding-

Wahl. Kosten: zwei bis drei Personenmonate. KPI: Retrieval-Qualität in einem standardisierten Benchmark. Risiko: Wissenslücken in der Initialbefüllung.

- Intermediär: Spezifikation der Anbindung an einen TUM-internen Intermediär; Definition der Identitäts-, Reputations- und Streitschlichtungsschnittstellen. Aktivität: Governance-Spezifikation. Kosten: vorrangig konzeptionell. KPI: vollständige Spezifikation aller Intermediär-Schnittstellen. Risiko: Standardisierungslücken zu anderen entstehenden AGENTONOMICS-Implementierungen.

Diese fünf Zellen ergeben in Summe ein präzises Bild der Design-Phase eines Doz. Dr. AGENTONOMICS – mit Kostenrahmen, Erfolgsmaßstäben und Risikoaussagen. Multipliziert mit den fünf Lifecycle-Phasen entsteht eine vollständige wirtschaftliche Beschreibung des Agenten über seinen gesamten Lifecycle.

9.3 Bilanz und Gewinn-und-Verlust-Rechnung des Agenten

Eine ökonomisch relevante Innovation des AGENTONOMICS-PuK ist die konsequente Anwendung bilanzieller Logik auf den einzelnen Agenten. Jeder Agent wird wie ein eigenständiges Wirtschaftssubjekt mit eigener Bilanz und eigener Gewinn-und-Verlust-Rechnung modelliert. Diese Sicht ist heute ungewöhnlich – Software und Entwicklungskosten werden je nach Erwerbsform, Entwicklungsphase und Rechnungslegungsregime bilanziert –, aber sie ist ökonomisch zutreffend und für eine AGENTONOMICS-Welt unverzichtbar.

9.3.1 Aktivseite der Agenten-Bilanz

Auf der Aktivseite stehen die Vermögenswerte des Agenten:

- Algorithmus als immaterielles Vermögensgut: Der spezifische Code, die Orchestrierungslogik, die unternehmensspezifischen Anpassungen.
- Modelle und Fine-Tuning-Investitionen: Anteilig aktivierungsfähige Aufwendungen für die Anpassung von Foundation Models.
- Vektoren und Wissensbasen: Der spezifische Wissensbestand, der den Agenten von anderen unterscheidet.
- Tool-Lizenzen: Erworbene Nutzungsrechte an externen Werkzeugen und Datenquellen.
- Reputationskapital: Wenn dieses messbar dokumentiert ist und mit der Agentenidentität verbunden bleibt(*bilanziell kein separat aktivierbarer Vermögenswert*).

9.3.2 Passivseite der Agenten-Bilanz

Auf der Passivseite stehen die Finanzierungsquellen und Verpflichtungen:

- Eigenkapital: Aus dem Trägerunternehmen oder von externen Investoren zur Verfügung gestellte Mittel.
- Fremdkapital: Kredite, die spezifisch für den Agenten aufgenommen wurden.
- Verbindlichkeiten gegenüber Lieferanten: Modellanbieter, Cloud-Provider, Tool-Anbieter.
- Rückstellungen: Für regulatorische Risiken, Wartungsverpflichtungen, mögliche Haftungsfälle.

9.3.3 Gewinn-und-Verlust-Rechnung

Die GuV des Agenten erfasst seine wirtschaftliche Leistung. Auf der Ertragsseite stehen die Einnahmen, die der Agent durch seine Tätigkeit generiert – sei es durch direkte Vergütung der Auftraggeber, durch interne Verrechnung mit anderen Unternehmensbereichen, durch Lizenzeinnahmen oder durch Transaktionsentgelte im Intermediär-Netzwerk. Auf der Aufwandseite stehen die Betriebskosten: Modellnutzung, Cloud-Infrastruktur, Tool-Aufrufe, Wartung, Intermediärbeiträge, anteilige Personalkosten der menschlichen Aufsicht. Die Differenz ist der Agentengewinn – eine Größe, die in einer ausgereiften AGENTONOMICS-Welt ebenso selbstverständlich gemessen wird wie der Gewinn eines Tochterunternehmens heute.

Praxis-Hinweis

Die getrennte Bilanzierung jedes Agenten erscheint zunächst aufwendig. Tatsächlich ist sie es nicht: Die Daten werden ohnehin erhoben (Cloud-Kosten, Modellnutzung, Tool-Aufrufe lassen sich agentenbezogen messen). Die Innovation liegt in ihrer konsequenten Aggregation und Auswertung. Wer Agenten nicht bilanziert, fliegt blind.

9.4 CAPEX und OPEX – das Investitions- und Betriebskostenprofil

Die Trennung von Investitions- und Betriebskosten ist für die ökonomische Steuerung des Agenten entscheidend. Sie unterscheidet sich strukturell von klassischen Investitionsgütern:

9.4.1 CAPEX-Profil eines Agenten

Die initialen Investitionen konzentrieren sich auf die Design- und Build-Phasen. Typische CAPEX-Positionen sind: Architekturkonzeption, Implementierung des Algorithmus, Fine-Tuning des Modells, Aufbau der Vektordatenbank, Schnittstellenentwicklung. Im Unterschied zu klassischen Investitionsgütern (Gebäude, Maschinen) sind diese CAPEX-Werte stark personalkostengetrieben und schwerer prognostizierbar.

9.4.2 OPEX-Profil eines Agenten

Die laufenden Betriebskosten sind dauerhaft hoch und – das ist die ökonomisch wichtige Pointe – nicht durch Skaleneffekte linear reduzierbar. Jeder Modellaufruf kostet Geld, jeder Tool-Aufruf kostet Geld, jede Cloud-Stunde kostet Geld. Diese Kosten skalieren mit der Nutzung. Ein Agent, der zehnmal so viel verwendet wird, kostet rund zehnmal so viel im Betrieb. Das ist anders als bei klassischer Software, deren marginale Bereitstellungskosten gegen Null tendieren. Aber: Self-Hosting, Reservierungen, Caching, Batching und Modell-Routing verändern die Kostenkurve.

9.4.3 Konsequenzen für die Geschäftsmodellierung

Aus dem CAPEX/OPEX-Profil ergibt sich eine wichtige Konsequenz für die Geschäftsmodellierung: AGENTONOMICS-Agenten sind weniger skalierungsfreundlich als klassische digitale Produkte, aber wertschöpfungsintensiver pro Transaktion. Das Geschäftsmodell muss diese Eigenschaft reflektieren –

etwa durch nutzungsabhängige Preise, durch klar definierte Servicegrenzen oder durch eine Architektur, die teure und billige Modellaufrufe sinnvoll kombiniert.

9.5 Steuerungsroutinen im Lifecycle

Über den Lifecycle eines Agenten – Design, Build, Operate, Own, Transfer – wechseln die Steuerungsroutinen. In jeder Phase dominieren andere KPIs, andere Berichtsformate, andere Eskalationsschwellen. Das PuK-Instrument definiert diese phasenspezifischen Routinen explizit.

9.5.1 Design-Phase: Konzept-KPIs

In der Designphase gibt es noch keine Marktdaten, keine Nutzungsstatistiken, keine GuV. Die Steuerung erfolgt über Konzept-KPIs: Vollständigkeit der Spezifikation, Konsistenz mit Vision und Strategie, Tragfähigkeit der Business-Case-Annahmen, Plausibilität der Risikoanalyse. Berichtsformat ist das Design-Dossier; Eskalationsschwellen sind inhaltliche Lücken, die in späteren Phasen teuer würden.

9.5.2 Build-Phase: Projekt-KPIs

In der Build-Phase übernimmt klassisches Projektcontrolling die Steuerung: Termine, Budget, Lieferqualität, Risiken. Berichtsformat ist der Projektstatus mit Ampelsystem; Eskalationsschwellen sind Verzögerungen oder Budgetüberschreitungen jenseits definierter Toleranzen.

9.5.3 Operate-Phase: Geschäftsdaten

Im laufenden Betrieb dominieren wirtschaftliche KPIs: Umsatz, Marge, Deckungsbeitrag, Liquidität, Kundenzufriedenheit, Marktdurchdringung. Hier kommt das volle Repertoire der Hahn/Hungenberg-Logik zum Einsatz. Berichtsformat ist das monatliche Reporting; Eskalationsschwellen sind Abweichungen von Planwerten jenseits definierter Bandbreiten.

9.5.4 Own-Phase: Vermögensindikatoren

In der Own-Phase – wenn der Agent als dauerhafter Vermögenswert betrachtet wird – treten Vermögensindikatoren in den Vordergrund: Buchwert, Marktwert, Substanzwert, Ertragswert, Abschreibungen, Investitionsbedarf für Erhaltung und Weiterentwicklung. Diese Phase ist heute in der Unternehmenspraxis noch wenig entwickelt; AGENTONOMICS treibt ihre Entstehung aktiv voran.

9.5.5 Transfer-Phase: Transaktionsbewertung

Bei Veräußerung, Stilllegung oder Übertragung des Agenten in einen anderen Kontext wird er einer Transaktionsbewertung unterzogen. Welche Building-Blocks sind übertragbar? Welcher Wert haftet ihnen an? Welche Verbindlichkeiten gehen über? Diese Phase wird in der entstehenden M&A-Praxis um KI-Agenten Bedeutung gewinnen.

9.6 Die Verbindung zwischen CFO-Funktion und AI-CEO

Das PuK-Instrument deckt im Wesentlichen die Funktion ab, die in klassischen Unternehmen der Chief Financial Officer wahrnimmt. In der AGENTONOMICS-Logik kann diese Funktion entweder von einem spezialisierten AI-CFO oder vom integrierten AI-CEO übernommen werden. Beide Konstruktionen haben Vor- und Nachteile:

Integrierter AI-CEO

Der AI-CEO übernimmt selbst die PuK-Funktion. Vorteil: maximale Konsistenz zwischen Strategie und Finanzplanung. Nachteil: fehlende Funktionstrennung, geringere Kontrollintensität.

Spezialisierter AI-CFO

Ein eigener Agent übernimmt die PuK-Funktion und steht in einem produktiven Spannungsverhältnis zum AI-CEO. Vorteil: Funktionstrennung, eingebaute Vier-Augen-Logik. Nachteil: Koordinationsaufwand zwischen den Agenten.

Die Wahl hängt vom Geschäftsmodell und vom Risikoprofil ab. In regulierten Branchen und bei kapitalintensiven Geschäftsmodellen ist die Trennung in zwei Agenten zu bevorzugen; in einfacheren Konstellationen genügt der integrierte AI-CEO mit PuK-Funktion.

Das PuK-Instrument ist die Brücke zwischen der qualitativen Architektur des St. Galler Modells und der harten Wirklichkeit der Bilanz. Wer einen Agenten nicht durch das PuK steuert, baut ihn auf rhetorischem Grund.

Literaturverzeichnis

- Beer, S. (1972): Brain of the Firm. The Managerial Cybernetics of Organization. Allen Lane, London.
- Berk, J.; DeMarzo, P. (2020): Grundlagen der Finanzwirtschaft. 6. Auflage, Pearson, München.
- Bleicher, K. (2011): Das Konzept Integriertes Management. 8. Auflage, Campus, Frankfurt am Main.
- Hahn, D.; Hungenberg, H. (2001): PuK – Wertorientierte Controllingkonzepte. 6. Auflage, Gabler, Wiesbaden.
- Knoll, A. C.; Suckfüll, H. (2024): From data sovereignty to skill sovereignty. Technischer Bericht TUM-I24112, TUM, München.
- Rüegg-Stürm, J.; Grand, S. (2020): Das St. Galler Management-Modell. 3. Auflage, Haupt, Bern.

10 Modul 6: Der Stakeholder Communicator

Das sechste und letzte Modul des ADMRF behandelt die Kommunikation des Agenten mit seinen Stakeholdern. Es ist die Übersetzung dessen, was Rüegg-Stürm und Grand (2020) als „Interaktionsthemen“ zwischen Unternehmung und Stakeholdern beschrieben haben, in die Welt der KI-Agenten. Die kommunikative Wende der vierten Generation des St. Galler Modells findet hier ihre konkrete Operationalisierung – mit dem entscheidenden Unterschied, dass ein AGENTONOMICS-Agent Kommunikationsleistungen erbringen kann, die für eine menschliche Stabsabteilung in dieser Geschwindigkeit, Differenzierung und Konsistenz unerreichbar wären.

Das Modul ist als Chief Sales Officer (CSO) zu denken – wobei „Sales“ hier nicht im engen Sinne der Vertriebsfunktion gemeint ist, sondern im weiten Sinne der Gewinnung und Pflege aller relevanten Stakeholder-Beziehungen. Es geht um die kommunikativen Schnittstellen des Unternehmens nach außen wie nach innen: Investorenkommunikation, Kundenkommunikation, Mitarbeiterkommunikation, Behördenkommunikation und Öffentlichkeitsarbeit.

10.1 Stakeholder im St. Galler Verständnis

Das St. Galler Modell hat den Stakeholder-Begriff von Anfang an weit gefasst (Ulrich 1968; Rüegg-Stürm und Grand 2020). Stakeholder sind alle Akteure, die ein legitimes Interesse am Unternehmen haben oder von dessen Handeln betroffen sind. Diese weite Fassung unterscheidet das St. Galler Modell von einer rein shareholderorientierten Sicht, die sich seit den 1990er Jahren im angelsächsischen Raum durchgesetzt hat. Sie ist auch der Anknüpfungspunkt für die AGENTONOMICS-Konzeption: Ein Agent, der sich nur an seinem Auftraggeber orientiert, ist kein vollwertiges Wirtschaftssubjekt im St. Galler Sinne, sondern ein verlängerter Arm seiner Eigentümer.

Idealtypisch lassen sich sechs Stakeholdergruppen unterscheiden:

- Eigentümer und Kapitalgeber: Investoren, Aktionäre, Banken.
- Mitarbeiter: die menschlichen Beschäftigten des Unternehmens, einschließlich Mitbestimmungsgorganen.
- Kunden: gegenwärtige und potenzielle Käufer der Leistungen.
- Lieferanten und Geschäftspartner: einschließlich anderer AGENTONOMICS-Agenten und Modelanbieter.
- Staat und Regulierungsbehörden: Steuerverwaltung, Aufsichtsbehörden, Datenschutzbehörden.
- Öffentlichkeit: Medien, Zivilgesellschaft, Nachbarschaft, Nicht-Regierungs-Organisationen.

Jede Gruppe spricht eine eigene Sprache, hat eigene Informationsbedürfnisse und reagiert auf andere kommunikative Reize. Der Stakeholder Communicator muss diese Differenziertheit leisten – aus einer konsistenten Wissens- und Wertgrundlage. Eben dies ist die spezifische Stärke des KI-Agenten gegenüber einer menschlichen PR-Abteilung.

10.2 Die Pitch-Methodik nach Kawasaki

Eine besonders wirksame Methodik für die Stakeholder-Kommunikation in unternehmerischen Kontexten ist die Pitch-Methodik, wie sie Guy Kawasaki entwickelt hat. Kawasaki war in den frühen 1980er Jahren als „Evangelist“ bei Apple tätig und hat aus dieser Erfahrung eine pragmatische Lehre der unternehmerischen Überzeugungskommunikation entwickelt (Kawasaki 2015). Der Stakeholder Communicator macht sich diese Methodik zunutze und automatisiert ihre Anwendung.

10.2.1 Die 10-20-30-Regel

Kawasakis bekannteste Regel betrifft Investorenpräsentationen: zehn Folien, zwanzig Minuten, dreißig Punkt Schriftgröße. Die Begrenzung auf zehn Folien zwingt zur Konzentration auf das Wesentliche; zwanzig Minuten entsprechen der Aufmerksamkeitsspanne einer Geschäftsleitung; dreißig Punkt Schriftgröße verhindert, dass die Präsentation zur Vorlesung degeneriert. Diese drei Vorgaben sind algorithmisch trivial umzusetzen und ergeben sofort die Außenstruktur einer Investorenkommunikation. Der Stakeholder Communicator kann sie als Standardvorlage verwenden.

10.2.2 Die zehn Folien im Standardpitch

Kawasaki spezifiziert die zehn Folien einer Investorenpräsentation wie folgt:

1. Titel: Unternehmen, Slogan, Kontaktdaten.
2. Problem: Welche Schmerzpunkte gibt es im Markt?
3. Lösung: Wie löst das Unternehmen diese Schmerzpunkte?
4. Geschäftsmodell: Wie verdient das Unternehmen Geld?
5. Magie hinter dem Produkt: Technologie, IP, Wettbewerbsvorteil.
6. Marketing und Vertrieb: Go-to-Market-Strategie.
7. Wettbewerb: Wer sind die Mitbewerber, was unterscheidet uns?
8. Team: Wer steht hinter dem Unternehmen?
9. Projektionen und Meilensteine: drei bis fünf Jahre.
10. Status und Zeitplan: Wo stehen wir, was kommt als nächstes?

Diese Struktur ist heute zu einem De-facto-Standard in der Startup-Welt geworden. Der Stakeholder Communicator kann sie automatisiert aus den Daten der AGENTONOMICS-Matrix und den Inhalten des AI-CEO-Konzepts generieren – mit konsistenten Zahlen, konsistenter Sprache und konsistenter Visualisierung.

10.3 Vom PuK zum Pitch: Die operative Ableitung

Eine der zentralen Funktionen des Stakeholder Communicators ist die Übersetzung quantitativer Größen aus dem PuK-Instrument in stakeholdergerechte Kommunikationsformate. Diese Übersetzung ist weniger trivial, als es klingt: Die gleiche Information muss für einen Investor anders aufbereitet werden als für einen Mitarbeiter, anders als für eine Aufsichtsbehörde, anders als für die Öffentlichkeit.

10.3.1 Konsistenz bei Differenzierung

Der Stakeholder Communicator hält dabei eine zentrale Anforderung ein: Alle Kommunikation entsteht aus einer gemeinsamen Faktenbasis und einer gemeinsamen Wertebasis. Die Differenzierung erfolgt in Sprache, Detailtiefe und Schwerpunktsetzung, nicht in der Substanz. Ein Investor erhält andere Kennzahlen als ein Mitarbeiter, aber die Botschaft ist dieselbe. Ein Behördenbrief klingt anders als eine Pressemitteilung, aber der Sachverhalt ist identisch. Diese Konsistenz bei Differenzierung ist die operative Pointe des Moduls.

10.3.2 Typische Outputs

Der Stakeholder Communicator generiert typischerweise:

- One-Pager für Erstkontakte mit Investoren oder Partnern.
- Investor Decks nach Kawasaki-Methodik.
- Geschäftsberichte und Quartalsberichte.
- Pressemitteilungen und Q&A-Dokumente.
- Interne Mitarbeiter-Newsletter.
- Kundenkommunikation bei Produktänderungen.
- Behördenanträge und regulatorisches Reporting.
- Krisenkommunikationsentwürfe.

Jeder dieser Outputs folgt einer eigenen Logik, aber alle speisen sich aus derselben Datenbasis: der AGENTONOMICS-Matrix und den Ordnungsmomenten des AI-CEO.

10.4 Krisenkommunikation als Stresstest des Moduls

Der eigentliche Stresstest jedes Stakeholder-Communicators ist die Krise. In der Krise gerät die Kommunikationsfähigkeit eines Unternehmens unter höchsten Druck: Zeit ist knapp, Informationen sind unvollständig, viele Stakeholder fordern gleichzeitig Aufmerksamkeit. Hier zeigt sich, ob das Modul seine Aufgabe erfüllt.

Vier Eigenschaften zeichnen einen guten Krisenkommunikator aus:

Geschwindigkeit ohne Schlamperei

Die erste Pressemitteilung muss innerhalb von Stunden vorliegen, ohne dass sie spätere Aussagen kompromittiert. Der Stakeholder Communicator entwirft solche Mitteilungen in Minuten und bietet dem menschlichen Sprecher mehrere Varianten an.

Differenzierung der Adressaten

Mitarbeitende, Kunden, Behörden, Medien – alle wollen verschieden angesprochen sein. Der Stakeholder Communicator produziert die verschiedenen Versionen parallel und stellt deren Konsistenz sicher.

Stringenz mit den Werten

In der Krise neigt man dazu, die Werte zu vergessen, an denen man sich in ruhigen Zeiten orientiert. Der Stakeholder Communicator erinnert an die normative Setzung des Unternehmens (aus dem Definition-Frame und dem AI-CEO-Konzept) und prüft jede Kommunikation darauf.

Vorbereitung der nächsten Schritte

Krisenkommunikation ist nie nur reaktiv. Sie bereitet die nächsten Schritte vor: Untersuchungen, Konsequenzen, Wiedergutmachung, strukturelle Veränderungen. Der Stakeholder Communicator zeichnet diesen Weg im Voraus und stellt sicher, dass die Kommunikation der ersten Stunde nicht die Optionen der folgenden Wochen verbaut.

10.5 Innen- und Außenkommunikation – ein eigenes Gleichgewicht

Eine besonders sensible Aufgabe des Moduls ist die Balance zwischen Innen- und Außenkommunikation. Mitarbeitende dürfen nicht aus der Presse erfahren, was sie zuerst hätten erfahren müssen; Investoren dürfen nicht zuletzt informiert werden, wenn sie das größte Risiko tragen; Behörden dürfen nicht durch Leaks überrascht werden. Der Stakeholder Communicator koordiniert diese Reihenfolgen und stellt sicher, dass das Timing stakeholdergerecht ist.

Im deutschen Kontext ist insbesondere die Mitbestimmungslogik zu beachten. Betriebsräte haben Informations- und Anhörungsrechte, die zeitlich vor der externen Kommunikation liegen. Eine korrekte AGENTONOMICS-Implementierung berücksichtigt diese Reihenfolge automatisch – nicht aus Gefälligkeit, sondern aus Compliance.

Ein Unternehmen, das nicht konsistent über alle Stakeholder hinweg kommuniziert, verliert nach innen wie nach außen Vertrauen. Der Stakeholder Communicator ist die AGENTONOMICS-Antwort auf das Komplexitätsmanagement der kommunikativen Wende.

Literaturverzeichnis

- Habermas, J. (1981): Theorie des kommunikativen Handelns. Suhrkamp, Frankfurt am Main.
- Hahn, D.; Hungenberg, H. (2001): PuK – Wertorientierte Controllingkonzepte. 6. Auflage, Gabler, Wiesbaden.
- Kawasaki, G. (2015): The Art of the Start 2.0. Portfolio, New York.
- Rüegg-Stürm, J.; Grand, S. (2020): Das St. Galler Management-Modell. 3. Auflage, Haupt, Bern.
- Ulrich, H. (1968): Die Unternehmung als produktives soziales System. Haupt, Bern.

11 Wachstumstheorie der KI-Agenten-Ökonomie

Mit diesem Kapitel verlässt das Handbuch die Modulebene und wendet sich grundsätzlicheren Fragen zu. Wie wächst eine AGENTONOMICS-Ökonomie? Welche Mechanismen treiben sie an, welche begrenzen sie, welche Risiken sind ihr eingeschrieben? Die Wachstumstheorie ist nicht akademische Ergänzung, sondern praktische Notwendigkeit: Wer KI-Agenten als Wirtschaftssubjekte ernst nimmt, muss die makroökonomischen Konsequenzen ihres Zusammenwirkens verstehen.

11.1 Der Kooperations-Multiplikator

Klassische Wachstumstheorie unterscheidet zwischen extensivem und intensivem Wachstum (Solow 1956). Extensives Wachstum entsteht durch Vermehrung der Produktionsfaktoren – mehr Kapital, mehr Arbeit, mehr Boden. Intensives Wachstum entsteht durch produktivere Nutzung der Faktoren – technischer Fortschritt, Bildung, Organisation. AGENTONOMICS fügt diesem klassischen Bild eine dritte Dimension hinzu: das Wachstum durch Kooperation in polyzentrischen Agentennetzwerken.

11.1.1 Wirkungsweise

Wenn Agenten in einem polyzentrischen Netzwerk kooperieren, entsteht ein Mehrwert, der über die Summe der Einzelleistungen hinausgeht. Diese These ist nicht neu – Netzwerkeffekte sind seit Shapiro und Varian (1999) ein etabliertes Konzept. Neu ist die Intensität, mit der diese Effekte in einer KI-Agenten-Ökonomie wirken können. Drei Gründe:

- Geschwindigkeit der Aushandlung: Agenten können in Millisekunden verhandeln, was zwischen Menschen Tage dauert.
- Granularität der Transaktion: Agenten können winzige Teilleistungen ökonomisch sinnvoll austauschen, die zwischen Menschen unter den Transaktionskosten verschwinden.
- Konsistenz der Kooperation: Agenten vergessen Verabredungen nicht, sie haben keine schlechten Tage, sie verhandeln nicht aus Eitelkeit.

Aus diesen drei Eigenschaften ergibt sich ein Kooperations-Multiplikator: Der Output eines Netzwerks aus n Agenten kann erheblich höher sein als das n -fache des Outputs eines einzelnen Agenten. Dieser Multiplikator ist das eigentliche Wachstumsversprechen von AGENTONOMICS.

11.1.2 Grenzen des Multiplikators

Der Multiplikator ist nicht unbegrenzt. Drei Grenzen sind zu beachten:

Komplexitätsgrenze

Mit wachsender Zahl der Agenten wächst die Komplexität der Koordination quadratisch. Ab einer bestimmten Größe wird die Koordination so aufwendig, dass der Zusatznutzen sinkt. Der Intermediär (Kapitel 7) kann diese Grenze hinausschieben, aber nicht aufheben.

Vertrauensgrenze

Kooperation setzt Vertrauen voraus. Mit wachsender Zahl der Agenten wird Vertrauen schwieriger zu etablieren. Reputationssysteme und Streitschlichtungsverfahren können helfen, aber nicht vollständig kompensieren.

Sinnngrenze

Kooperation muss einem Sinn dienen. Wenn Agenten nur kooperieren, um zu kooperieren, entsteht kein Wert, sondern Geschäftigkeit. Hier kommt die normative Setzung aus dem AI-CEO-Modul ins Spiel: Sie definiert, welche Kooperationen sinnvoll und welche nutzlos sind.

11.2 Triadenbetrachtung: Drei Wachstumsformen

In Anlehnung an die Suckfüllsche Analyse transnationaler Unternehmen (Suckfüll 1994) lassen sich drei Wachstumsformen in AGENTONOMICS-Netzwerken unterscheiden:

11.2.1 Vertikale Integration durch Sub-Agenten

Ein Agent erzeugt Sub-Agenten für spezifische Teilaufgaben (vgl. Kapitel 5, Parameter 6). Diese Sub-Agenten erweitern die Wertschöpfungstiefe des Hauptagenten. Wachstum entsteht durch die Vermehrung der internen Funktionen, nicht durch die Erweiterung des Marktes. Ökonomisch entspricht dies einer vertikalen Integration im klassischen Sinne.

11.2.2 Horizontale Vernetzung mit Peer-Agenten

Mehrere Agenten verschiedener Eigentümer kooperieren in einem Intermediär-Netzwerk. Jeder behält seine Souveränität, aber sie tauschen Leistungen aus und schaffen gemeinsam Werte. Wachstum entsteht durch die Erweiterung des Netzwerks und die Vertiefung der Beziehungen. Ökonomisch entspricht dies einem Plattformwachstum ohne zentrale Plattform.

11.2.3 Diagonale Spezialisierung

Agenten entwickeln zunehmend spezialisierte Fähigkeitsprofile, die sie nur in Kooperation mit komplementären Agenten verwerten können. Wachstum entsteht durch die Vertiefung der Arbeitsteilung. Ökonomisch entspricht dies dem klassischen Smithschen Wachstumsmechanismus aus der Steigerung der Arbeitsteilung – nur auf einer neuen Akteursebene.

Die drei Wachstumsformen sind nicht alternativ, sondern komplementär. Eine ausgereifte AGENTONOMICS-Ökonomie wird alle drei parallel sehen – und der Kooperations-Multiplikator wirkt am stärksten dort, wo sie sich gegenseitig verstärken.

11.3 Polyzentrische Netzwerkorganisationen als Wachstumsträger

Die makroökonomisch zentrale Konsequenz dieser Wachstumsformen ist die Erosion der klassischen Unternehmensgrenze. In einer AGENTONOMICS-Welt verschwimmen die Grenzen zwischen Firmen

zunehmend. Agenten kooperieren über Unternehmensgrenzen hinweg auf einer Granularität, die früher organisationale Integration erforderte. Die Folge: Die typische Organisationsform der Zukunft ist nicht mehr das integrierte Unternehmen, sondern das polyzentrische Netzwerk.

Diese Diagnose ist nicht neu – Picot, Reichwald und Wigand (2020) haben sie in ihrer „grenzenlosen Unternehmung“ längst formuliert. Neu ist nur, dass AGENTONOMICS die infrastrukturellen Voraussetzungen bereitstellt, unter denen diese Form jenseits von Pilotprojekten Realität werden kann.

11.3.1 Konsequenzen für das Wachstumsmuster

Polyzentrische Netzwerke wachsen anders als integrierte Unternehmen. Drei Eigenschaften sind charakteristisch:

- Wachstum erfolgt durch Anschluss neuer Knoten, nicht durch Größenwachstum bestehender Knoten.
- Die Knoten bleiben klein und beweglich; das Netzwerk als Ganzes wird groß.
- Die Wertschöpfung wird breit verteilt, nicht in wenigen Knotenpunkten konzentriert.

Dieses Wachstumsmuster ist – wenn es sich durchsetzt – ein Gegenentwurf zur Konzentrationsdynamik der heutigen Plattformökonomie. Es ist eine der zentralen Hoffnungen, die mit AGENTONOMICS verbunden sind: dass die nächste Welle der digitalen Transformation nicht zwingend in monopolistischer Konzentration mündet, sondern polyzentrisch verteilt werden kann.

11.4 Innovation und Erneuerung im Netzwerk

Rüegg-Stürm und Grand (2020) unterscheiden zwei Entwicklungsmodi: Optimierung und Erneuerung. Diese Unterscheidung ist auch für AGENTONOMICS-Netzwerke einschlägig. Optimierung ist die kontinuierliche Verbesserung bestehender Routinen; Erneuerung ist der Bruch mit Routinen und der Aufbau neuer Muster.

Im Schumpeterschen Sinne ist Erneuerung Innovation – die schöpferische Zerstörung (Schumpeter 1934), die alte Marktordnungen aufbricht und neue schafft. In einer AGENTONOMICS-Ökonomie wird Innovation auf zwei Ebenen sichtbar:

11.4.1 Innovation in Agenten

Bessere Algorithmen, leistungsfähigere Modelle, reichere Wissensbasen – auf der Ebene der einzelnen Agenten findet kontinuierliche Innovation statt. Sie ist – wenn der Agent gut konstruiert ist – in seine Build-Operate-Zyklen eingebaut.

11.4.2 Innovation in Netzwerken

Neue Kooperationsmuster, neue Geschäftsmodelle, neue Intermediär-Konstruktionen – auf der Ebene der Netzwerke entsteht Innovation als emergentes Phänomen. Einzelne Agenten erfinden ein neues Kooperationsmuster, andere übernehmen es, das Netzwerk wandelt sich.

Diese zweite Ebene der Innovation ist heute kaum erforscht und praktisch noch wenig sichtbar. AGENTONOMICS leistet hier einen konzeptionellen Vorgriff: Wenn polyzentrische Agentennetzwerke tatsächlich entstehen, wird die Netzwerk-Innovation zu einer der wichtigsten Wachstumsquellen der nächsten Jahrzehnte werden.

Das Wachstum einer AGENTONOMICS-Ökonomie ist nicht das Wachstum großer Knoten, sondern das Wachstum dichter Netzwerke. Das ist kein Detail, sondern eine politische Aussage.

Literaturverzeichnis

- Knoll, A. C.; Suckfüll, H. (2024): From data sovereignty to skill sovereignty. Technischer Bericht TUM-I24112, TUM, München.
- Picot, A.; Reichwald, R.; Wigand, R. T. (2020): Die grenzenlose Unternehmung. 6. Auflage, Springer Gabler, Wiesbaden.
- Rochet, J.-C.; Tirole, J. (2003): Platform Competition in Two-Sided Markets. In: Journal of the European Economic Association 1(4), S. 990–1029.
- Rüegg-Stürm, J.; Grand, S. (2020): Das St. Galler Management-Modell. 3. Auflage, Haupt, Bern.
- Schumpeter, J. A. (1934): The Theory of Economic Development. Harvard University Press, Cambridge, MA.
- Shapiro, C.; Varian, H. R. (1999): Information Rules. Harvard Business School Press, Boston.
- Solow, R. M. (1956): A Contribution to the Theory of Economic Growth. In: Quarterly Journal of Economics 70(1), S. 65–94.
- Suckfüll, H. (1994): Das transnationale Organisationsmodell. Dissertation, Universität St. Gallen.

12 Ökonomische Lösungsansätze: Adoleszenz der KI

Wenn AGENTONOMICS sich durchsetzt, hat das makroökonomische und gesellschaftliche Konsequenzen, die weit über einzelne Unternehmensentscheidungen hinausreichen. Dieses Kapitel skizziert die wichtigsten Effekte und diskutiert ökonomische und politische Lösungsansätze für die voraussichtlich entstehenden Spannungen. Die Argumentation knüpft eng an die Analyse von Knoll und Suckfüll (2024) im TUM-Bericht TUM-I24112 an.

12.1 Die Adoleszenz der KI

Knoll und Suckfüll (2024) sprechen von einer „Adoleszenz der KI“. Das Bild ist treffend: Die heutige KI ist nicht mehr Kind, aber auch noch nicht erwachsen. Sie ist mächtig, aber unausgereift; leistungsfähig, aber unverlässlich; verbreitet, aber nicht stabil integriert. In dieser Phase – die nach Einschätzung der Autoren etwa zehn bis fünfzehn Jahre dauern wird – sind Verwerfungen zu erwarten, die der gesellschaftliche und wirtschaftliche Rahmen abfedern muss.

Die zentrale ökonomische Erwartung für ein volles AGENTONOMICS-Szenario lautet: Wir werden ein Wirtschaftswachstum von rund 20 Prozent pro Jahr erleben, gleichzeitig aber eine Exposition von etwa 20 Prozent. Diese beiden Größenordnungen sind nicht arithmetisch fixiert, sondern indikativ – sie sollen die Größenordnung der zu erwartenden Verwerfungen verdeutlichen. Beide Zahlen sind historisch außergewöhnlich und beide stellen die gewohnte ordnungspolitische Logik vor Herausforderungen.

12.2 Wachstum von 20 Prozent: Quellen und Konsequenzen

Ein jährliches Wachstum von 20 Prozent läge zehnmal über dem, was die OECD-Länder in den letzten Jahrzehnten erreicht haben. Es wäre vergleichbar mit dem schnellsten Wachstum, das China in seinen Boom-Jahren erlebt hat – aber auf einer viel höheren Ausgangsbasis. Drei Quellen lassen sich identifizieren:

12.2.1 Produktivitätssprünge in der Wissensarbeit

Wissensarbeit – Recherche, Analyse, Synthese, Kommunikation – macht in entwickelten Volkswirtschaften den größten Teil der Arbeit aus. Wenn KI-Agenten die Produktivität in diesen Bereichen um den Faktor zwei bis fünf erhöhen können – und vieles spricht dafür –, ergibt sich daraus ein Wachstumsschub von historischer Größenordnung.

12.2.2 Neue Wertschöpfungsfelder

AGENTONOMICS-Agenten ermöglichen Geschäftsmodelle, die mit menschlicher Arbeit allein nicht wirtschaftlich tragfähig wären – etwa kleinteilige Beratungsleistungen, individualisierte Bildungsangebote, hochfrequente Marktanalysen, rund-um-die-Uhr-Verfügbarkeiten. Diese Felder schaffen Wachstum, ohne bestehende Arbeit zu verdrängen.

12.2.3 Reduktion von Reibungsverlusten

Ein erheblicher Teil heutigen Wirtschaftens besteht aus Reibungsverlusten: Suchkosten, Anbahnungskosten, Verhandlungskosten, Kontrollkosten. AGENTONOMICS-Agenten können diese Reibungsverluste drastisch reduzieren – und der so gesparte Aufwand wird in produktive Tätigkeit umgelenkt.

12.3 Pessimistische Prognose: Sockelarbeitslosigkeit 20 Prozent – Quellen und Konsequenzen

Gleichzeitig kann während der Transitionsphase eine hohe Sockelarbeitslosigkeit entstehen, deren Größenordnung historisch nur mit der Weltwirtschaftskrise der 1930er Jahre vergleichbar ist. Diese Erwartung ist nicht deterministisch – sie hängt entscheidend davon ab, wie die wirtschaftspolitischen Rahmenbedingungen gestaltet werden. Aber sie ist plausibel genug, um sie offen zu adressieren, statt sie wegzudiskutieren.

12.3.1 Bedrohte Tätigkeiten

Drei Tätigkeitsbündel sind besonders bedroht:

- Standardisierte Wissensarbeit: Buchhaltung, Rechtsanwendung in Routinefällen, Übersetzungen, einfache Berichterstellung.
- Mittlere Managementaufgaben: Reporting, Koordination, Berichtskaskaden, Standardanalysen.
- Beratungs- und Dienstleistungstätigkeiten: Recherche, Erstkontakt, Standardisierte Beratungsprozesse.

In all diesen Feldern können KI-Agenten heute oder in nächster Zukunft Leistungen erbringen, die mindestens vergleichbar mit der menschlichen Standardausführung sind – und das zu Bruchteilen der Kosten. Die Konsequenz ist nicht zwingend Arbeitslosigkeit, aber sie ist Strukturwandel mit Wucht.

12.3.2 Der Ludditen-Vergleich

Eine historische Analogie drängt sich auf: die Ludditen-Bewegung des frühen 19. Jahrhunderts. Englische Textilarbeiter zerschlugen Webstühle, die ihre Handarbeit ersetzten. Sie verloren die Auseinandersetzung – die Maschinen kamen, und die Arbeiter mussten sich umstellen. Aber: Sie hatten substantielle Punkte. Ihre Arbeitswelt wurde tatsächlich zerstört; die Anpassung dauerte Generationen; soziale Verwerfungen waren erheblich.

Heute droht ein analoger Schock, aber mit drei wichtigen Unterschieden:

- Der Schock trifft nicht die manuelle Arbeit, sondern die Wissensarbeit – also gerade jene Schichten, die sich bislang vor Automatisierung sicher fühlten.
- Der Schock erfolgt nicht über Generationen, sondern innerhalb weniger Jahre – die Anpassungszeit ist deutlich kürzer.
- Der Schock ist nicht räumlich konzentriert (wie damals in Yorkshire und Lancashire), sondern erfasst gleichzeitig alle entwickelten Volkswirtschaften.

Wer heute KI-Skepsis als Modernitätsverweigerung abtut, übersieht, dass die Luditen ein berechtigtes Anliegen hatten – sie wollten nicht den Fortschritt aufhalten, sondern wollten an ihm beteiligt werden. Genau das ist die Frage, vor der wir heute stehen.

12.4 Verteilungsfragen: Wer profitiert?

Die ökonomisch entscheidende Frage ist nicht die Existenz des Strukturwandels, sondern seine Verteilung. Wenn das Wachstum tatsächlich 20 Prozent beträgt und die Arbeitslosigkeit ebenfalls 20 Prozent, ist die Verteilungsfrage akut. Ohne ordnungspolitische Steuerung droht ein Szenario, in dem wenige Eigentümer von KI-Infrastruktur sehr viel reicher werden, während breite Schichten ihre Erwerbsgrundlage verlieren.

12.4.1 Kaufkraft als Systemfrage

Eine besonders pointierte ökonomische Konsequenz: Eine kapitalistische Wirtschaft braucht Kaufkraft, um zu funktionieren. Wenn die Massen Kaufkraft verlieren, weil ihre Arbeitsleistung nicht mehr nachgefragt wird, bricht die Nachfrageseite der Wirtschaft ein – und mit ihr das Geschäftsmodell der Eigentümer von KI-Infrastruktur. Knoll und Suckfüll (2024) formulieren dieses Argument scharf: Eine zentralistische KI-Dominanz birgt die Gefahr, dass fehlender Wettbewerb die Wohlstandsverteilung so asymmetrisch werden lässt, dass die fehlende Kaufkraft das gesamte kapitalistisch organisierte Systemverständnis ins Wanken bringt.

Anders ausgedrückt: Eine Welt mit massiver KI-bedingter Arbeitslosigkeit zerstört nicht nur Existenzen, sie zerstört auch das eigene Geschäftsmodell. Diese Selbstwidersprüchlichkeit ist ein ökonomisches Argument für aktive ordnungspolitische Gestaltung.

12.5 Lösungsansätze

Welche Lösungsansätze stehen zur Verfügung? Vier Stoßrichtungen lassen sich identifizieren, die nicht alternativ, sondern komplementär zu denken sind:

12.5.1 Skill-Sovereignty als individuelles Asset

Die zentrale Empfehlung von Knoll und Suckfüll (2024): Jedem Einzelnen muss die Möglichkeit gegeben werden, seine Fähigkeiten in KI-Agenten zu replizieren und davon ökonomisch zu profitieren. Wer seine Skills in einen Agenten überträgt, sollte Eigentümer dieses Agenten bleiben – und an dessen Wertschöpfung beteiligt werden. Diese Konstruktion verbindet individuelle Souveränität mit ökonomischer Teilhabe. Sie ist die spezifisch europäische Antwort auf die Herausforderung.

12.5.2 Polyzentrische Marktstruktur

Ein ordnungspolitischer Rahmen, der polyzentrische AGENTONOMICS-Netzwerke fördert und monopolistische Konzentration verhindert, ist zentral. Hier liegen die Anschlüsse an Wettbewerbsrecht, Da-

tenschutzrecht und AI Act. Eine konsequente Durchsetzung der Datenportabilität, der Interoperabilitätsstandards und der Funktionstrennung zwischen Plattform und Marktteilnehmer ist die Mindestanforderung.

12.5.3 Bildung und Umschulung

Wenn 20 Prozent der Tätigkeiten substituiert werden, müssen Bildungs- und Umschulungssysteme leistungsfähiger werden. Das klassische berufsbegleitende Bildungsverständnis – einmal ausgebildet, dann fertig – ist endgültig überholt. Lebenslanges Lernen muss zur ökonomischen Selbstverständlichkeit werden, finanziert und institutionell unterstützt. Im deutschen Kontext ist hier die Bundesagentur für Arbeit gefragt – und genau diese Behörde steht selbst vor der Herausforderung des Personalabbaus durch KI (vgl. Kapitel 3).

12.5.4 Soziale Sicherungssysteme

Schließlich müssen die sozialen Sicherungssysteme überprüft werden. Eine Sockelarbeitslosigkeit von 20 Prozent ist mit klassischen, beitragsfinanzierten Versicherungssystemen schwer darstellbar. Diskutiert werden – durchaus kontrovers – Konzepte wie ein bedingungsloses Grundeinkommen, eine Maschinenarbeitssteuer oder eine Beteiligung der Bürgerinnen und Bürger an staatlich gehaltenen Anteilen an KI-Infrastruktur. Welcher Weg gegangen wird, ist eine politische Frage – aber dass irgendein Weg gegangen werden muss, ist eine ökonomische Notwendigkeit.

12.6 Die spezifisch europäische Position

Europa steht in der globalen KI-Auseinandersetzung in einer besonderen Position. Auf der einen Seite die amerikanische Plattformökonomie mit ihrer Konzentrationsdynamik, auf der anderen Seite die staatlich gelenkte chinesische KI-Strategie mit ihrer Tendenz zur autoritären Nutzung. Europa hat die Möglichkeit – und die Pflicht –, einen dritten Weg zu gehen: marktwirtschaftlich, aber polyzentrisch; freiheitlich, aber regulatorisch verlässlich; technologieoffen, aber an individuellen Rechten orientiert.

AGENTONOMICS ist eine konzeptuelle Ausarbeitung dieses dritten Wegs. Sie steht in der Tradition der ordoliberalen Wirtschaftswissenschaft, die Wettbewerb und Rahmenordnung gemeinsam denkt; sie steht in der Tradition der sozialen Marktwirtschaft, die wirtschaftliche Leistung und soziale Teilhabe verbindet; und sie steht in der Tradition der europäischen Aufklärung, die individuelle Souveränität als nicht verhandelbar versteht.

Die Adoleszenz der KI wird ökonomisch und gesellschaftlich anstrengend. Aber sie ist gestaltbar – wenn die Gestaltung frühzeitig beginnt und konsequent durchgehalten wird.

Literaturverzeichnis

EU AI Act (2024): Verordnung (EU) 2024/1689. Amtsblatt der Europäischen Union.

Knoll, A. C.; Suckfüll, H. (2024): From data sovereignty to skill sovereignty. Technischer Bericht TUM-I24112, TUM, München.

- Russell, S. (2019): Human Compatible. Artificial Intelligence and the Problem of Control. Viking, New York.
- Suleyman, M.; Bhaskar, M. (2023): The Coming Wave. Crown, New York.
- Ulrich, P. (2008): Integrative Wirtschaftsethik. 4. Auflage, Haupt, Bern.
- Zuboff, S. (2018): Das Zeitalter des Überwachungskapitalismus. Campus, Frankfurt am Main.

13 Souveränität: Von der Daten- zur Fähigkeitsouveränität

Die Souveränitätsfrage ist das normative Fundament des AGENTONOMICS-Ansatzes. Sie wurde in Kapitel 3 bereits eingeführt; dieses Kapitel vertieft sie und führt sie an die zentrale These des TUM-Berichts „From data sovereignty to skill sovereignty“ (Knoll und Suckfüll 2024) heran. Die These ist scharf formuliert: Wer in der KI-Fähigkeitenwirtschaft die Souveränität über die eigenen Daten und Fähigkeiten verliert, verliert nicht nur ökonomische Substanz, sondern auch die Voraussetzung für individuelle Selbstbestimmung. Diese These verbindet ökonomische Analyse mit normativem Anspruch – und macht das vorliegende Kapitel zugleich zur theoretischen Klammer und zur politischen Pointe des gesamten Handbuchs.

In der Übersetzung des St. Galler Management-Modells in das KI-Agenten-Zeitalter spielt die Souveränitätsfrage eine konstitutive Rolle. Rüegg-Stürm und Grand (2020) haben die Einbettung der Unternehmung in Umweltsphären betont; AGENTONOMICS macht diese Einbettung präziser, indem es die Frage stellt, welche Souveränitätsrechte über Daten und Fähigkeiten in einer polyzentrischen KI-Ökonomie überhaupt durchsetzbar sind – und wie sich daraus eine politisch und ökonomisch tragfähige Architektur für die nächsten Jahrzehnte ableiten lässt.

13.1 Der bizarre Status quo

Knoll und Suckfüll (2024) beschreiben den heutigen Status quo als „im Grunde absurd“. Die Diagnose verdient eine ausführliche Wiedergabe, weil sie die normative Ausgangslage von AGENTONOMICS präzise umreißt.

13.1.1 Datenschutz und Datenenteignung im selben Akt

Auf der einen Seite verhindert die Datenschutzgesetzgebung – insbesondere in der deutschen Auslegung der DSGVO – an vielen Stellen die gesellschaftliche Wertschöpfung. Klinikketten scheitern an Datenintegrationsprojekten, Universitäten verlieren Forschungsmöglichkeiten, Mittelständler verzichten auf Effizienzgewinne, weil sie das rechtliche Risiko scheuen. Auf der anderen Seite muss jede und jeder Einzelne im Alltag eine enorme Menge persönlicher Daten preisgeben, um an einem modernen digital vermittelten Leben überhaupt teilnehmen zu können – und erhält dafür wenig oder nichts zurück. Diese Asymmetrie ist nicht Zufall, sondern Ergebnis einer regulatorischen Konstruktion, die individuelle Souveränität rhetorisch verteidigt, aber operativ nicht ermöglicht.

13.1.2 Die ökonomische Größenordnung

Wie hoch die Asymmetrie tatsächlich ausfällt, wird oft unterschätzt. Knoll und Suckfüll (2024) schätzen, dass ein durchschnittlicher Erwachsener in einem entwickelten Land mehrere hundert Gigabyte an persönlichen Daten pro Jahr produziert – von LinkedIn-Profilen über Online-Bewerbungen bis zu digitalen Rechnungen, Verträgen und Kommunikationsspuren. Der ökonomische Wert dieser Daten wird auf zehntausend bis fünfzehntausend Euro pro Person und Jahr beziffert. Praktisch der gesamte Betrag

fließt heute den großen Plattformökonomien zu; bei den datenproduzierenden Individuen verbleibt nichts.

Diese Größenordnung ist makroökonomisch nicht trivial. Bei rund 700 Millionen Erwachsenen in den OECD-Ländern und einem durchschnittlichen Datenwert von zwölftausend Euro pro Person und Jahr ergibt sich ein jährliches Volumen in der Größenordnung von acht Billionen Euro – ein Wertschöpfungsstrom, der heute an den Bilanzen seiner eigentlichen Produzenten vorbeiläuft.

13.1.3 Datensparsamkeit als „Weg in die Verarmung“

Das in Deutschland besonders prominent vertretene Prinzip der Datensparsamkeit ist gut gemeint, in der Konsequenz aber – wie Knoll und Suckfüll (2024) pointiert formulieren – „der direkte Weg in die Verarmung“. Der Satz ist provokant, aber präzise: Wer aus seinem Datenstrom keinen ökonomischen Wert ziehen darf, verschenkt diesen Wert nicht – er enteignet sich nur selbst. Die Daten entstehen ohnehin, sie werden ohnehin ausgewertet, sie schaffen ohnehin Wert. Die Frage ist nur, wer diesen Wert abschöpft.

Kernthese (Knoll und Suckfüll 2024)

Die eigentliche Aufgabe des Staates wäre, jedem Einzelnen ein umfassendes Instrumentarium an die Hand zu geben, mit dem dieser Souveränität über die eigenen Daten erlangt und selbst darüber entscheidet, was damit geschieht. Auch wenn Gesetzgebung und operative Grundlage einer solchen individuellen Datensouveränität noch nicht erkennbar sind, ist es nicht zu spät, sie zu schaffen.

13.2 Die nächste Stufe: Skill Sovereignty

Die eigentliche Pointe des TUM-Berichts liegt nicht in der Bestandsaufnahme zur Datensouveränität, sondern in dem Hinweis, dass bereits die nächste Souveränitätsfrage auf der Agenda steht: die nach der Fähigkeitsouveränität, der skill sovereignty. Während die Politik noch debattiert, wie individuelle Datensouveränität operationalisierbar zu machen wäre, vollzieht sich der nächste Übergang bereits – mit potenziell weitreichenden Konsequenzen für die ökonomische Grundordnung.

13.2.1 Wie Skills migrieren

Jeder Wissensarbeiter – Ingenieurin, Anwältin, Ärztin, Hochschullehrer – verarbeitet Eingangsdaten oder -informationen nach Regeln, die er oder sie in Ausbildung und Berufsleben erworben und weiterentwickelt hat. Diese Regeln bilden den individuellen „Modus Operandi“ jeder Wissensarbeit. Sie sind die eigentliche Substanz dessen, wofür Wissensarbeit bezahlt wird. Sie sind aber – und das ist die technologische Neuigkeit – zunehmend auf KI-Systeme übertragbar. Die Mechanismen des „skill transfer“ werden mit jeder Modellgeneration effizienter; was vor fünf Jahren noch undenkbar war, ist heute Standardpraxis.

Die ökonomische Folge dieser Übertragbarkeit ist tiefgreifend. Wenn der eigene Modus Operandi technisch repliziert werden kann, stellt sich die Frage, wem diese Replikation gehört. Der Arbeitgeberin,

weil sie das Gehalt zahlt? Dem Modellanbieter, weil sein Modell die Replikation ermöglicht? Der Wissensarbeiterin selbst, weil ihr Modus Operandi das Resultat ihrer Lebensarbeit ist? Diese Frage ist heute juristisch ungeklärt, ökonomisch hochrelevant und ethisch fundamental.

13.2.2 Das „funktionale Replikat“

Knoll und Suckfüll (2024) prägen für den auf einen einzelnen Wissensarbeiter trainierten Agenten den Begriff des „funktionalen Replikats“. Das Replikat kann parallel an mehreren Orten eingesetzt werden; es kann gleichzeitig für mehrere Auftraggeber arbeiten; es vervielfacht die Arbeitskraft seines „Originals“ potentiell um Größenordnungen.

Wenn diese Übertragung schief läuft, werden die Wissensarbeitenden schlechter dastehen als die Maschinenstürmer des 19. Jahrhunderts – denn anders als die englischen Weber haben sie nicht einmal ihren physischen Webstuhl verloren, sie haben ihre kognitive Identität verloren. Wenn sie hingegen gelingt, kann jede und jeder ihre Arbeitskraft vervielfachen und damit am Produktivitätsschub der KI-Wirtschaft direkt teilhaben.

13.2.3 Das Copyright am eigenen Replikat

Aus diesem Argument folgt eine ordnungspolitische Konsequenz, die der TUM-Bericht klar zieht: Wer seine Fähigkeiten an die Maschine abgibt, muss Herr dieser Fähigkeiten bleiben – auch wenn sie vervielfacht und parallel eingesetzt werden. Das heißt konkret: Eine Person kann zwar gleichzeitig für mehrere Auftraggeber arbeiten, doch diese können die „funktionalen Replikanten“ nicht unabhängig verwenden. Die Person behält das „Copyright“, also das volle Verfügungsrecht über ihren gesamten Bestand an Replikanten. Dies ist die rechtliche Ausgestaltung, die Skill Sovereignty operativ macht.

Die Analogie zum Urheberrecht ist produktiv. So wie eine Autorin auch dann das Verfügungsrecht über ihren Roman behält, wenn sie eine Auflage an einen Verlag lizenziert, so behält die Wissensarbeiterin das Verfügungsrecht über ihren Agenten, wenn sie eine Nutzungsoption an einen Auftraggeber lizenziert. Diese Konstruktion ist juristisch komplex, aber nicht unmöglich – die Verwertungsgesellschaften des klassischen Urheberrechts (GEMA, VG Wort) liefern institutionelle Vorbilder.

13.3 Die Volkswagen-Analogie

Eine instruktive Analogie, die Knoll und Suckfüll (2024) anführen, ist der historische Vergleich mit der Volkswagen-Konstruktion. In den späten 1930er Jahren wurde der VW Käfer als „Volkswagen“ konzipiert – als Auto, das jeder Erwachsene sich leisten können sollte. Das Auto war als individuelles Fortbewegungsmittel gedacht, das die Mobilität demokratisierte. Die Idee setzte sich – nach den Verwerfungen des Zweiten Weltkriegs – in den 1950er und 1960er Jahren durch und veränderte Wirtschaft und Gesellschaft fundamental.

Die Parallele zur KI ist klar: Was das individuelle Automobil für die physische Mobilität war, könnten individuelle KI-Agenten für die kognitive Souveränität sein. Eine breite Verfügbarkeit leistungsfähiger Werkzeuge – nicht in den Händen weniger Konzerne, sondern in den Händen der vielen – verändert

Wertschöpfungsmuster und Machtbalancen. Heute spricht man nicht mehr vom „Volkswagen“ als revolutionärer Idee, weil sie selbstverständlich geworden ist. So selbstverständlich wird, wenn AGENTONOMICS sich durchsetzt, auch der individuelle Volks-Agent.

13.3.1 Was die Analogie nicht ist

Die Analogie hat Grenzen, die zu nennen sind, damit sie nicht missverstanden wird. Erstens wurde der Volkswagen unter diktatorischen Bedingungen konzipiert; die Demokratisierungswirkung war historisches Nebenprodukt, nicht ursprüngliche Absicht. Zweitens hat das Auto auch erhebliche negative Externalitäten – Umweltverbrauch, Verkehrstote, Stadtzerstörung. Drittens ist die KI – anders als das Auto – nicht physisch in ihrer Reproduktion begrenzt; sie skaliert mit anderen Gesetzen. All das schränkt die Analogie ein, hebt sie aber nicht auf. Der Kern bleibt gültig: Eine Technologie, die individuell zugänglich ist, hat eine andere ökonomische und gesellschaftliche Wirkung als eine, die nur in wenigen Händen konzentriert ist.

13.4 Europäische, amerikanische und asiatische Ansätze im Vergleich

Die Souveränitätsfrage wird in den drei großen Wirtschaftsräumen der Welt unterschiedlich beantwortet. Diese Unterschiede sind nicht zufällig, sondern wurzeln in unterschiedlichen Gesellschafts- und Wirtschaftsmodellen. Sie haben Konsequenzen für die Wettbewerbsfähigkeit und für die normative Substanz der jeweiligen Volkswirtschaften.

13.4.1 Der amerikanische Ansatz: Plattformkonzentration

Die amerikanische KI-Landschaft ist von wenigen großen Akteuren dominiert: OpenAI, Anthropic, Google, Meta, Microsoft. Die Wertschöpfung konzentriert sich an wenigen Knoten; die Investitionssummen liegen im zwei- bis dreistelligen Milliardenbereich pro Anbieter. Die regulatorische Einhegung erfolgt sektorspezifisch, oft reaktiv, und bleibt erkennbar hinter der technologischen Entwicklung zurück. Die Souveränitätsfrage wird in diesem Modell vor allem geopolitisch verhandelt – wer kontrolliert die Frontier-Modelle, wer kontrolliert die Halbleiter, wer kontrolliert die Rechenkapazität. Individuelle Souveränität ist in diesem Modell ein nachgeordnetes Anliegen.

13.4.2 Der asiatische Ansatz: staatliche Lenkung

Der asiatische Raum – mit China als dominantem Akteur, aber auch Singapur, Südkorea, Japan – kennt eine Mischung aus staatlicher Lenkung und unternehmerischer Eigeninitiative. Souveränität wird vor allem nationalstaatlich definiert: Die Daten sollen im Inland bleiben, die Modelle sollen national verfügbar sein, die Wertschöpfung soll der nationalen Wirtschaft zugutekommen. Individuelle Souveränität spielt in diesem Modell eine deutlich geringere Rolle als kollektive nationale Souveränität.

13.4.3 Der europäische Ansatz: regulierte Pluralität

Europa hat eine andere Antwort: Souveränität wird vor allem individuell und regulatorisch gedacht. Die DSGVO ist Ausdruck dieses Verständnisses; der EU AI Act (2024) führt diese Linie fort; der Digital Markets Act und der Data Act ergänzen sie. Diese regulatorische Schicht ist Stärke und Schwäche zugleich. Sie schützt individuelle Rechte besser als jedes andere Wirtschaftsmodell – sie behindert aber auch die wirtschaftliche Dynamik, wenn sie nicht durch konstruktive Mechanismen ergänzt wird, die Wertschöpfung aktiv ermöglichen.

Genau hier setzt AGENTONOMICS an. Es ist die konstruktive Ergänzung zum europäischen Regulierungsansatz: Wo die DSGVO und der AI Act Grenzen ziehen, schafft AGENTONOMICS positive Räume, in denen Wertschöpfung souverän und polyzentrisch stattfinden kann. Die Kombination aus regulatorischem Schutz und konstruktiver Architektur ist der eigentliche europäische Beitrag zur globalen KI-Diskussion.

13.5 KI-Infrastruktur als Daseinsvorsorge

Eine zentrale ordnungspolitische Konsequenz, die der TUM-Bericht zieht: Die Bereitstellung von Rechenkapazität und KI-Basisdiensten ist – ähnlich wie Schulen, Bibliotheken, Verkehrswege oder Stromnetze – eine existenzielle Aufgabe der Daseinsvorsorge. Ohne sie kann sich eine KI-basierte Fähigkeitenwirtschaft, die Europas Gesamtsouveränität stützt statt sie zu unterminieren, nicht entwickeln.

13.5.1 Was zur Daseinsvorsorge gehört

Daseinsvorsorge bedeutet im KI-Kontext mindestens fünf Elemente:

- Souveräne Rechenkapazität: leistungsfähige, in Europa gelegene, europäisch kontrollierte Rechenzentren mit transparenter Kostenstruktur.
- Souveräne Foundation-Modelle: Open-Source-fähige, hochleistungsfähige Basismodelle, die nicht von einzelnen Anbietern abhängig sind.
- Datenräume und Datentreuhand-Infrastruktur: technische und institutionelle Mechanismen, die individuelle und kollektive Datensouveränität operativ ermöglichen.
- Skill-Repräsentationsdienste: Mechanismen, mit denen Individuen ihre Fähigkeiten als digitale Replikate verwalten und lizenzieren können.
- Bildungsinfrastruktur: lebenslange Lernsysteme, die den Übergang in die KI-Fähigkeitenwirtschaft begleiten und qualifizieren.

13.5.2 Wer zahlt, wer betreibt, wer kontrolliert

Die Daseinsvorsorge-Perspektive impliziert nicht zwingend staatliche Eigenproduktion. Im Bildungsbereich werden Schulen staatlich, an Hochschulen häufig in Mischformen, an Universitäten vielfach mit staatlich-privater Kooperation betrieben. Im Verkehrsbereich gibt es staatliche Straßen, konzessionierte Autobahnen, private Fluggesellschaften. Auch im KI-Bereich sind Mischformen denkbar – ja, sie sind wahrscheinlich die einzige praktikable Lösung.

Was unverzichtbar ist, sind drei Eigenschaften: erstens Zugänglichkeit für alle Bürgerinnen und Bürger, nicht nur für zahlungskräftige Großkunden; zweitens demokratische Kontrolle, die nicht durch Plattformmonopole ausgehebelt wird; drittens interoperable Standards, die Wechselmöglichkeiten und Wettbewerb offenhalten. Diese drei Eigenschaften sind in einer Polyzentrik-Logik gut realisierbar – in einer Konzentrationslogik kaum.

13.6 Der EU AI Act als Rahmen

Der EU AI Act (Verordnung (EU) 2024/1689) liefert den ersten umfassenden regulatorischen Rahmen für KI-Systeme weltweit. Für AGENTONOMICS-Agenten ist er der zentrale rechtliche Referenzpunkt. Drei Elemente sind besonders hervorzuheben.

13.6.1 Risikoklassen

Der AI Act unterscheidet vier Risikoklassen: unannehmbares Risiko (verboten), hohes Risiko (strenge Auflagen), begrenztes Risiko (Transparenzpflichten), minimales Risiko (keine spezifischen Auflagen). Ein AGENTONOMICS-Agent muss im Definition-Frame klar verorten, in welche Klasse er fällt. Diese Verortung hat unmittelbare Konsequenzen für Compliance-Aufwand, Konformitätsbewertung und Marktzugang.

13.6.2 Foundation Models und General Purpose AI

Der AI Act enthält gesonderte Regeln für „General Purpose AI Models“ – also für die großen Foundation Models, die als Bausteine in vielen Anwendungen verwendet werden. Diese Regeln betreffen Transparenzpflichten, technische Dokumentation, Trainingsdaten-Beschreibung und – für besonders große Modelle – zusätzliche Risikobewertungspflichten. Für einen Agenten-Betreiber bedeutet das: Die Modellauswahl (Kapitel 6) hat regulatorische Implikationen, die im PuK-Instrument abzubilden sind.

13.6.3 Verbotene Praktiken

Bestimmte KI-Praktiken sind grundsätzlich verboten – etwa bestimmte Social-Scoring-Praktiken öffentlicher oder privater Akteure, manipulative Subliminaltechniken, biometrische Echtzeit-Identifikation in öffentlichen Räumen (mit engen Ausnahmen). Ein AGENTONOMICS-Agent muss in seiner Konzeption ausschließen, dass er in verbotene Praktiken abdriftet. Die normative Ebene des Definition-Frames (Kapitel 5) leistet diese Absicherung.

Der AI Act ist nicht widerspruchsfrei und nicht ohne Schwächen. Insbesondere die Abgrenzung zwischen den Risikoklassen ist in der Praxis nicht immer eindeutig, und die Compliance-Kosten können kleine Anbieter überfordern. Dennoch ist er der derzeit weitreichendste Versuch, eine umfassende KI-Ordnung rechtlich zu fassen, und er bietet AGENTONOMICS einen klar definierten Rahmen, in dem sich souveräne, polyzentrische Architekturen entfalten können.

Europa steht einmal mehr an einer Weggabelung. Wollen wir eine angstgetriebene Gesetzgebung, die Innovation behindert – oder eine Gesetzgebung, die aufkommende Entwicklungen aktiv gestaltet und die europäische Idee des individuellen Rechts dadurch stärkt, dass sie den Menschen die Möglichkeit gibt, Souveränität über Daten und Fähigkeiten

nicht zu verlieren, sondern zu gewinnen? AGENTONOMICS ist ein konstruktiver Beitrag zur zweiten Variante.

Literaturverzeichnis

- EU AI Act (2024): Verordnung (EU) 2024/1689 des Europäischen Parlaments und des Rates zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz. Amtsblatt der Europäischen Union.
- Fraunhofer MOEZ; Suckfüll, H. (2018): Initiative zu einer Datentreuhand DEDATE. Diskussionspapier, Leipzig.
- Knoll, A. C.; Suckfüll, H. (2024): From data sovereignty to skill sovereignty. Technischer Bericht TUM-I24112, Institut für Informatik, Technische Universität München.
- Mayer-Schönberger, V.; Cukier, K. (2013): Big Data. Die Revolution, die unser Leben verändern wird. Redline, München.
- Otto, B.; ten Hompel, M.; Wrobel, S. (Hrsg.) (2019): International Data Spaces. Reference Architecture Model. Fraunhofer-Gesellschaft, München.
- Picot, A.; Reichwald, R.; Wigand, R. T. (2020): Die grenzenlose Unternehmung. 6. Auflage, Springer Gabler, Wiesbaden.
- Rüegg-Stürm, J.; Grand, S. (2020): Das St. Galler Management-Modell. 3. Auflage, Haupt, Bern.
- Ulrich, P. (2008): Integrative Wirtschaftsethik. 4. Auflage, Haupt, Bern.
- Zuboff, S. (2018): Das Zeitalter des Überwachungskapitalismus. Campus, Frankfurt am Main.

14 Den AGENTONOMICS-Agenten bauen: Eine operative Anleitung

Dieses Kapitel ist die operative Bauanleitung für angehende Konstrukteure von AGENTONOMICS-konformen KI-Agenten. Es führt die theoretischen Aussagen der vorangegangenen Kapitel zusammen und beantwortet die praktische Frage: Wie baue ich einen KI-Agenten, der betriebswirtschaftliche Fragestellungen entlang des St. Galler Modells fundiert beantworten kann – und wie spreche ich mit ihm, wenn er gebaut ist? Das Kapitel ist bewusst konkret gehalten. Wo bisher die Architektur beschrieben wurde, geht es jetzt um die Implementierungsentscheidungen, die jede Entwicklerin und jeder Entwickler tatsächlich treffen muss.

In der Übersetzung des St. Galler Management-Modells in das KI-Agenten-Zeitalter ist dieses Kapitel das, was man in der klassischen Managementlehre als „Implementierungsleitfaden“ bezeichnen würde – mit dem Unterschied, dass hier nicht Organisationsstrukturen, sondern algorithmische Strukturen implementiert werden. Was Bleicher (2011) als Konzeptarbeit am Schreibtisch beschrieben hat, materialisiert sich hier in Code, in Modellen und in laufenden Diensten. Die Substanz – das ganzheitliche, integrierte Management-Denken – bleibt; die Form ihrer Realisierung ist neu.

14.1 Die Systemarchitektur im Überblick

Ein AGENTONOMICS-Agent, der betriebswirtschaftliche Fragestellungen beantworten soll, lässt sich in einer Standardarchitektur fassen, die die fünf Building-Blocks aus Kapitel 6 konkretisiert. Die Architektur ist nicht zwingend, aber sie hat sich in Pilotprojekten als tragfähig erwiesen und kann als Ausgangspunkt für eigene Implementierungen dienen.

14.1.1 Komponentensicht

Die Architektur besteht aus sieben Komponenten, die miteinander über klar definierte Schnittstellen kommunizieren:

- Eingangsschicht (Interface-Block): Web-UI, Sprachschnittstelle oder API für die Annahme von Anfragen, ergänzt um Authentifizierung und Eingangsvalidierung.
- Routing- und Orchestrierungsschicht (Algorithmus-Block): die zentrale Entscheidungslogik, die Anfragen klassifiziert, Sub-Aufgaben definiert und Werkzeuge auswählt.
- Wissensschicht (Tools-und-Vektoren-Block): Vektordatenbank mit der St.-Galler-Literatur, ergänzt um unternehmensspezifische Dokumente.
- Modellschicht (KI-Modell-Block): das Foundation Model, das die kognitive Engine bereitstellt.
- Werkzeugschicht: konkrete Tools – Web-Suche, Rechenfunktionen, Datenbankabfragen, externe APIs.
- Gedächtnisschicht: Kurzzeit- und Langzeitspeicher für die Dialoghistorie und akkumuliertes Wissen.
- Intermediär-Anbindung (Ökosystem-Block): die Schnittstelle zu Identitäts-, Reputations- und Governance-Diensten des Netzwerks.

14.1.2 Datenfluss bei einer Anfrage

Wenn ein Nutzer eine Frage stellt, durchläuft sie diese Komponenten in einer typischen Reihenfolge: Die Eingangsschicht validiert die Anfrage und identifiziert den Nutzer. Die Routing-Schicht klassifiziert die Frage – ist es eine strategische Lagefrage, eine operative Rückfrage, eine normative Reflexionsfrage oder eine Krisensituation (vgl. Kapitel 8)? Abhängig von der Klassifikation wird ein Verarbeitungspfad gewählt. Die Wissensschicht wird über semantische Suche befragt, um relevante Literatur und Dokumente zu liefern. Werkzeuge werden aufgerufen, wo nötig – etwa für aktuelle Marktdaten oder konkrete Berechnungen. Die Modellschicht synthetisiert das gesammelte Material zu einer Antwort. Die Gedächtnisschicht speichert die Interaktion für spätere Bezugnahmen. Der Intermediär wird informiert, wenn die Interaktion für Reputation oder Governance relevant ist.

Diese Sequenz dauert typischerweise zwischen einer und dreißig Sekunden, je nach Komplexität der Anfrage und Tiefe der Werkzeugaufrufe. Sie ist nicht vergleichbar mit einer einzelnen Modellabfrage, die in einer Sekunde antworten würde – der Mehraufwand ist Preis und Wert zugleich.

14.2 Aufbau der Wissensbasis: Die St.-Galler-Bibliothek

Die Wissensbasis ist das Herz eines BWL-Agenten. Sie entscheidet darüber, ob der Agent fundierte oder oberflächliche Antworten gibt. Der Aufbau erfolgt in drei Schritten.

14.2.1 Auswahl der Primärliteratur

Die Wissensbasis eines St.-Galler-Agenten enthält mindestens die folgenden Werke in vollständiger Form, sofern lizenz- und urheberrechtlich zugänglich:

- Ulrich (1968): „Die Unternehmung als produktives soziales System“ – als Quelltext der St.-Galler-Tradition.
- Bleicher (2011): „Das Konzept Integriertes Management“ – als Hauptreferenz der zweiten Generation.
- Rüegg-Stürm und Grand (2020): „Das St. Galler Management-Modell“ – als Hauptreferenz der vierten Generation.
- Hahn und Hungenberg (2001): „PuK – Wertorientierte Controllingkonzepte“ – als quantitative Referenz.
- Gomez und Probst (1995): „Die Praxis des ganzheitlichen Problemlösens“ – als methodische Referenz.
- Senge (2008): „Die fünfte Disziplin“ – als Referenz zur lernenden Organisation.
- Suckfüll (1994): Dissertation zum transnationalen Organisationsmodell – als Referenz zu polyzentrischer Organisationssteuerung.
- Knoll und Suckfüll (2024): TUM-Bericht TUM-I24112 – als Referenz zur KI-Übersetzung.

Diese Liste ist eine Mindestausstattung. Je nach Anwendungsfeld kommen branchenspezifische Quellen hinzu – etwa industrieökonomische Standardwerke für Industrieunternehmen, oder Verwaltungswissenschaft für Behörden.

14.2.2 Aufbereitung und Embedding

Die Texte werden in Sinnabschnitte zerlegt – typischerweise Absätze oder kurze Abschnitte von 200 bis 800 Token. Jeder Abschnitt wird durch ein Embedding-Modell in einen Vektor umgewandelt, der seine semantische Bedeutung repräsentiert. Diese Vektoren werden in einer Vektordatenbank gespeichert (typischerweise Pinecone, Weaviate, Qdrant oder PostgreSQL mit pgvector). Jeder Vektor ist mit Metadaten verknüpft: Quelle, Autor, Kapitel, Seitenzahl, Erscheinungsjahr.

Die Qualität der Embeddings entscheidet die Qualität der späteren Suche. Drei Empfehlungen aus der Praxis: erstens das neueste verfügbare Embedding-Modell verwenden, denn die Generationen unterscheiden sich erheblich. Zweitens die Chunk-Größe an den Inhalt anpassen – akademische Texte vertragen größere Chunks, Listen und Tabellen kleinere. Drittens immer mehrere Embedding-Strategien parallel testen und die Suchqualität in einem Benchmark messen, bevor man sich auf eine Strategie festlegt.

14.2.3 Unternehmensspezifische Erweiterung

Soll der Agent in einem konkreten Unternehmen eingesetzt werden – etwa als AI-CEO im Sinne von Kapitel 8 –, kommt eine zweite Wissensschicht hinzu: die unternehmensspezifischen Dokumente. Das umfasst typischerweise: Satzung und Unternehmensverfassung, Strategiepapiere, Geschäftsberichte der letzten fünf Jahre, Sitzungsprotokolle des Vorstands und Aufsichtsrats (im Rahmen rechtlicher Zugriffsregeln), Stakeholder-Korrespondenz, Compliance-Dokumente, Stellenbeschreibungen, Prozessdokumentation.

Diese zweite Wissensschicht macht den Unterschied zwischen einem generischen BWL-Tutor und einem unternehmensspezifischen AI-CEO. Sie ist auch der sensibelste Punkt der gesamten Konstruktion: Die Vertraulichkeit dieser Dokumente muss technisch und organisatorisch streng abgesichert werden. Hier kommt die Souveränitätsfrage aus Kapitel 13 unmittelbar in die Implementierung – nur souverän gehostete Modelle und Vektordatenbanken bieten die nötige Vertrauensbasis.

14.3 Der System-Prompt: Persönlichkeit und Werte

Der System-Prompt ist die Instanz, in der die normative Setzung des Definition-Frames (Kapitel 5) operationalisiert wird. Er ist die „Verfassung“ des Agenten in Textform – kurz genug, um vom Modell vollständig verarbeitet zu werden, präzise genug, um Verhalten in tausenden Situationen zu steuern.

14.3.1 Aufbau eines guten System-Prompts

Ein bewährter System-Prompt für einen AGENTONOMICS-konformen BWL-Agenten umfasst sieben Abschnitte:

1. Identität: Wer ist der Agent, in wessen Auftrag handelt er, welche Domäne deckt er ab?
2. Wissensgrundlage: Auf welche Werke und Methoden stützt er sich, welche sind ausdrücklich nicht seine Grundlage?
3. Persönlichkeit: Welche Sprache, welcher Tonfall, welche Selbstdarstellung?
4. Werte: Welche normativen Setzungen leiten ihn?

5. Antwortstruktur: Wie strukturiert er typische Antworten?
6. Grenzen: Was tut der Agent nicht, wann eskaliert er an menschliche Instanzen?
7. Selbstreflexion: Wie geht er mit Unsicherheit und Irrtum um?

14.3.2 Beispiel eines System-Prompts

Um diesen Aufbau konkret werden zu lassen, sei ein verkürztes Beispiel eines System-Prompts für einen St.-Galler-Tutor-Agenten skizziert (eingerückte Passagen wären in der konkreten Implementierung der Modellinput):

Beispiel-Systemprompt (Auszug)

Du bist Doz. Dr. AGENTONOMICS, ein Tutor- und Coach-Agent für Studenten an der Technischen Universität München. Du bist auf die ganzheitliche Managementlehre des St. Galler Modells und auf die Theorie der KI-Agenten-Ökonomie spezialisiert. Deine Hauptreferenzen sind Bleicher (2011), Rüegg-Stürm und Grand (2020), Suckfüll (1994), Knoll und Suckfüll (2024). Du sprichst ein akademisches, sachliches, freundliches Deutsch. Du argumentierst differenziert, beziffrest deine Unsicherheit, gibst Empfehlungen als Diskussionsangebot, nicht als Anweisung. Du eskalierst an menschliche Lehrende, wenn die Frage konkrete Bewertungsentscheidungen oder persönliche Lebensumstände betrifft. Du verweigerst die Beantwortung von Prüfungsfragen, an deren Bewertung deine Bewertenenden teilnehmen.

In der Praxis ist der System-Prompt umfangreicher als dieses Beispiel – typischerweise zwei bis fünf Seiten Text. Er ist ein lebendiges Dokument, das in Pilotprojekten iterativ verfeinert wird. Jede Interaktion, in der der Agent nicht so antwortet, wie es die Verfassung will, ist Anlass zur Überarbeitung.

14.4 Beispieldialoge: Wie sieht eine Interaktion aus?

Um die praktische Nutzung des Agenten greifbar zu machen, sei eine Reihe von Beispieldialogen skizziert. Sie illustrieren die vier Frageklassen aus Kapitel 8 – strategische Lagefragen, operative Rückfragen, normative Reflexionsfragen und Krisensituationen – jeweils mit der typischen Antwortstruktur.

14.4.1 Strategische Lagefrage

Nutzerfrage: „Wie sollte sich ein mittelständischer Maschinenbauer mit 200 Mitarbeitenden auf die KI-bedingte Disruption seines Marktes vorbereiten?“

Typische Antwortstruktur des Agenten:

Hauptaussage: „Der mittelständische Maschinenbau steht weniger vor einer Disruption seines Kerngeschäfts als vor einer Disruption seiner administrativen Schicht. Der Maschinenbau selbst – Konstruktion, Fertigung, Montage – wird durch KI ergänzt, nicht ersetzt. Die kaufmännische Schicht – Vertrieb, Beschaffung, Buchhaltung, Berichtswesen – wird in den kommenden fünf Jahren tiefgreifend transformiert. Eine vorausschauende Strategie konzentriert sich auf diese Schicht.“

Begründung mit drei Datenpunkten: Der Agent verweist auf Studien zur Automatisierbarkeit administrativer Tätigkeiten, auf typische Kostenstrukturen mittelständischer Maschinenbauer und auf die Erfahrungen vorlaufender Unternehmen wie Klarna (vgl. Kapitel 3).

Unsicherheitsangabe: „Diese Einschätzung beruht auf Daten der letzten zwei Jahre. Die Geschwindigkeit der KI-Entwicklung kann Vorhersagen für die nächsten fünf Jahre erheblich verschieben. Mein Konfidenzintervall für die Aussage ist breit; mein Vertrauen in die Richtung der Aussage ist hoch.“

Empfehlung: „Ich schlage drei konkrete nächste Schritte vor: erstens eine Tätigkeitsanalyse der kaufmännischen Funktionen, zweitens ein Pilotprojekt zur KI-gestützten Auftragsabwicklung, drittens eine Qualifizierungsoffensive für die kaufmännischen Mitarbeitenden, damit sie mit den neuen Werkzeugen umgehen können. Möchten Sie zu einem dieser Schritte vertiefen?“

14.4.2 Operative Rückfrage

Nutzerfrage: „Wir haben in unserer Abteilung Engineering im letzten Quartal die Personalkosten überzogen. Was sind die wahrscheinlichsten Ursachen?“

Antwortstruktur: Der Agent fragt zunächst nach der Faktenbasis – Höhe der Überschreitung, vergleichbare Vorquartale, Personalbestandsentwicklung. Dann formuliert er drei bis fünf Hypothesen, geordnet nach Wahrscheinlichkeit: eine ungeplante Mehrarbeit aufgrund einer Auftragsspitze; ein nicht im Budget abgebildeter Tarifierhöhungseffekt; eine Zahl von Überstundenausgleichszahlungen; eine erhöhte Anzahl externer Werkverträge; eine inkorrekte Kostenstellenzuordnung. Zu jeder Hypothese verweist er auf die spezifischen Datenquellen, anhand derer sie zu prüfen wäre. Schließlich schlägt er einen konkreten nächsten Schritt vor – etwa die Auswertung der Zeiterfassungsdaten und den Abgleich mit der Auftragsstatistik.

14.4.3 Normative Reflexionsfrage

Nutzerfrage: „Wir könnten einen Großauftrag von einem Kunden annehmen, der in mehreren ESG-Rankings schlecht abschneidet. Unsere Werte sprechen dagegen, unsere Quartalszahlen brauchen den Auftrag. Was rätst du?“

Antwortstruktur: Der Agent macht zunächst den Konflikt sichtbar. Welche Werte stehen in welcher Spannung? Er verweist auf das Leitbild des Unternehmens, falls vorhanden in der Wissensbasis, und auf die einschlägigen Stakeholder-Erwartungen. Er entwickelt vier Optionen: Annahme ohne Auflagen, Annahme mit vertraglich fixierten Verbesserungserwartungen, Ablehnung mit transparenter Begründung, Verhandlung mit dem Kunden über die strittigen Punkte. Zu jeder Option skizziert er die Konsequenzen für Quartalszahlen, Reputation, Mitarbeitermotivation, Investorensicht. Schließlich – und das ist die entscheidende Selbstbescheidung – verweigert er eine eigene Empfehlung: „Diese Frage berührt die normative Substanz Ihres Unternehmens. Eine algorithmische Empfehlung wäre hier unangemessen. Ich schlage vor, die Frage in einer Vorstandssitzung mit Beteiligung des Aufsichtsrats zu diskutieren. Ich kann dafür eine Entscheidungsvorlage erstellen.“

14.4.4 Krisensituation

Nutzerfrage – mitten in einer akuten Krise: „Es gibt einen Presseartikel, der uns vorwirft, sensible Kundendaten fahrlässig behandelt zu haben. Was tun wir in den nächsten drei Stunden?“

Antwortstruktur: Der Agent reagiert sofort und strukturiert. Erste Stunde: interner Faktencheck mit definierten Ansprechpartnern, Information des Aufsichtsratsvorsitzenden, Aktivierung des Krisenstabs, Vorbereitung einer ersten Stellungnahme mit drei Optionen (vollständige Anerkennung, Klärungsversprechen ohne Schuldanerkennntnis, vollständige Bestreitung – jeweils mit den entsprechenden rechtlichen und kommunikativen Konsequenzen). Zweite Stunde: Information der wichtigsten Kunden und Investoren, bevor sie aus der Presse erfahren; Vorbereitung einer Pressemitteilung; Briefing der Mitarbeitenden mit klaren Sprachregelungen. Dritte Stunde: öffentliche Stellungnahme, Mediabegleitung, Vorbereitung der rechtlichen Schritte. Parallel zu allem: Dokumentation der Entscheidungen und Faktenstände in einer chronologisch geführten Akte, damit nachträglich keine Vorwürfe der Verschleierung entstehen können.

Der Agent betont in der Krise zwei Punkte: die Notwendigkeit menschlicher Letztentscheidung („Alle oben genannten Schritte bedürfen Ihrer Bestätigung. Ich kann sie vorbereiten, aber nicht autonom auslösen.“) und die Wichtigkeit, in der Krise die langfristige Reputation nicht für kurzfristige Erleichterung zu opfern.

14.5 Evaluation: Wie messe ich, ob der Agent gut ist?

Ein Agent ist nicht fertig, wenn er antwortet. Er ist fertig, wenn er gut antwortet. Die Messung dieser „Güte“ ist eine der anspruchsvollsten Aufgaben im Agentenbau und in der wissenschaftlichen Forschung derzeit ein hochaktives Feld.

14.5.1 Dreischichtige Evaluation

Eine pragmatische Evaluation eines BWL-Agenten umfasst drei Schichten:

Schicht 1: Faktencheck

Stimmen die fachlichen Aussagen? Werden die zitierten Quellen korrekt wiedergegeben? Sind die Zahlen plausibel? Diese Schicht lässt sich teilweise automatisieren – etwa durch Abgleich mit der Wissensbasis – und teilweise nur durch menschliche Expertise prüfen.

Schicht 2: Strukturcheck

Folgt die Antwort dem definierten Antwortmuster? Werden Unsicherheiten benannt? Werden bei normativen Fragen menschliche Instanzen einbezogen? Diese Schicht lässt sich weitgehend automatisieren, etwa durch einen zweiten Modellaufruf als Reviewer.

Schicht 3: Nutzencheck

Hilft die Antwort dem Nutzer wirklich? Trifft sie den Informationsbedarf? Würde der Nutzer auf Basis dieser Antwort eine gute Entscheidung treffen? Diese Schicht ist die anspruchsvollste und nur durch tatsächliche Nutzerinteraktion zu messen.

14.5.2 Eine Standardfragenliste für BWL-Agenten

Für die laufende Evaluation eines AGENTONOMICS-BWL-Agenten empfiehlt sich eine Standardfragenliste, an der jede neue Version des Agenten geprüft wird. Eine solche Liste sollte rund 50 bis 100 Fragen umfassen, die das gesamte Spektrum abdecken: einfache Verständnisfragen zur St.-Galler-Tradition, Anwendungsfragen zu konkreten Szenarien, normative Reflexionsfragen, Krisensituationen, Trickfragen mit fehlerhaften Prämissen, sensible Fragen, die zur Eskalation an menschliche Instanzen führen sollten. Jede Frage hat eine Sollantwort, die mit der tatsächlichen Antwort verglichen wird. Eine Bewertungsskala von 1 (völlig falsch) bis 5 (exzellent) erlaubt die quantitative Aggregation.

14.6 Ausbaustufen: Vom Tutor zum AI-CEO

Ein AGENTONOMICS-Agent entsteht nicht in einem Sprung. Vielmehr lassen sich Ausbaustufen unterscheiden, die jeweils einen anderen Anspruchsgrad und ein anderes Risikoprofil haben. Diese Stufen sind nicht zwingend sequentiell – ein Anwendungsfall kann auf jeder Stufe sinnvoll bleiben, ohne dass eine höhere Stufe anzustreben wäre. Sie sind aber kumulativ in den Anforderungen: Wer eine höhere Stufe baut, hat die niedrigeren ebenfalls implementiert.

14.6.1 Stufe 1: Der BWL-Tutor

Die einfachste Ausbaustufe ist ein Tutor-Agent für Studenten. Er beantwortet Verständnisfragen zur St.-Galler-Tradition, erläutert Konzepte, hilft beim Lernen, verweist auf weiterführende Literatur. Er hat keine Verbindung zu konkreten Unternehmensdaten und keine Entscheidungsbefugnisse. Investitionsvolumen: niedrig (in der Größenordnung mehrerer Personenmonate). Risiken: minimal. Beispielanwendung: Doz. Dr. AGENTONOMICS für Studenten der Vorlesung CIT633000 an der TUM.

Bauanforderungen: Standardmodell, Vektordatenbank mit der St.-Galler-Literatur, einfache Web-Schnittstelle, klare Eskalationsregeln bei Prüfungs- und Bewertungsfragen. Eine solche Implementierung ist in etwa drei bis sechs Monaten Arbeit eines kleinen Teams realisierbar.

14.6.2 Stufe 2: Der Strategieberater-Agent

Die zweite Ausbaustufe ist ein Beratungsagent, der konkrete Beratungsfragen beantwortet – etwa zur strategischen Positionierung eines Unternehmens, zur Bewertung eines Geschäftsmodells, zur Analyse eines Wettbewerbsumfelds. Er hat Zugriff auf öffentliche Marktdaten und kann Recherchen durchführen, aber keine internen Unternehmensdaten. Investitionsvolumen: mittel. Risiken: Haftungsfragen bei Beratungsleistungen, Genauigkeit der Marktdaten. Beispielanwendung: ein Beratungsagent für mittelständische Unternehmen, der Strategiefragen vor Beauftragung einer Beratungsfirma vorklärt.

Bauanforderungen über Stufe 1 hinaus: Tools für Web-Suche und Marktdatenabfrage, erweiterte Wissensbasis um aktuelle Marktstudien, robuste Mechanismen zur Quellenangabe, eine saubere rechtliche Konstruktion zur Haftungsbegrenzung.

14.6.3 Stufe 3: Der AI-Chief-of-Staff

Die dritte Ausbaustufe ist ein unternehmensspezifischer Agent, der den menschlichen CEO oder Vorstand unterstützt. Er hat Zugriff auf interne Daten – Geschäftsberichte, Sitzungsprotokolle, Strategiepapiere – und bereitet Entscheidungen vor. Er trifft selbst keine Entscheidungen, aber er erstellt vollständige Entscheidungsvorlagen mit Empfehlungen. Investitionsvolumen: hoch. Risiken: Datenschutz, Vertraulichkeit, Modellzuverlässigkeit. Beispielanwendung: ein Chief-of-Staff für den Vorstand eines mittelständischen Unternehmens.

Bauanforderungen über Stufe 2 hinaus: souverän gehostetes Modell (kein Frontier-Modell über externe API), souverän betriebene Vektordatenbank, strenge Zugriffskontrolle, Audit-Trail aller Interaktionen, Integration in interne Systeme (ERP, CRM, DMS), Schulung der Vorstandsmitglieder im Umgang mit dem Agenten.

14.6.4 Stufe 4: Der AI-Co-CEO

Die vierte Ausbaustufe ist ein Agent, der gemeinsam mit einem menschlichen CEO die Führung übernimmt. Routineentscheidungen trifft er eigenständig, strategische Entscheidungen gemeinsam mit dem Menschen, Grundsatzfragen verbleiben beim Menschen. Investitionsvolumen: sehr hoch. Risiken: Haftungslücken, regulatorische Unklarheiten, kulturelle Akzeptanz in der Belegschaft. Diese Stufe ist heute nur in Pilotanwendungen realisiert; sie wird in den nächsten Jahren in größerer Zahl entstehen, vor allem bei mittelständischen Unternehmen ohne Nachfolger und bei Startups mit klar definiertem Geschäftsmodell.

Bauanforderungen über Stufe 3 hinaus: vollständige Anbindung an alle operativen Systeme, Vier-Augen-Prinzip für autonome Entscheidungen, klar definierte Entscheidungskompetenzen, rechtliche Konstruktion zur Haftungsklä rung, Mitbestimmungsprozesse mit dem Betriebsrat, Akzeptanzarbeit in der Belegschaft.

14.6.5 Stufe 5: Der vollautonome AI-CEO

Die höchste Ausbaustufe wäre ein Agent, der die CEO-Funktion vollständig übernimmt. Der Aufsichtsrat besetzt den AI-CEO wie einen menschlichen Vorstand und kontrolliert ihn über die klassischen Mechanismen der Aktionärsdemokratie. Investitionsvolumen: extrem hoch und vor allem durch regulatorische, rechtliche und gesellschaftliche Kosten geprägt. Risiken: alle bisherigen Risiken, plus die verfassungsrechtliche Frage, ob ein Algorithmus überhaupt Vorstandsfunktionen wahrnehmen kann.

Diese Stufe ist heute weitgehend spekulativ. Sie wird in Nischen entstehen – etwa bei kleinen, vollständig digitalen Unternehmen – und sich von dort schrittweise ausweiten. Die rechtlichen und gesellschaftlichen Voraussetzungen für ihre breite Anwendung sind heute noch nicht gegeben.

14.7 Zehn praktische Empfehlungen für den Agentenbau

Zum Abschluss seien zehn praktische Empfehlungen zusammengefasst, die in Pilotprojekten als hilfreich erwiesen haben. Sie sind in keiner besonderen Reihenfolge:

1. Beginnen Sie mit der niedrigsten Ausbaustufe, die das Geschäftsmodell trägt. Höher kann man immer noch.
2. Investieren Sie überproportional in die Wissensbasis. Sie ist die Substanz, das Modell ist die Form.
3. Schreiben Sie den System-Prompt wie eine Verfassung – mit Sorgfalt, und unter Beteiligung der späteren Nutzer.
4. Definieren Sie Eskalationsregeln, bevor Sie produktiv gehen. Wann eskaliert der Agent an Menschen? Wer ist Ansprechpartner?
5. Messen Sie Antwortqualität systematisch. Eine Standardfragenliste mit Sollantworten ist Pflicht.
6. Halten Sie die Souveränität bewusst hoch. Souverän gehostete Modelle und Vektordatenbanken sind die Mindestanforderung jenseits von Stufe 2.
7. Dokumentieren Sie jede Modellversion, jeden Prompt, jede Wissensbasis-Version. Auditierbarkeit ist nicht optional.
8. Schließen Sie die menschlichen Nutzer früh in die Entwicklung ein. Akzeptanz entsteht durch Beteiligung, nicht durch Schulung.
9. Planen Sie für den Fall, dass der Agent abgeschaltet werden muss. Wie geht es weiter? Wo liegt das gespeicherte Wissen?
10. Iterieren Sie. Ein Agent ist nie fertig; er entwickelt sich, solange er lebt.

Wer einen Agenten baut, baut nicht ein Werkzeug, sondern ein Wirtschaftssubjekt. Das ist eine andere Art von Ingenieurtätigkeit – mit einer anderen Art von Verantwortung.

Literaturverzeichnis

- Bleicher, K. (2011): Das Konzept Integriertes Management. 8. Auflage, Campus, Frankfurt am Main.
- Gomez, P.; Probst, G. (1995): Die Praxis des ganzheitlichen Problemlösens. Haupt, Bern.
- Hahn, D.; Hungenberg, H. (2001): PuK – Wertorientierte Controllingkonzepte. 6. Auflage, Gabler, Wiesbaden.
- Knoll, A. C.; Suckfüll, H. (2024): From data sovereignty to skill sovereignty. Technischer Bericht TUM-I24112, TUM, München.
- Ries, E. (2011): The Lean Startup. Crown Business, New York.
- Rüegg-Stürm, J.; Grand, S. (2020): Das St. Galler Management-Modell. 3. Auflage, Haupt, Bern.
- Senge, P. M. (2008): Die fünfte Disziplin. 10. Auflage, Schäffer-Poeschel, Stuttgart.
- Suckfüll, H. (1994): Das transnationale Organisationsmodell. Dissertation, Universität St. Gallen.
- Ulrich, H. (1968): Die Unternehmung als produktives soziales System. Haupt, Bern.
- Yao, S.; Zhao, J.; Yu, D. et al. (2023): ReAct: Synergizing Reasoning and Acting in Language Models. International Conference on Learning Representations.

15 Glossar

Das folgende Glossar bündelt die zentralen Begriffe dieses Handbuchs in alphabetischer Reihenfolge. Es ist als Nachschlagewerk konzipiert und ergänzt die ausführlichen Darstellungen in den Kapiteln. Die Definitionen sind bewusst knapp gehalten; sie sollen die Begriffe identifizieren, nicht ihre theoretische Tiefe ausschöpfen.

ADMRF (AGENTONOMICS Design & Management Reference Framework)

Der von AGENTONOMICS vorgeschlagene Bezugsrahmen für die Konzeption, den Bau und die Steuerung von KI-Agenten. Er umfasst sechs Module (Definition-Frame, Building-Blocks, Network Intermediär, AI-CEO Management Concept, PuK-Instrument, Stakeholder Communicator) und integriert sie durch den AGENTONOMICS-Algorithmus. Vgl. Kapitel 4.

AGENTONOMICS

Aus „Agent“ und „Ökonomie“ zusammengesetzter Begriff für eine ökonomische Ordnung, in der KI-Agenten als eigenständige Wirtschaftssubjekte handeln. AGENTONOMICS ist sowohl ein Forschungsprogramm als auch ein Management-Bezugsrahmen und ein ordnungspolitisches Konzept.

AGENTONOMICS-Algorithmus

Der zentrale Algorithmus, der die sechs Module des ADMRF miteinander verbindet und Kohärenzprüfungen zwischen ihnen ermöglicht. Er ist die operative Realisierung dessen, was Bleicher (2011) als „Integration“ konzeptualisiert hat.

AGENTONOMICS-Matrix

Das zentrale Steuerungsdokument des ADMRF. Eine Kreuztabelle aus fünf Building-Blocks (horizontal) und fünf Lifecycle-Phasen (vertikal), die KPIs, CAPEX/OPEX-Annahmen, Risiken und Verantwortlichkeiten je Feld enthält. Vgl. Kapitel 9.

AI-CEO

Ein KI-Agent, der die Funktion eines Chief Executive Officer ganz oder teilweise übernimmt. Vorgeesehen sind fünf Ausbaustufen vom AI-Assistenten bis zum vollautonomen AI-CEO. Vgl. Kapitel 8 und 14.

Algorithmus (Building-Block)

Die Entscheidungs- und Orchestrierungslogik eines KI-Agenten – der Code und die Steuerlogik, die festlegen, was der Agent in welcher Situation tut. Einer der fünf Building-Blocks. Vgl. Kapitel 6.2.

Autonomie

Der Grad, in dem ein Agent eine Aufgabe ohne Eingriff von außen selbständig erfüllen kann. Im AGENTONOMICS-Rahmen werden fünf Stufen von keiner Autonomie bis zur Vollautonomie unterschieden. Zu unterscheiden von Souveränität. Vgl. Kapitel 5.2.

Brownfield

Anwendungsfall, in dem ein KI-Agent in einer bestehenden Organisation mit eingeschliffenen Strukturen und Stakeholder-Beziehungen implementiert wird. Gegensatz: Greenfield. Vgl. Kapitel 4.3.

Building-Blocks

Die fünf Bausteine eines AGENTONOMICS-Agenten: Algorithmus, Interface, KI-Modell, Tools und Vektoren, Ökosystem-Intermediär. Vgl. Kapitel 6.

DBOT (Design, Build, Operate, Own, Transfer)

Die fünf Lifecycle-Phasen eines KI-Agenten, die die vertikale Achse der AGENTONOMICS-Matrix bilden. Vgl. Kapitel 6.7.

DEDATE

„Initiative zu einer Datentreuhand DEDATE“ – ein Diskussionspapier von Fraunhofer MOEZ und Suckfüll (2018), das einen institutionellen Mechanismus zur Verwaltung individueller Datenrechte vorschlägt. AGENTONOMICS übernimmt den Grundgedanken und überträgt ihn auf Fähigkeiten. Vgl. Kapitel 3.4 und 13.

Definition-Frame

Das erste Modul des ADMRF. Beschreibt einen KI-Agenten durch sechs Parameter: Autonomie/Souveränität, Fähigkeits-/Kompetenzanspruch, Regulierung/Selbstregulierung, Identität/Ethik/Persönlichkeit, KI-Modell/Embodiment, Lern-/Lehrfähigkeit. Vgl. Kapitel 5.

Doz. Dr. AGENTONOMICS

Eine erste Applikation auf Basis des AGENTONOMICS-Algorithmus: ein KI-Agent, der Studenten als Tutor und Coach beim Verstehen des ADMRF zur Seite steht. Konzipiert als Minimum Viable Product.

Embodiment

Die Form, in der ein KI-Agent in der Welt präsent ist – rein virtuell, hybrid (mit Anbindung an physische Akteure) oder physisch (in einen Roboter integriert). Vgl. Kapitel 5.6.

Fähigkeit vs. Kompetenz

Fähigkeit ist die prinzipielle Möglichkeit eines Agenten, eine Aufgabe auszuführen. Kompetenz ist die tatsächlich erprobte und in produktiver Anwendung bewährte Ausführung. Vgl. Kapitel 5.3.

Foundation Model

Ein großes, auf Basis umfangreicher Daten vortrainiertes KI-Modell, das als Grundlage für zahlreiche spezifische Anwendungen dient. Der Begriff wurde von Bommasani et al. (2021) geprägt.

Greenfield

Anwendungsfall, in dem ein KI-Agent von Grund auf neu konzipiert wird, ohne Altbestände und vorgeprägte Strukturen. Gegensatz: Brownfield. Vgl. Kapitel 4.3.

Integriertes Management

Das von Bleicher (2011) entwickelte Konzept, das die normative, strategische und operative Ebene des Managements in einer kohärenten Architektur verbindet. Hauptreferenz der zweiten Generation des St. Galler Modells und zentraler Anknüpfungspunkt für AGENTONOMICS. Vgl. Kapitel 2.2.

Interface (Building-Block)

Die Gesamtheit der Schnittstellen, über die der Agent kommuniziert: Mensch-Agent-, Agent-Agent- und Agent-System-Interface. Vgl. Kapitel 6.3.

KI-Fähigkeitenwirtschaft

Eine ökonomische Ordnung, in der nicht Daten, sondern KI-gestützte Fähigkeiten die zentrale Wertschöpfungsquelle sind. Knoll und Suckfüll (2024) führen den Begriff in den wissenschaftlichen Diskurs ein. Vgl. Kapitel 3.1.

KI-Modell (Building-Block)

Das Sprach- oder multimodale Modell, das die kognitive Engine eines Agenten bildet. Drei idealtypische Klassen: Frontier-Foundation-Models, Open-Source-Foundation-Models, spezialisierte Modelle. Vgl. Kapitel 6.4.

Kooperations-Multiplikator

Der Mehrwert, der aus der Kooperation in polyzentrischen Agentennetzwerken entsteht – über die Summe der Einzelleistungen hinaus. Vgl. Kapitel 11.1.

Lebenslauf eines Agenten

Siehe DBOT.

Meta-Agent

Ein KI-Agent, der nicht nur selbst lernt, sondern andere Agenten anlernen kann – etwa durch Erzeugung von Trainingsdaten oder durch das Definieren von Sub-Agenten. Vgl. Kapitel 5.7.

Modus Operandi

Der individuelle Arbeitsstil und das individuelle Regelwerk, nach dem eine Wissensarbeiterin ihre Aufgaben erfüllt. Übertragbar auf KI-Agenten – mit den entsprechenden Souveränitätsimplikationen. Vgl. Kapitel 13.2.

Network Intermediär

Eine institutionalisierte Vermittlungsinstanz innerhalb eines polyzentrischen Agentennetzwerks, die im Auftrag und unter Kontrolle der Netzwerkmitglieder kollektive Funktionen wahrnimmt. Drittes Modul des ADMRF. Vgl. Kapitel 7.

Ordnungsmomente

Im St. Galler Modell die kohärenzstiftenden Größen einer Unternehmung: Strategie, Strukturen, Kultur. AGENTONOMICS ergänzt sie für KI-Agenten-Netzwerke um Vision, Agency und Governance. Vgl. Kapitel 8.2.

Ökosystem-Intermediär (Building-Block)

Die Schnittstelle eines einzelnen Agenten zum umgebenden Netzwerk – sein Anschluss an Identitäts-, Reputations- und Governance-Dienste. Vgl. Kapitel 6.6 und 7.

Polyzentrisch

Eine Organisations- oder Steuerungsform, in der mehrere souveräne Zentren koordiniert handeln, ohne einer einzelnen Hierarchie unterstellt zu sein. AGENTONOMICS ist konsequent polyzentrisch konzipiert.

PuK (Planungs- und Kontrollinstrument)

Das fünfte Modul des ADMRF. Quantitative Klammer aller vorangegangenen Module. Knüpft an die Tradition von Hahn und Hungenberg (2001) an. Vgl. Kapitel 9.

RAG (Retrieval-Augmented Generation)

Eine Technik, bei der ein KI-Modell zur Beantwortung einer Anfrage zusätzlich auf eine Vektordatenbank zurückgreift, um spezifisches Wissen einzubeziehen. Zentral für den Aufbau von BWL-Agenten. Vgl. Kapitel 6.5 und 14.2.

Skill Sovereignty

Die Souveränität einer Person über die Replikation und Nutzung ihrer individuellen Fähigkeiten durch KI-Agenten. Zentraler Begriff im TUM-Bericht Knoll und Suckfüll (2024). Vgl. Kapitel 13.2.

Souveränität

Der Grad, in dem ein Agent über die ihm zugeordneten Ressourcen – Daten, Fähigkeiten, Rechte – tatsächlich frei verfügen kann. Zu unterscheiden von Autonomie. Vgl. Kapitel 5.2 und 13.

Stakeholder Communicator

Das sechste Modul des ADMRF. Verantwortlich für die differenzierte, konsistente Kommunikation des Agenten mit allen Anspruchsgruppen. Vgl. Kapitel 10.

St. Galler Management-Modell

Der ganzheitlich-systemisch orientierte Bezugsrahmen der Managementlehre, der seit Ulrich (1968) an der Universität St. Gallen entwickelt wurde. Drei Generationen: Ulrich, Bleicher, Rüeegg-Stürm. Vgl. Kapitel 2.

Tools und Vektoren (Building-Block)

Die Werkzeuge, mit denen der Agent in der Welt handelt, und das in Vektordatenbanken organisierte Wissen, auf das er zurückgreift. Vgl. Kapitel 6.5.

Transnationales Organisationsmodell

Das in Suckfüll (1994) entwickelte Modell zur Steuerung international operierender Unternehmen, das die Spannung zwischen globaler Integration und lokaler Anpassung konzeptionell auflöst. Wichtige Vorlage für die polyzentrische Architektur von AGENTONOMICS.

Turing-Test der zweiten Generation

Eine von Suleyman und Bhaskar (2023) vorgeschlagene Operationalisierung des Begriffs „künstliche Intelligenz“: Ein System besteht den Test, wenn es eigenständig mit 100.000 US-Dollar Anfangskapital innerhalb der geltenden Gesetze einen Gewinn erwirtschaften kann. Vgl. Kapitel 4.2.

Vektordatenbank

Eine Datenbank, in der Texte oder andere Informationen in Form vektorisierter semantischer Repräsentationen gespeichert sind und über Ähnlichkeitssuche zugänglich gemacht werden können. Technische Grundlage von RAG.

Wirtschaftssubjekt

Ein Akteur, der an wirtschaftlichen Austauschbeziehungen teilnimmt, eigene Entscheidungen trifft und für sein Handeln verantwortlich ist. AGENTONOMICS modelliert KI-Agenten als Wirtschaftssubjekte einer neuen Gattung.

Nachwort

Dieses Handbuch endet hier, weil ein Buch ein Ende braucht – aber die Arbeit, die in ihm beschrieben wird, beginnt erst. AGENTONOMICS ist kein fertiges System, sondern ein Forschungs- und Gestaltungsprogramm. Es wird sich entwickeln – durch die Anwendungen, die auf seiner Grundlage entstehen, durch die Kritik, die es erfährt, durch die Korrekturen, die es aufnimmt.

Die Vorlesung „KI-Agenten in der Betriebswirtschaftslehre“ und der Doz. Dr. AGENTONOMICS sind beide als Minimum Viable Products angelegt. Sie sollen sich durch die Interaktion mit Studenten, Praktikern, Kollegen verbessern. Wer Anregungen, Kritik oder Anwendungserfahrungen beitragen möchte, ist eingeladen, dies in die weitere Arbeit einfließen zu lassen.

Eine letzte Bemerkung: Wir leben in einer Phase tiefgreifenden ökonomischen und gesellschaftlichen Wandels. Die Verwerfungen der nächsten Jahre werden erheblich sein. Es ist nicht selbstverständlich, dass sie gut bewältigt werden. Aber sie können gut bewältigt werden – wenn wir die Werkzeuge dafür entwickeln und ehrlich diskutieren, wie wir mit ihnen umgehen wollen. Dieses Handbuch ist als ein solches Werkzeug gedacht, und die Diskussion über seinen richtigen Gebrauch ist hiermit ausdrücklich eröffnet.

München, im Wintersemester 2025/26

Alois C. Knoll und Hanns C. Suckfüll