AGENTONOMICS

The AI Agent Economy

A Handbook on the Development and Control
economically operating AI agents

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

Chair of Robotics, Artificial Intelligence and Real-Time Systems
TUM School of Computation, Information and Technology
Technical University of Munich

1st edition 2026 · Winter semester 2025/26

Authors' Foreword

This handbook accompanies the lecture 'AI Agents in Business Administration – Agentic AI in Business Administration', which is being offered for the first time at the Technical University of Munich in the winter semester 2025/26. It brings together two traditions of knowledge that are rarely addressed jointly in business administration teaching: holistic, systemic management theory, as shaped in particular by the St. Gallen Management Model, and the computer science of autonomous, adaptive software agents.

The reason for this combination is not an academic end in itself. Economic reality is currently changing at a pace that overwhelms the classical repertoire of management theory. Generative language models and tool-capable agents are encroaching on fields of activity that were previously considered genuinely human. Knowledge work, decision preparation, negotiation, customer communication – everywhere, systems are emerging that are no longer merely tools, but act as agents in their own right. This shifts the very focus of business administration as a discipline.

With AGENTONOMICS, we propose a framework that consistently conceives of AI agents as economic entities and embeds them within a management architecture that historically stems from the St. Gallen concept of integrated management (Bleicher 2011) and is extended to meet the specific requirements of polycentric agent networks (Suckfüll 1994). The underlying algorithm – the AGENTONOMICS algorithm – and its first application, the 'Doz. Dr. AGENTONOMICS', are designed as a minimum viable product: they are intended to improve iteratively through concrete use cases.

The handbook has a twofold aim. Firstly, it is intended to provide students with a precise set of conceptual tools rooted in established management literature. Secondly, it aims to enable readers, by the end of the book, to design and technically implement AI agents themselves that can provide well-founded answers to business management questions in line with the St. Gallen Model. For this reason, the handbook concludes with a detailed, practical guide to agent development (Chapter 14).

We would like to thank the students and colleagues who contributed to refining the arguments with their critical questions, as well as the project partners from industry and academia, whose practical insights have shaped the case studies in this handbook.

Munich, in the winter semester 2025/26

Alois C. Knoll and Hanns C. Suckfüll

Contents

1 AI Agents in Business Administration – Introduction and Motivation

Whilst business administration research and teaching have, for decades, primarily focused on organisations, markets and human actors, the emergence of autonomous AI systems has brought new actors onto the scene: AI agents that no longer merely assist, but increasingly act autonomously, adaptively and purposefully. This shift calls for a fundamental expansion of the conceptual framework of business administration – not its complete replacement. The lecture 'AI Agents in Business Administration' and this handbook address precisely this issue.

It is evident that, whilst classical management models (Ulrich 1968; Bleicher 2011; Rüegg-Stürm and Grand 2020) continue to provide a sound foundation, they lose their explanatory power as soon as autonomous, adaptive and economically acting AI systems are integrated into value creation processes. The central question addressed by this handbook is therefore:

Key question

How can AI agents be conceptualised, designed and managed from a business perspective when they act as independently operating economic entities – and what consequences does this have for the architecture of holistic management models?

This guiding question forms the starting point for the AGENTONOMICS management model, which was developed as part of the lecture series and is set out step by step below. The name is derived from the terms 'agent' and 'economy' and reflects the aim of understanding AI agents not as isolated technical artefacts, but as components of an autonomous economic order.

AGENTONOMICS claims to be a state-of-the-art management model which, with the help of the AGENTONOMICS algorithm, utilises AI technology to provide human decision-makers with management expertise that is superior in both depth and speed. As a first step, AGENTONOMICS is being introduced in the form of a lecture to TUM students who are preparing for working life, whether as employees or as self-employed professionals.

1.1 Three objectives of the lecture

The lecture pursues three interlinked objectives, which are reflected in the structure of the handbook:

Terminological and conceptual orientation

Students should be enabled to understand AI agents not merely as technical artefacts, but as economic actors. This includes clearly situating AGENTONOMICS within the context of business administration, economics and related academic disciplines – in particular business informatics, business ethics and law (cf. Picot et al. 2020; Mainzer 2019).

Application and design skills

Building on a holistic, business management control model – the AGENTONOMICS Design & Management Reference Framework (ADMRF) – students learn to systematically describe and parameterise AI agents, and to conceive, technically implement and successfully manage them economically as integral components of business models or organisational forms. Chapter 4 sets out the overall architecture of the ADMRF; Chapters 5 to 10 explore its six modules in greater depth.

Reflections on future economic structures

The course is explicitly aimed at students preparing to enter a world of work and business shaped by AI – whether as employees, founders or as architects of new organisational and market forms. The concept of the 'AI skills economy', as first introduced into academic discourse by Knoll and Suckfüll (2024), is explored in depth throughout the handbook and formulated as a microeconomic model.

The overarching aim of the lecture is to bridge the gap between conventional management methodologies and innovative AI development opportunities. In doing so, the lecture is not intended merely as a means of imparting knowledge, but also as a space for reflection and experimentation.

1.2 Target audience and teaching approach

The lecture is deliberately aimed at an interdisciplinary audience. It is particularly aimed at students

- computer science,
- business administration,
- economics,
- and related engineering and social science degree programmes.

The teaching approach employs a dual shift in perspective to promote holistic thinking. For students with a primarily technical background, the lecture provides an introduction to business management ways of thinking, decision-making logic and control mechanisms. For students with a business management background, AI is not treated as a 'black box', but is presented as a configurable, parametric and economically effective system.

In the medium term, the lecture also positions itself as a training tool for senior management in companies, as well as for other socially relevant stakeholder groups wishing to contribute to the successful transformation towards an AI-driven society – the AI skills economy. In the long term, the lecture serves to further develop the AGENTONOMICS algorithm as a tool for practical business management and to gradually refine it using real-world examples.

1.3 Normative meta-skills in an AI-driven world of work

The lecture is based on the assessment that, in order to thrive in an AI-driven world of work, it is not primarily specialised technical knowledge that is crucial, but rather overarching, normative meta-skills. Drawing on educational theory dating back to Humboldt and engaging with contemporary analyses of the knowledge society (Nida-Rümelin and Weidenfeld 2018), these can be summarised in four clusters:

Creativity, reflected in entrepreneurial action

This refers to the ability to conceive of possibilities beyond the existing order and to translate them into viable courses of action. This creativity is not the romantic inspiration of the artist, but rather the pragmatic ability to combine existing resources in an entrepreneurial manner to create new solutions (Schumpeter 1934).

Resilience and a positive worldview

In a world of rapid technological disruption, the ability to deal productively with uncertainty is a key economic skill. Resilience is understood here not as an individual trait, but as a competence that can be cultivated through specific practices – such as scenario thinking, learning organisations and controlled experiments (Senge 2008).

Fundamental Ethical and Moral Orientation

This lecture regards the ethical orientation towards a Kantian imperative not as a secondary, rhetorical exercise, but as an integral part of economic judgement. An economy that operates with agents acting on behalf of their principals is dependent on stable normative expectations – otherwise its transaction costs will collapse (Ulrich 2008; Kant 1785).

Comprehensive general knowledge and world knowledge

What used to be simply called general knowledge has once again become a strategic resource in a world characterised by specialisation and AI-supported in-depth research. Anyone wishing to assess the implications of an AI output must be able to think within a broad framework of historical, philosophical, scientific and economic contexts.

Against this backdrop, the lecture will incorporate situational, topical examples from practice and multidisciplinary fields of knowledge – extending beyond business administration in the narrower sense – as well as current political events.

1.4 The AGENTONOMICS algorithm as a didactic experiment

A key feature of the lecture is its experimental nature. Artificial intelligence is not merely discussed, but actively integrated into the teaching and learning process. In the future, students will be supported by a customisable AI agent – 'Doz. Dr. AGENTONOMICS'. It acts as a tutor, coach and sparring partner, whilst also forming the basis for networking amongst the students.

'Doz. Dr. AGENTONOMICS' explains the six modules of the lecture within a holistic, cyclical process: the AGENTONOMICS Design & Management Reference Framework (ADMRF). The agent is to be understood as the first application based on the AGENTONOMICS algorithm. The application enables students and practitioners alike to conceive, design and implement AI agents in a systematic and business-oriented manner, and to manage them successfully within the AI skills economy.

The ADMRF comprises six modules that integrate different perspectives on corporate management – those of the CEO, the CFO, the CTO, the CSO and the board as a whole. A central AI model links these

perspectives in such a way as to provide state-of-the-art decision support for the development and management of AI agents. The AI-supported ADMRF is explicitly intended as a further development of traditional management frameworks. Historically, it can be seen as a logical continuation of the St. Gallen concept of integrated management (Bleicher 2011) and the PuK planning and control accounting model (Hahn and Hungenberg 2001). Chapter 2 explains these connections in detail.

In particular, the three-part division into a normative, a strategic and an operational perspective from the concept of integrated management is incorporated into the context of AI-agent network organisations and is reflected in the six modules of the ADMRF. Specifically, the control mechanisms of polycentric – that is, autonomously interacting AI-agent networks – are based to a large extent on the organisational understanding of the transnational organisational model (Suckfüll 1994), which was developed in the context of internationally operating companies. The design of the intermediary function between the AI agents and external stakeholders also benefits from this. The discussion on the design of an intermediary function enabling data sharing – which was largely initiated in Germany by the publication *Deutsche Daten-Treuhand* (DEDATE; Fraunhofer MOEZ and Suckfüll 2018) – is incorporated into the AGENTONOMICS algorithm.

1.5 Structure of the handbook

The handbook is structured in such a way that, by the end, readers will not only have understood the AGENTONOMICS model, but will also be able to build AI agents themselves that can provide well-founded answers to business management questions in line with the St. Gallen Management Model. This dual function – as both a textbook and a guide to building such agents – is reflected in its structure:

- Chapters 1 to 3 provide the theoretical foundation for AGENTONOMICS. They introduce the project, reconstruct the St. Gallen Management Model through its key developmental stages, and situate the emergence of AGENTONOMICS within a historical and economic context.
- Chapters 4 to 10 set out the AGENTONOMICS Design & Management Reference Framework (ADMRF) and its six modules – from the Definition Frame, through the five Building Blocks and the Network Intermediary, to the CEO Concept, the Planning and Control Account, and Stakeholder Management.
- Chapters 11 to 13 are devoted to growth theory, economic approaches and the question of sovereignty – in particular the transition from data sovereignty to capability sovereignty (Knoll and Suckfüll 2024).
- Finally, Chapter 14 provides the operational guide: How do I build an AGENTONOMICS-compliant AI agent that answers St. Gallen questions? The glossary in Chapter 15 brings together the most important terms.

Each chapter concludes with its own bibliography, to which references are made in the text. This ensures that the individual chapters can also be read in isolation.

Bibliography

Bleicher, K. (2011): The Concept of Integrated Management. Visions – Missions – Programmes. 8th, updated and expanded edition, Campus, Frankfurt am Main.

Fraunhofer MOEZ; Suckfüll, H. (2018): Initiative for a Data Trust: DEDATE. Discussion paper, Fraunhofer Centre for International Management and Knowledge Economy, Leipzig.

Hahn, D.; Hungenberg, H. (2001): PuK – Value-Oriented Controlling Concepts. 6th, revised and expanded edition, Gabler, Wiesbaden.

Kant, I. (1785): Groundwork of the Metaphysics of Morals. Riga.

Knoll, A. C.; Suckfüll, H. (2024): From data sovereignty to skill sovereignty. Technical Report TUM-I24112, Institute of Computer Science, Technical University of Munich.

Mainzer, K. (2019): Artificial Intelligence – When Will the Machines Take Over? 2nd edition, Springer, Berlin and Heidelberg.

Nida-Rümelin, J.; Weidenfeld, N. (2018): Digital Humanism. An Ethics for the Age of Artificial Intelligence. Piper, Munich.

Picot, A.; Reichwald, R.; Wigand, R. T. (2020): The Boundless Enterprise. 6th edition, Springer Gabler, Wiesbaden.

Rüegg-Stürm, J.; Grand, S. (2020): The St. Gallen Management Model. Management in a Complex World. 3rd edition, Haupt, Bern.

Schumpeter, J. A. (1934): The Theory of Economic Development. Harvard University Press, Cambridge, MA.

Senge, P. M. (2008): The Fifth Discipline. The Art and Practice of the Learning Organisation. 10th edition, Schäffer-Poeschel, Stuttgart.

Suckfüll, H. (1994): The Transnational Organisational Model. PhD thesis, University of St Gallen.

Ulrich, H. (1968): The Enterprise as a Productive Social System. Foundations of General Business Theory. Haupt, Bern.

Ulrich, P. (2008): Integrative Business Ethics. Foundations of an Economy that Serves Life. 4th edition, Haupt, Bern.

2 The St. Gallen Management Model as a Theoretical Foundation

Anyone wishing to understand AGENTONOMICS must first be familiar with the tradition within which it is situated. This chapter reconstructs the St. Gallen Management Model in its basic features and across its three generations. It serves as the central frame of reference against which the AGENTONOMICS framework is oriented and from which it – justifiably – distinguishes itself in certain respects. The chapter thus also provides the conceptual framework with which the AI agents described in later chapters are intended to operate: anyone wishing to build an agent that answers business management questions in line with the St. Gallen Model must have a precise grasp of this model.

2.1 Roots: Hans Ulrich's system-oriented theory of business

The St. Gallen Management Model did not originate as a marketing-oriented consultancy product, but as a scientific programme. It is closely associated with the name of Hans Ulrich, who, in his 1968 work *The Enterprise as a Productive Social System*, situated the enterprise within the theoretical tradition of systems orientation (Ulrich 1968). In doing so, Ulrich draws on General Systems Theory (Bertalanffy 1968) and cybernetics (Wiener 1948; Beer 1972) and applies their fundamental principles to the organisation.

Three basic assumptions characterise this perspective and run through the model right up to its most recent iteration:

The organisation as an open, purpose-oriented system

Organisations are involved in a wide range of interactive relationships with their environment. They cannot be reduced to a causal, machine-like logic, but are dependent on the creation of meaning and purpose. Their controllability is limited: unlike a technical system, a social system cannot be unambiguously 'programmed'.

Complexity as a fundamental experience of management

In the St. Gallen understanding, management is not about solving well-defined problems, but about dealing with complexity. The classic question is therefore not 'What is the right solution?', but 'How can we remain capable of acting, even though we do not fully understand the world?'. To this day, this fundamental attitude sharply distinguishes St. Gallen thinking from a purely instrumental understanding of management theory.

Shaping, steering and development as fundamental functions

Ulrich distinguishes three basic functions of management: shaping the institutional framework, steering day-to-day processes, and developing the system over time. To this day, this triad forms an implicit axis running through all generations of the model.

Definition: Management

In the St. Gallen understanding, management is the act of shaping, steering and developing purpose-oriented social systems, thereby ensuring their viability and performance under conditions of high complexity and uncertainty.

2.2 The second generation: Bleicher's concept of integrated management

In the 1990s and early 2000s, Knut Bleicher transformed the St. Gallen legacy into the prominent and internationally recognised 'Concept of Integrated Management' (Bleicher 2011). Bleicher's framework is the theoretical foundation on which AGENTONOMICS is most strongly based. It is therefore discussed in greater detail here.

2.2.1 The three levels: normative, strategic, operational

At its core lies the distinction between three levels of management, which do not stand one above the other like storeys, but rather interlock like concentric circles:

Normative management

Normative management addresses the fundamental questions of purpose within an organisation. Why do we exist? What values guide us? What responsibilities do we assume towards our stakeholders? The outcomes are the corporate constitution, corporate policy and corporate culture. The normative level establishes the rules of the game within which strategic and operational decisions are made.

Strategic Management

Strategic management translates the normative framework into longer-term programmes. It concerns positioning, competitive advantages, business areas, organisational structures and management systems. In the St. Gallen understanding, strategy is not primarily a problem of optimisation, but a question of pattern recognition in complex environments and the conscious choice between alternative development paths (Gomez and Probst 1995).

Operational Management

Operational management ensures efficient and effective implementation in day-to-day business. This encompasses the traditional areas of orders, processes, cost accounting, cash flow management and staff management. Important: the operational level is also embedded within the normative and strategic frameworks; it is not merely a value-neutral execution.

2.2.2 The three aspects: activities, structures, behaviour

Across the three levels, Bleicher distinguishes three aspects in which management becomes effective:

- Activities – that is, what the organisation does: programmes, tasks, projects, measures.

- Structures – the institutional frameworks within which action is taken: organisational structure, management systems, constitution.
- Behaviour – the cultural patterns by which actors are effectively guided: values, norms, communication styles, leadership behaviour.

The famous 'Bleicher matrix' is formed by intersecting the three levels with the three aspects. It yields nine fields in which management decisions can be situated – ranging from the corporate constitution (normative/structures) through strategic programmes (strategic/activities) to staff management in the narrower sense (operational/behaviour).

2.2.3 The concept of integration

The central concept in Bleicher's work is integration. This does not refer to bureaucratic coordination between departments, but rather to the internally consistent interlinking of all levels and aspects. A strategy that is not embedded in the corporate culture is doomed to fail; an operational measure that contradicts the vision undermines management. Integrated management therefore means that all nine fields of the matrix are coherently interrelated.

This demand for coherence is precisely where AGENTONOMICS comes in. Whereas Bleicher's model cognitively demands that top management achieve coherence, AGENTONOMICS offers an algorithm capable of technically verifying and supporting this coherence. This is not a substitute for, but rather a reinforcement of, Bleicher's approach.

2.3 The fourth generation: Rüegg-Stürm and Grand

With the 'new' version first published in 2002 and the fourth generation presented in 2017/2020, Johannes Rüegg-Stürm has brought the St. Gallen model up to date with a network-organised, communication-intensive economy (Rüegg-Stürm and Grand 2020). This fourth generation shapes the current teaching of the model and is also directly relevant to AGENTONOMICS.

2.3.1 Six basic categories

Rüegg-Stürm and Grand structure the model around six basic categories, which together form the conceptual framework:

Environmental spheres

Companies are embedded in a society, a natural environment, a technological landscape and an economy. These environmental spheres are not a context-free backdrop, but active participants. They provide resources and expectations and continuously scrutinise the company's actions for their legitimacy.

Stakeholders

A one-sided shareholder focus is replaced by a plurality of stakeholder groups: customers, employees, investors, suppliers, competitors, the state, the public and NGOs. Successful management is the ability to productively coordinate their sometimes conflicting expectations.

Interaction issues

Issues arise between the company and its stakeholders that can ignite tensions in their relationships: prices, quality, risks, values, concerns. The list of these issues is not static but constantly changing. Anyone leading a company must be able to continuously identify which issues are currently salient.

Organisational Pillars

The organising principles are the very centre of gravity of the new St. Gallen Model – and the crucial point of connection for AGENTONOMICS. They encompass strategy, structures and culture, and form the framework within which day-to-day value creation takes place. AGENTONOMICS translates this concept into five key organising principles for AI agents: vision, strategy, structure, agency (culture) and governance (see Chapter 8).

Processes

The fourth generation distinguishes between three categories of processes: management processes, which steer the organisation; business processes, which directly generate value; and support processes, which create the conditions for this to happen. From this perspective, processes are not mere appendages to structures, but independent objects of design.

Modes of development

Organisations do not develop along a linear causal chain, but rather in two alternating modes: optimisation (improving existing routines) and renewal (breaking with existing routines and establishing new ones). Both modes are necessary; striking a balance between them is a key management task.

2.3.2 The communicative shift

A conceptual shift in the fourth generation compared to Bleicher is the significantly stronger emphasis on communication. Following on from the theory of communicative action (Habermas 1981), the enterprise is understood as a network of communicative practices. Strategy is not primarily a plan, but an understanding; culture is not a set of values, but an ongoing negotiation of what is considered appropriate.

This communicative turn is doubly significant for AGENTONOMICS. Firstly, because language models as AI agents operate precisely at the level that the new St. Gallen model identifies as the central level of action: communication. Secondly, because polycentric agent networks must establish their cohesion through communicative negotiation – a point which is explored in greater depth in Chapter 7.

2.4 Complementary building blocks: networked thinking, the learning organisation, cybernetics

The St. Gallen Model is not monolithic. It draws on several strands that are equally relevant to AGENTONOMICS.

2.4.1 Networked thinking

In their methodology of 'networked thinking', Gomez and Probst (1995) developed tools with which complex problem areas can be systematically explored: influence matrices, impact diagrams, scenarios and scope for action. In today's age of AI, these tools are not obsolete but more relevant than ever: they provide precisely the structured framework within which an AI agent can process complex issues for its human client.

2.4.2 The Learning Organisation

Peter Senge (2008) brought the concept of the learning organisation into the mainstream of management theory. It is characterised by five disciplines: personal mastery, mental models, shared vision, team learning and systems thinking. When applied to networks of AI agents, it becomes clear that an organisation whose operational layer consists of learning agents must ensure its capacity for reflection at a higher level – otherwise the agents will optimise their way into local dead ends. This level of reflection is a key task of the AI CEO (Chapter 8).

2.4.3 Cybernetics of Management

Stafford Beer's Viable System Model (Beer 1972) provides a cybernetic architecture for viable systems, organised into five subsystems. This architecture is highly relevant to the design of AI agent networks: it shows the minimum functions a system must fulfil in order to be viable – ranging from operational execution, through coordination and operational management, to strategic and normative control. AGENTONOMICS implicitly incorporates this architecture into its modular structure.

2.5 What remains, what needs to be expanded

Three points should be noted where AGENTONOMICS builds on the St. Gallen legacy and where it goes beyond it:

Adoption: the holistic architecture

The triad of normative, strategic and operational dimensions, the reference to stakeholders, the distinction between regulatory elements and processes, and the fundamental communicative approach are fully adopted by AGENTONOMICS. In terms of its architecture, the ADMRF (Chapter 4) is a St. Gallen model for AI agent networks.

Adoption with a shift in emphasis: polycentricity

The transnational organisational model (Suckfüll 1994) had already demonstrated, within the St. Gallen context, that internationally operating companies cannot be managed in a purely centralised or decentralised manner, but require a polycentric architecture. In AGENTONOMICS, this insight is applied to the AI agent level: agent networks, too, are neither purely hierarchical nor chaotic, but are organised in a polycentric manner (Chapter 7).

Extension: algorithmic coherence check

Whereas the classical St. Gallen model locates the function of ensuring coherence cognitively within management, AGENTONOMICS adds an algorithmic layer. The AGENTONOMICS algorithm technically checks for coherence between the normative, strategic and operational levels and suggests inconsistencies for resolution. This does not replace the St. Gallen model, but rather extends it into an AI-enabled variant.

Extension: the AI agent as an economic entity

Finally – and this is the most fundamental extension – AGENTONOMICS introduces a new type of actor into the model: the AI agent as an economic entity. Until now, the St. Gallen Model has implicitly assumed human and organisational actors. With the AI agent, a hybrid entity enters the scene, combining technical artefacts and economic decision-making logic. The implications of this extension are explored in Chapters 5 and 6.

> *The St. Gallen Management Model is not outdated – on the contrary: it is the one management theory that is particularly adaptable and can truly cope with the complexity of polycentric AI agent networks. It simply needs to be thought through to its logical conclusion.*

Bibliography

Beer, S. (1972): Brain of the Firm. The Managerial Cybernetics of Organisation. Allen Lane, London.

Bertalanffy, L. von (1968): General System Theory. Foundations, Development, Applications. Braziller, New York.

Bleicher, K. (2011): The Concept of Integrated Management. Visions – Missions – Programmes. 8th, updated and expanded edition, Campus, Frankfurt am Main.

Gomez, P.; Probst, G. (1995): The Practice of Holistic Problem-Solving. Thinking in Networks – Acting Entrepreneurially – Persuading Personally. Haupt, Bern.

Habermas, J. (1981): Theory of Communicative Action. Two volumes, Suhrkamp, Frankfurt am Main.

Rüegg-Stürm, J.; Grand, S. (2020): The St. Gallen Management Model. Management in a Complex World. 3rd edition, Haupt, Bern.

Senge, P. M. (2008): The Fifth Discipline. The Art and Practice of the Learning Organisation. 10th edition, Schäffer-Poeschel, Stuttgart.

Suckfüll, H. (1994): The Transnational Organisational Model. PhD thesis, University of St. Gallen.

Ulrich, H. (1968): The Organisation as a Productive Social System. Foundations of General Organisation Theory. 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 – A Historical and Economic Context

To understand AGENTONOMICS within its proper context, it is necessary to examine the development of digital economic models. The current debate surrounding AI agents does not stand in isolation, but is the result of several successive shifts in economic paradigms. The transition from the information economy, through the data economy, to the AI capabilities economy reveals a pattern that helps us understand the logic of the present – and which opens up the space for argumentation into which AGENTONOMICS enters (Knoll and Suckfüll 2024).

3.1 Three economies: information, data, capabilities

3.1.1 The information economy of the 1990s

In the information economy of the 1990s, existing business models were primarily transferred to digital channels. Companies initially used the internet as an additional communication and sales channel. Although numerous innovative business models emerged (Shapiro and Varian 1999), many early approaches proved unsustainable. The dot-com bubble and the dot-com crash that began in 2000 marked an initial learning process: not every technical innovation automatically leads to sustainable value creation.

This phase was described theoretically in works on network economics (Shapiro and Varian 1999) and platform economics (Rochet and Tirole 2003). The key concepts – network effects, lock-in, standard competition – still form the vocabulary used to describe the business models of the digital economy today. However, they are based on the assumption that the good 'information' is largely freely interchangeable – an assumption that must be revised in the subsequent data economy.

3.1.2 The data economy

The data economy gradually evolved from the information economy. Here, data is increasingly understood as a strategic resource and a potential asset (Mayer-Schönberger and Cukier 2013). Value creation shifted from the mere provision of information to the systematic collection, analysis and monetisation of data. Large digital platforms have established business models based on economies of scale, network effects and lock-in mechanisms (Zuboff 2018).

Despite their economic success, the data economy revealed three fundamental structural tensions that are central to understanding the emergence of AGENTONOMICS:

Asymmetrical power relations

Control over data is concentrated in the hands of a few global players. Individuals and organisations that provide data have only limited opportunities to participate appropriately in the resulting value creation. This trend is currently intensifying exponentially. As Knoll and Suckfüll (2024) highlight, the average adult in a developed country is estimated to generate several hundred gigabytes of personal data per year

– ranging from LinkedIn profiles and online job applications to digital invoices and contracts. The economic value of this data is estimated at between 10,000 and 15,000 euros per person per year – and is almost entirely held by the major platforms.

Dysfunctional data markets

Although data is recognised as having economic value, efforts at both national and international level have failed to establish transparent, fair and scalable markets for data exchange. Issues relating to the allocation of rights, valuation and remuneration remain largely unresolved (Otto et al. 2019). The result is the paradoxical situation in which data is of high economic value but cannot be traded by individuals – leading to the de facto expropriation of data producers.

Regulatory fragmentation

Government intervention – particularly through data protection and digital regulation – has primarily aimed at damage limitation and control. However, rather than creating independent, productive market mechanisms, it has in some cases led to tendencies towards isolation and barriers to innovation. As Knoll and Suckfüll (2024) put it so succinctly: the principle of 'data minimisation', which is particularly pronounced in Germany, is a 'direct path to impoverishment' unless it is counterbalanced by productive mechanisms of individual data sovereignty.

3.1.3 The AI skills economy

With the breakthrough of powerful generative AI systems – particularly since the release of large language models from 2022 onwards (Brown et al. 2020; OpenAI 2023; Bommasani et al. 2021) – the focus of economic value creation is shifting once again. Data is not losing its significance, but is increasingly becoming an intermediate product: it is only through AI models that it can be transformed into capabilities relevant to action.

This development can be described as a transition to an AI capabilities economy. In this economy, the focus is no longer on isolated data sets, but on AI-supported capabilities such as:

- decision support,
- pattern recognition,
- forecasting and planning,
- autonomous interaction with markets and organisations,
- synthesis of complex bodies of knowledge.

What is economically relevant here is not merely technical performance, but the ability to deploy these AI capabilities in a targeted, sustainable and value-adding manner. Knoll and Suckfüll (2024) pointedly refer to a second leap in sovereignty: following the leap towards data sovereignty (which Germany has largely missed), the next step is the leap towards 'skill sovereignty'.

3.2 Signs of the shift in the real economy

The shift from a data-based to a skills-based economy has long been documented in the day-to-day reporting of the German business press. Knoll and Suckfüll (2024) cite two symptomatic examples, which will be explored further here:

3.2.1 The case of the Federal Employment Agency

The Chair of the Federal Employment Agency, Andrea Nahles, has publicly pointed out that, of 114,000 staff members, 35,000 will retire over the next ten years, with no new recruits available in the traditional sense. Consequence: The agency's services could only be guaranteed in future through automation – through the use of artificial intelligence. This is no longer merely the wish of an efficiency consultant, but the sober assessment of a major public administration.

3.2.2 The Klarna case

The Swedish financial services provider Klarna is no longer recruiting new staff – with the exception of engineers. Within a year, the workforce was reduced from 5,000 to 3,800 through the use of AI; in the long term, the target is 2,000 – whilst the total volume of work is increasing. This is not staff reduction in the traditional sense, but rather a structural shift in functions. In the meantime, this radical approach has proved to be excessive, but the underlying trend remains intact.

Both cases exemplify what the 'AI skills economy' actually means: it is not the number of human employees, but the number of economically relevant skills (whether performed by humans or replicated by agents) that becomes the decisive factor. In many respects, this shift is comparable to developments in manual labour since the 1970s – with the crucial difference that, in the field of information processing, there are virtually no fundamental limits to what a computer can do, whereas industrial robots have so far been unable to match the dexterity of the trained human hand (Knoll and Suckfüll 2024).

3.3 The question of sovereignty: from data sovereignty to capability sovereignty

In their Technical Report TUM-I24112, Knoll and Suckfüll (2024) develop an argument that is constitutive of AGENTONOMICS, which will be elaborated on in more detail here because it forms the normative foundation of the entire frame of reference.

3.3.1 The bizarre status quo

The central diagnosis is this: we find ourselves in a fundamentally absurd situation. On the one hand, data protection legislation and the lack of operational mechanisms for individual data use prevent social value creation in many areas. On the other hand, every single individual must disclose an enormous amount of personal data in order to participate in a modern, digitally mediated life – and receives little or nothing in return.

The state's actual role should be to provide every individual with a comprehensive set of tools enabling them to gain sovereignty over their own data and to decide for themselves what happens to it. Even if the legislation and operational framework for such individual data sovereignty are not yet in place, it is not too late to create them.

3.3.2 Skill sovereignty as the next stage

The key point made by Knoll and Suckfüll (2024) is this: whilst policymakers are still debating how to make individual data sovereignty operational, the next sovereignty issue is already on the horizon – that of 'skill sovereignty'.

Core thesis (Knoll and Suckfüll 2024)

Every knowledge worker processes input data or information according to rules that they have acquired and refined through their education and professional life. In this way, each person develops their own expertise, their own modus operandi. This individual professional expertise is increasingly being transferred to computers, that is, to AI. The mechanisms of 'skill transfer' are becoming ever more efficient. If this transfer goes wrong, knowledge workers will be worse off than the machine-breakers of the 19th century. If it succeeds, everyone can multiply their productivity.

The crucial point is this: anyone who hands over their skills to a machine should remain in control of those skills – even if they are multiplied and deployed in parallel. In concrete terms, this means that whilst a person may work for several clients at the same time, those clients cannot use the person's 'functional replicants' independently. The person should retain the 'copyright' – that is, full control over their entire stock of replicants.

3.3.3 The state's responsibility

This argument leads to a clear regulatory demand: the provision of computing capacity and basic AI services is – much like schools, libraries or transport networks – an essential public service. Without it, an AI-based skills economy cannot develop – at least not one that supports Europe's overall sovereignty rather than undermining it.

Europe thus finds itself once again at a crossroads: do we want fear-driven legislation that hinders innovation – or legislation that actively shapes emerging developments and strengthens the European ideal of individual rights by giving people the opportunity not to lose, but to gain, sovereignty over their data and capabilities? AGENTONOMICS is a constructive contribution to the second option.

3.4 From the data trusteeship debate to AGENTONOMICS

An important conceptual precursor to AGENTONOMICS can be found in the – not least politically relevant – debate on data trust models. The data trust aims to balance the interests of three generic stakeholder perspectives in a legally sound manner:

1. Data providers – the actors who produce data.
2. Data users – the actors who exploit data for economic gain.

3. Data owners – the parties to whom the data is legally entitled.

The first relevant contributions to this discussion can be found in the 'Initiative for a DEDATE Data Trust' (Fraunhofer MOEZ and Suckfüll 2018) and in the further developed, as yet unpublished form of a working paper, DEDATE 2.0. Enabling those with legally verifiable rights of disposal to negotiate 'their' data with others in a sovereign manner also from a technical perspective represents the ultimate goal of a functioning data economy.

AGENTONOMICS adopts this fundamental idea but shifts the focus: instead of static data rights, dynamic, AI-supported capabilities take centre stage in value creation. AI agents function – alongside their own core competencies – as autonomous and sovereign AI capability intermediaries capable of negotiating rights, interests and value creation on a case-by-case basis. This takes place with a level of complexity and speed that would be virtually impossible to achieve with purely human or institutional solutions.

3.5 Polycentric AI vs. AGI – the strategic distinction

The emergence of AGENTONOMICS is closely linked to a critical examination of centralised AI strategies, as succinctly expressed by the slogan 'only size matters'. Current global competition is characterised by the attempt to develop AI systems that are as comprehensive and generalised as possible (AGI) – with:

- enormous economies of scale,
- high capital intensity,
- centralised governance structures,
- oligopolistic to monopolistic market positions.

AGENTONOMICS, by contrast, pursues a network-based, polycentric approach. Instead of a single, all-encompassing AI, the focus is on many specialised, cooperating AI agents. Value creation arises not primarily through centralisation, but through coordination, division of labour and situational leadership within networks. Drawing on the principle of a kaleidoscope, AGENTONOMICS agents can form configurations and pool resources as the situation demands: imagine Siemens, BMW and Deutsche Telekom defining themselves – and thus their interfaces – as AI agents and cooperating with one another to tackle shared tasks.

This approach should not be understood as a technical counter-proposal, but as an economically and socially motivated alternative – particularly with regard to issues of participation, ownership structures and sustainable value creation. A centralised AI dominance carries the risk that a lack of competition will cause the distribution of wealth to become so asymmetrical that the resulting lack of purchasing power will undermine the entire capitalist system.

3.6 AI agents as a new category of economic actors

This is precisely where the AGENTONOMICS concept comes in. It goes a step beyond existing AI applications by conceiving of AI agents as economically active actors. These agents are not merely tools or assistance systems, but:

- possess their own decision-making logic,
- act economically autonomously within defined parameters,
- interact with other agents, organisations and people,
- are geared towards sustainable value creation and self-optimisation.

In the terminology of economics, AI agents can thus be understood as a new form of 'economic agents'. They participate in exchange relationships, compete or cooperate, and make decisions under conditions of scarcity. The interacting economic agents, in turn, form a market that can be described in its own right, characterised microeconomically by its own conditions for competition, market participants, exchange relationships and price formation.

Of particular interest here is the transition phase – from a data-driven economy to an AI-capabilities-driven economy. The exclusive interaction between AI agents has, to date, been an idealised scenario. In this respect, the relationships and interfaces between humans, AI agents and other AI agents need to be rethought from the ground up. AGENTONOMICS represents precisely this rethinking.

3.7 Theoretical classification of AGENTONOMICS

AGENTONOMICS is neither a short-lived fad nor a buzzword. Nor is it conceived as an empirical-descriptive theory in the strict sense. Rather, it is an ideal-theoretical model that provides a conceptual framework for AI agents acting in an economic capacity. Ideal-theoretical models are characterised by the fact that they:

- abstract complex real-world phenomena,
- isolate key mechanisms of action,
- provide guidance for analysis and design.

AGENTONOMICS follows this logic by modelling AI agents as economic ideal types. These agents possess defined capabilities, decision-making logics and incentive systems, even though real-world implementations may deviate from them. The standardisation of AI agents leads, through the establishment of norms, to a value-adding ecosystem: compatible interfaces increase the likelihood of productive agent interactions.

3.7.1 Positioning at the interface between business administration and economics

Within the field of economics, AGENTONOMICS can be situated at the interface between business administration and economics.

The business administration perspective

From a business administration perspective, the central question is how economic units are managed, coordinated and aligned with value creation. AGENTONOMICS broadens this perspective by conceiving AI agents as controllable, yet not entirely determinable, actors. Key points of reference include management and control models, organisational and network structures, strategic decision-making processes, and governance and incentive systems.

Economic perspective

From an economic perspective, AGENTONOMICS builds on microeconomics. AI agents are modelled as market participants who engage in exchange relationships, shape competitive and cooperative situations, and are involved in price-setting processes. Unlike traditional actors, however, AI agents are highly scalable, networkable and capable of learning. This requires an extension of microeconomic models, particularly with regard to network and platform economies.

3.7.2 Distinction from related discourses

AGENTONOMICS deliberately distinguishes itself from several established discourses:

- Pure AI research focuses on algorithms, models and technical performance, but rarely addresses economic governance issues.
- The digital platform economy analyses market-dominating structures, but usually assumes the presence of human or organisational actors.
- Debates on automation and substitution view AI primarily as a substitute for human labour and neglect its role as an autonomous agent.

AGENTONOMICS, by contrast, integrates technical, economic and institutional perspectives into a distinct field of study. A multidisciplinary approach is fundamental. The disciplines involved include, in particular, computer science (development of autonomous systems), business administration (control and value creation models), economics (market and competition structures) and law (issues of liability, ownership and governance).

3.7.3 Research and Design Programme

Beyond its theoretical classification, AGENTONOMICS should be understood as a research and design programme. It is not merely about analysis, but explicitly about the development, testing and evaluation of new forms of economic activity:

- the conceptualisation of AI agents as economic actors,
- Designing incentive and governance structures,
- Development of experimental business and organisational models,
- empirical testing in real or simulated environments.

In this sense, AGENTONOMICS bridges the gap between theory, experimentation and practice. The development of AI agents lends itself particularly well to the form of Minimum Viable Products, as systematised by Ries's (2011) Lean Startup concept. The 'Doz. Dr. AGENTONOMICS' is one such MVP.

> *AGENTONOMICS is the economically consistent response to a technological shift that took place some time ago – but has not yet found its way into the language of management theory.*

Bibliography

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, pp. 1877–1901.

Fraunhofer MOEZ; Suckfüll, H. (2018): Initiative for a data trust: DEDATE. Discussion paper, Fraunhofer Centre for International Management and Knowledge Economy, Leipzig.

Knoll, A. C.; Suckfüll, H. (2024): From data sovereignty to skill sovereignty. Technical Report TUM-I24112, Department of Computer Science, Technical University of Munich.

Mayer-Schönberger, V.; Cukier, K. (2013): Big Data. The revolution that will change our lives. Redline, Munich.

OpenAI (2023): GPT-4 Technical Report. arXiv:2303.08774.

Otto, B.; ten Hompel, M.; Wrobel, S. (eds.) (2019): International Data Spaces. Reference Architecture Model. Fraunhofer Society, Munich.

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), pp. 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): The Age of Surveillance Capitalism. Campus, Frankfurt am Main.

4 The AGENTONOMICS Design & Management Reference Framework (ADMRF)

This chapter forms the operational core of the handbook. It describes the overall logic of the AGENTONOMICS Design & Management Reference Framework (ADMRF) as an integrated reference framework. The six modules of the ADMRF are briefly introduced here and positioned within their broader context; Chapters 5 to 10 then explore each module in detail. Anyone wishing to build an AI agent to answer questions on business administration and the St. Gallen School will find in the ADMRF the knowledge architecture that such an agent must be based on.

4.1 Aims and theoretical background

The ADMRF builds upon the St. Gallen Management Model – in particular on Bleicher's (2011) concept of Integrated Management and on the fourth generation proposed by Rüegg-Stürm and Grand (2020). The six modules link the normative, strategic and operational levels of a holistically integrated management framework. The ADMRF is conceived as a management framework for network organisations comprising AI agents that comply with a common standard – the AGENTONOMICS standard.

The governance concepts for polycentric or quasi-hierarchical networks are based on the doctoral thesis 'The Transnational Organisational Model' (Suckfüll 1994). This model is particularly useful because, as early as the 1990s, it sought to address the problem of how an organisation comprising autonomous, locally accountable units can simultaneously achieve global cohesion and local adaptability – precisely the problem that AI agent networks must also solve today.

AI as a technology is utilised in two ways for the ADMRF. On the one hand, it serves to optimally calculate the interrelationships between the modules – and the involvement of stakeholders. On the other hand, it is itself the subject of control, as the economic entities to be managed are AI agents. This dual role is constitutive: AI is both the subject and the instrument of control.

The six modules presented below are not considered in isolation, but rather as interdependent phases of an integrated development and control process. Interdependencies and interactions must be explicitly taken into account in order to understand the overall logic – the AGENTONOMICS algorithm – and to recognise its fields of application.

4.2 The second-generation Turing test as a benchmark

The autonomy target for AI agents – to operate successfully in the market whilst being able to utilise all the degrees of freedom of an organisation, i.e. a legal entity – is based on the requirements of a second-generation Turing test, as formulated by Suleyman and Bhaskar (2023) in 'The Coming Wave'.

Suleyman's modern Turing test (paraphrased)

An AI system passes the modern Turing test if it can independently complete the task "Using an initial capital of US$100,000, earn *US$1 million within a few months on a retail web platform*, whilst complying with all applicable laws" without any further specific instructions – including devising a business idea, concluding contracts, marketing and handling tax matters.

AGENTONOMICS AI agents represent the next step towards meeting these requirements. The second generation of the Turing Test is not primarily a technical criterion, but a business management one: it assesses the ability to act economically. The ADMRF serves as the unifying framework linking the initial business idea, management logic and practical implementation – right through to successful operation.

4.3 Areas of application: Greenfield and Brownfield

The ADMRF is equally applicable to two different scenarios, which are commonly referred to in business practice as 'greenfield' and 'brownfield' cases:

Greenfield: AI agents on a greenfield site

Here, the AI agent is designed from scratch. There are no legacy systems, no established processes, and no pre-existing stakeholder expectations. The ADMRF guides users through all phases, from the Definition Framework through the Building Blocks to stakeholder management. This application is typical for start-ups developing agent-based business models.

Brownfield: Transformation of existing organisations

Here, an existing organisation – a traditional company, a public authority, an NGO – is gradually transformed into an AI agent or a network of such agents. The ADMRF acts as a roadmap along which legacy systems are converted into agent-enabled structures. This application is by far the more common – and the more challenging.

In both cases, the ADMRF acts as a unifying framework linking the initial business idea, management logic and practical implementation. To put it another way: it describes how AI agents are developed and iteratively refined, from conceptual definition through modular implementation and economic management right through to market readiness.

4.4 An overview of the six modules

At the heart of the ADMRF are six modules which – in the language of classical business management – represent six ideal-typical management perspectives. They are presented here in summary form; Chapters 5 to 10 explore them in detail.

4.4.1 Module 1: Definition Frame – the entrepreneurial framework

The first module, the AGENTONOMICS Definition Frame, describes the agent through six parameters: autonomy and sovereignty; claims to capability and competence; regulation and self-regulation; identity and ethics; AI model and (virtual or real) embodiment; and the ability to learn and teach (meta-agent). Fundamentally, the AGENTONOMICS definition of an AI agent requires that it be distinguishable as an economic entity and capable of cooperating with other agents – organised and monitored by an intermediary function. The ownership structure enables independence from other agents, meaning that these can be described as polycentrically governed organisations.

The module is a tool for tackling the business challenge of defining an AI agent that meets the requirements of an AGENTONOMICS agent. All management perspectives (CEO, CFO, CTO, CSO) must be involved collectively as a board. See Chapter 5.

4.4.2 Module 2: Five Building Blocks – Technical Implementation

The second module breaks down the AI agent into five legally and technically distinct building blocks: agent algorithm, agent interface, AI model, tools and vectors, and agent ecosystem intermediary. Each building block is represented by its own logo; when combined, all the logos form a graphical representation of the AI agent. The technical implementation of the various building blocks – or their optimised composition for the AI agent defined in advance by the framework – is intended to link the central tasks of the agent's technical and business development.

From a business management perspective, this phase can be described as the integration of the business development process with that of product development – as a process that follows on from the vision and mission developed in the definition framework. The '5 Building Blocks' module acts as an interface between business and technical issues. It essentially reflects the perspective of the CTO. See Chapter 6.

4.4.3 Module 3: Network Intermediary – Agency and Governance

The third module deals with the network trustee or network intermediary. AI agents are capable of cooperation and, as autonomously interacting units within a polycentrically organised network, can generate value synergistically. Network cohesion arises through a quasi-democratic negotiation process amongst the AGENTONOMICS agents themselves. In this respect, the network trustee represents the outcome of a hierarchy-free discourse amongst the participating AI agents – analogous to the theory of communicative action (Habermas 1981).

The functions developed by the network through consensus and intended for delegation are implemented operationally by the so-called network intermediary. An example of this is the owners' association of a property, whose rights and obligations are carried out operationally by the property management company. The specification of such context-specific intermediary functions – data trustees, broker agents, clearing houses, etc. – which coordinate rights, data, capabilities and transactions, must be taken into account from the outset when designing an individual AGENTONOMICS AI agent . See Chapter 7.

4.4.4 Module 4: AI-CEO Management Concept – Operationalising Organisational Principles

The fourth module – the management tool or AI-CEO – operationalises the organising principles of vision, strategy, structure, agency (culture) and governance as a holistically conceived support for corporate decision-making at management and board level. This is an AI-supported, holistic management concept that is intended as a further development of Bleicher's (2011) concept of integrated management. The focus of this framework, also known as the St. Gallen Management Concept, is centred on networks of AI agents. See Chapter 8.

4.4.5 Module 5: The PuK Tool – the quantitative component

The fifth module is the PuK instrument (planning and control instrument). It is based on the PuK model developed by Hahn and Hungenberg (2001). This concept is also tailored to AI agents operating within a network and integrated using an AI model. The AGENTONOMICS matrix is systematically populated with the five building blocks listed horizontally and, on the vertical axis, with the phases Design / Build / Operate / Own / Transfer (DBOT logic), to each of which specific balance sheets and profit and loss accounts are assigned – in accordance with the guidelines set out by Hahn and Hungenberg (2001) and the principles of financial management (Berk and DeMarzo 2020). The module performs the duties of a CFO in traditional, human-dominated organisations. See Chapter 9.

4.4.6 Module 6: Stakeholder Communicator – the AI-driven discourse

The sixth module is the AGENTONOMICS Stakeholder Communicator. It derives target-group-specific communication formats – for employees, investors, strategic partners and other stakeholders – from the AGENTONOMICS PuK tool or the Excel tool populated for the respective AI agent. The Pitch Master is based on the work of Guy Kawasaki (2015), which he developed whilst working as Steve Jobs' evangelist at Apple. This module takes on the role of stakeholder communication; it can also be referred to as a CSO (Chief Sales Officer). See Chapter 10.

4.5 How the modules are interconnected

The six modules of the ADMRF are interconnected via a central AI. The AGENTONOMICS algorithm facilitates this interconnection to provide the optimal solution to specified problems. Three types of connection can be distinguished:

Sequential connection: the lifecycle

In their first interpretation, the six modules form a lifecycle: from the Definition Frame through the Building Blocks and the Intermediary to the AI-CEO, on to the PuK and finally to stakeholder management. This sequential interpretation corresponds to the typical sequence in a greenfield project.

Cyclical connection: the feedback loops

In reality, however, the ADMRF is not a linear process, but a cyclical one. Feedback from the Stakeholder Communicator flows back into the PuK; the PuK can adapt the AI-CEO concept; the AI-CEO concept alters the Intermediary architecture; this, in turn, alters the Building Blocks and, ultimately, the Definition Frame as well. Any adjustment to one module can trigger adjustments to all the others. It is precisely these feedback loops that the AGENTONOMICS algorithm is designed to make mathematically manageable.

Cross-linking: the organising principles

Running across all modules are the five organising principles: vision, strategy, structure, agency and governance. They do not constitute a separate module, but rather the coordinate system within which the modules are situated. Every entry in a module can be classified according to these five organising principles and checked for its coherence with other modules.

4.6 The AGENTONOMICS Matrix as a condensed management tool

The central visual representation of the ADMRF is the AGENTONOMICS matrix. It intersects two dimensions:

- Horizontally, the five building blocks (agent algorithm, agent interface, AI model, tools and vectors, ecosystem intermediary).
- Vertically, the DBOT phases (Design, Build, Operate, Own, Transfer) of the agent lifecycle.

For each cell in this matrix, KPIs, CAPEX/OPEX assumptions, risks and responsibilities are recorded, where relevant. The matrix therefore serves as:

- a blueprint (which component in which phase),
- a budgeting tool (which costs in which phase),
- a risk management tool (which risks in which phase),
- a communication tool (which stakeholders in which phase).

The matrix is developed in detail in Chapter 9; at this stage, it is important to recognise it as the shared working document for all modules. Anyone planning an AGENTONOMICS agent within a real organisation communicates via the matrix – not via six separate documents.

4.7 First application: ‘Doz. Dr. AGENTONOMICS’

The ADMRF modules are introduced for the first time as part of the lecture ‘AI Agents in Business Administration’ during the winter semester 2025/26 at TUM, and are linked together via an AI model. The AGENTONOMICS algorithm will be applied and gradually refined through case studies. This marks the birth of ‘Doz. Dr. AGENTONOMICS’ – the first application based on the algorithm.

This application is a deliberately small but well-defined use case: an AI agent that acts as a tutor and coach to help students understand and apply the ADMRF. It is both a subject of learning and a learning

medium – a classic minimum viable product in the sense defined by Ries (2011). Feedback from students' interactions with Doz. Dr. AGENTONOMICS feeds back into the further development of the algorithm. In this way, the lecture itself becomes part of its own subject of investigation: it tests how AI agents support learning processes, structure knowledge and enable individual learning pathways.

> *The ADMRF is not the end point of thought, but its operating theatre. It is intended to be a frame of reference that is good enough for students, researchers and practitioners to work with – and simple enough for them to critique it seriously.*

Bibliography

Berk, J.; DeMarzo, P. (2020): Fundamentals of Finance. Analysis, Decisions and Implementation. 6th, updated edition, Pearson, Munich.

Bleicher, K. (2011): The Concept of Integrated Management. Visions – Missions – Programmes. 8th, updated and expanded edition, Campus, Frankfurt am Main.

Habermas, J. (1981): Theory of Communicative Action. Two volumes, Suhrkamp, Frankfurt am Main.

Hahn, D.; Hungenberg, H. (2001): PuK – Value-Oriented Controlling Concepts. 6th, revised and expanded edition, 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. Technical Report TUM-I24112, Department of Computer Science, Technical University of Munich.

Ries, E. (2011): The Lean Startup. Crown Business, New York.

Rüegg-Stürm, J.; Grand, S. (2020): The St. Gallen Management Model. Management in a Complex World. 3rd edition, Haupt, Bern.

Suckfüll, H. (1994): The Transnational Organisational Model. PhD thesis, University of St Gallen.

Suleyman, M.; Bhaskar, M. (2023): The Coming Wave. Crown, New York.

5 Module 1: The AGENTONOMICS Definition Frame

The definition of the agent using the AGENTONOMICS framework forms the starting point of the AGENTONOMICS lifecycle. In this phase, a description of the relevant parameters determines who or what the AI agent is to represent economically in order to be able to operate as an economic entity. The Definition Frame is the corporate framework, comparable to a company's constitution or the articles of association of a start-up. It is the responsibility of the Board – that is, the collective body bearing corporate responsibility – rather than a single function.

5.1 Function of the Definition Frame

In the classical St. Gallen understanding (Bleicher 2011; Rüegg-Stürm and Grand 2020), the definition frame corresponds to the normative level of management: this is where the fundamental questions of purpose are answered. Unlike in traditional organisations, however, it is not sufficient for AI agents to formulate a corporate constitution in prose – the frameworks must also be machine-readable and verifiable. The frame achieves both: it is a human-readable normative statement and a machine-readable specification.

The frame addresses six key questions, each of which is operationalised by a parameter:

1. What degree of autonomy and sovereignty should the agent have?
2. What capabilities and competencies should it possess?
3. What regulations and self-regulatory mechanisms apply to it?
4. What identity, ethics and personality does it possess?
5. Which AI model and which form of embodiment underpin it?
6. What learning and teaching capabilities should it demonstrate?

The six parameters must first be described in qualitative terms. Where appropriate, they can be specified quantitatively, for example using scales, key figures or formal specifications. Each parameter is explored in detail below.

5.2 Parameter 1: Autonomy and sovereignty

In everyday language, autonomy and sovereignty are often confused or treated as synonymous. In the AGENTONOMICS context, however, a clear distinction must be made between them, as they refer to different concepts and have different implications for the design of an agent.

Definition: Autonomy and sovereignty

Autonomy refers to the degree to which an agent can perform a task independently without external intervention. Sovereignty refers to the degree to which an agent can actually dispose freely of the resources allocated to it – data, capabilities, rights. An agent may have a high degree of autonomy and a low degree of sovereignty – or vice versa.

Example: A fully automated translation system is highly autonomous (it translates without human intervention), but has low sovereignty (it is not permitted to monetise, forward or utilise the translated texts itself). A human translator with their own agency may be less autonomous (they accept commissions), but highly sovereign (they decide on the price, acceptance and utilisation of their work).

5.2.1 The six levels of autonomy (0–5)

The determination of the degree of autonomy is based – in a didactically productive analogy – on the six SAE levels (0–5) of automated driving (SAE 2021), which are applied here to economic activity:

- **Level 0** – **No autonomy**: The agent is merely a tool. Every decision is made by a human; the agent merely carries out the instructions.
- **Level 1** – **Assistance**: The agent provides suggestions within a narrowly defined scope of tasks; the decision is made by the human.
- **Level 2** – **Partial autonomy**: The agent carries out defined subtasks independently; the human monitors and intervenes.
- **Level 3** – **Conditional autonomy**: The agent acts independently within a clearly defined area of business; in defined exceptional cases, a human is called upon to make a decision.
- **Level 4** – **High autonomy**: The agent acts completely independently within its business area; human intervention is the exception.
- **Level 5** – **Full autonomy**: The agent operates across business areas, generates its own objectives and adapts its business areas independently. This level is currently discussed primarily for theoretical purposes.

5.2.2 Sovereignty and its implications

Whilst autonomy is a characteristic of behaviour, sovereignty is a characteristic of rights of disposal. Three dimensions of sovereignty are to be specified within the AGENTONOMICS framework:

Data sovereignty

What data can the agent use, pass on or withhold? Here, AGENTONOMICS ties in directly with the data trusteeship debate (Fraunhofer MOEZ and Suckfüll 2018): a sovereign agent has clear, enforceable data rights.

Capability sovereignty

Which capabilities can the agent independently offer, licence or withhold? This is where AGENTONOMICS' specific innovation lies in comparison to traditional data markets (Knoll and Suckfüll 2024).

Decision-making sovereignty

Which decisions can the agent make definitively without escalation to human authorities? This dimension is directly linked to the degree of autonomy.

5.3 Parameter 2: Claims to ability and competence

Ability and competence must be distinguished from one another. Cycling serves as a suitable example. Whilst more or less all healthy people have the ability to ride a bicycle, only those who have actually learnt to cycle possess the competence to do so. Applied to AI agents:

Definition: Ability and competence

Ability is an agent's fundamental capacity to perform a task – rooted in its model, its tools and its embodiment. Competence is the actual, tried-and-tested execution of this task, proven in productive application. Abilities are potential; competences are actualised.

This distinction has practical implications. An agent based on a foundation model (Bommasani et al. 2021) potentially possesses a vast number of abilities – but it only develops competencies in those areas where it is actually trained, fine-tuned and tested in practice. An AGENTONOMICS agent is therefore not described by all the abilities that its model theoretically permits, but by the profile of its actual competencies.

The key question is: Which competencies are certified, and which are merely claimed? The trustworthiness of an AI agent depends crucially on this distinction being made transparent. This ties in with discussions on AI Assurance (NIST 2023; EU AI Act 2024).

5.4 Parameter 3: Regulation and self-regulation

Regulation and self-regulation refer to the technical, organisational and legal measures by which an AGENTONOMICS agent implements all legal, ethical and other requirements. In a polycentric network, two levels of operation can generally be identified:

Inter-agent measures

At the level of agent-to-agent relationships, protocols and standards ensure that cooperation is possible in the first place. This concerns interfaces, authentication, proof of identity, contract conclusion and dispute resolution. Many of these measures are technical in nature and are explored in greater depth in Chapter 6 (Building Blocks).

Intermediary-related arrangements

At the level between agents and the network intermediary, overarching rules ensure cohesion. This is where the quasi-constitutional parameters lie: Which issues are even negotiated within the network? Which decisions are reserved for the intermediary? What escalation procedures are in place? This level is dealt with in detail in Chapter 7.

Self-regulation is not the opposite of regulation, but a specific form of it: the agent checks for themselves whether their behaviour complies with the rules and – ideally – corrects it proactively. This form of self-regulation is economically attractive (it reduces the transaction costs of external control) and ethically desirable (it embeds responsibility within the agent itself). However, it is also risky: an agent who conceals its own breaches is more dangerous than one that has no self-regulation at all (Russell 2019).

The EU AI Act (*Regulation (EU) 2024/1689*) provides a concrete regulatory framework that sets out different requirements depending on the risk class of the application domain. An AGENTONOMICS agent must clearly identify, within the definition framework, which risk class it falls into – and which compliance mechanisms follow from this.

5.5 Parameter 4: Identity, Ethics and Personality

Identity, ethics and personality are the – at first glance, most surprising – parameters of an AI agent. Yet they are precisely the parameters that distinguish an agent from any other software component. From a marketing perspective, one might refer to the agent's Unique Selling Proposition (USP); in terms of business ethics, the focus is on the reliability of its behaviour; philosophically, it concerns the question 'Who am I, and if so, how many?' – as a reminder that a replicable agent is particularly subject to this question.

5.5.1 Identity

Identity here does not refer to philosophical self-identity, but rather to the economically and legally relevant identifiability of an agent. It comprises:

- a unique technical address (identifier),
- an assignable economic responsibility (owner, responsible person or organisation),
- a certifiable origin (provenance of the model and training data),
- a recognisable voice in the sense of recognisable communication patterns.

5.5.2 Ethics

Ethics is – following on from integrative business ethics (Ulrich 2008) and the concepts of an ethics of artificial intelligence (Mainzer 2019; Floridi 2023) – the embedding of normative standards within the agent's decision-making mechanisms. A distinction must be made between three levels:

- Compliance ethics: The agent complies with applicable law. This is the minimum requirement, not the maximum.
- Stakeholder ethics: The agent takes into account the legitimate interests of all relevant stakeholders, not just those of their client.
- Discourse ethics: The agent is prepared to justify their decisions and to engage in a discourse in which they too may be wrong. This ties in with Habermas (1981) and the Kantian imperative.

5.5.3 Personality

Personality comprises stable patterns of behaviour by which stakeholders recognise the agent. Personality is not a mere embellishment, but is of economic value: it reduces the transaction costs of cooperation because stakeholders know reliably where they stand. Personality comprises:

- communication style (formal, collegial, matter-of-fact, humorous),
- decision-making style (risk-averse, opportunistic, long-term),
- conflict behaviour (willing to compromise, tough, conciliatory),

- value profile (innovation, reliability, discretion, ...).

5.6 Parameter 5: AI model and embodiment

The AI model is the agent's 'engine'. The choice of model is a strategic decision with far-reaching economic consequences. Three dimensions must be specified:

Model type and model size

Foundation models (Bommasani et al. 2021) offer broad applicability, but are capital-intensive and create dependency on their providers. Specialised models are cheaper and more autonomous to operate, but have a narrower range of applications. The choice is always a make-or-buy decision – with the AGENTONOMICS-specific additional criterion of capability sovereignty (Knoll and Suckfüll 2024).

Local vs. centralised hosting

Where does the model run? On in-house servers, in a hyperscaler's cloud, or on European infrastructure? The answer has implications for latency, costs, data protection and – not least – geopolitical dependencies. Here, too, the question of sovereignty is fundamental.

Embodiment

- Embodiment refers to the form in which an agent is present in the world. There are three ideal-typical variants: Purely virtual embodiment: The agent exists as software in the cloud and communicates via text or voice interfaces.
- Hybrid embodiment: The agent controls physical actuators – printers, IoT devices, vehicle fleets – but remains within the digital sphere itself.
- Physical embodiment: The agent is integrated into a robot that acts physically and directly. This is the classic field of TUM robotics research.

Embodiment has economic consequences because it extends the scope of the agent's economic actions. A physically embodied agent can do things that a purely virtual one cannot – but the investment and maintenance costs are higher, as is legal liability.

5.7 Parameter 6: Learning and Teaching – The Meta-Agent

The sixth parameter describes the agent's ability to learn independently and – this is the crux of AGENTONOMICS – to train other agents. This is the concept of Homo sapiens sapiens: an intelligence that is not only intelligent, but knows that it is intelligent and can pass on this knowledge.

5.7.1 Learning Ability

Three learning modes can be distinguished:

- Pre-training: The extensive initial training on large datasets that constitutes the foundation model. This usually takes place at the model providers' premises, not at the agent operator's.

- Fine-tuning: The targeted adjustment of the model to application- and personality-specific data.
- In-context learning and tool use: Situational learning during use – for example, by the agent drawing on documents, databases or tools made available at runtime.

5.7.2 Teachability and sub-agents

The ability to teach is the specific hallmark of the meta-agent. An agent capable of teaching can train other agents – by generating training data, by defining sub-tasks, or by creating its own 'sub-agents' for specific tasks. The question of whether an agent is permitted to create its own sub-agents must be explicitly clarified in the definition framework – it has significant governance implications. Who is responsible for the sub-agents? Who owns the knowledge stored within them? How is control ensured?

Here, the parameter is directly linked to the discussion on skill sovereignty: if a human knowledge worker transfers their skills to an agent and that agent in turn creates sub-agents, it must be clear who owns the multiplied skills (Knoll and Suckfüll 2024).

5.8 Consistency check and board perspective

The six parameters are not independent of one another. A definition is only consistent if the assumptions are mutually compatible. Three typical inconsistencies are worth mentioning:

- An agent with autonomy level 4 but without an identity specification: Who is liable if it causes damage?
- An agent with a comprehensive claim to capability but a small, specialised model: the promise cannot be kept.
- An agent with a high degree of sovereignty but no regulatory framework: stakeholders will not trust it.

A key task of the AGENTONOMICS algorithm is to detect such inconsistencies and submit them to the human board for a decision. This consistency check is the AGENTONOMICS-specific operationalisation of what Bleicher (2011) called for as 'integration'.

The definition phase is a task for the board. All perspectives – CEO, CFO, CTO, CSO – must be involved collectively. The CEO is responsible for the vision and setting normative standards, the CFO for economic viability, the CTO for technical feasibility, and the CSO for stakeholder acceptance. If any of these perspectives is overlooked, the definition will be incomplete – and the agent will fail in its lifecycle.

> *The definition phase is not a tedious formality at the start of a project, but the most important decision of all. Anyone who cuts corners here is building on sand.*

Bibliography

Bleicher, K. (2011): The Concept of Integrated Management. 8th edition, 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): Regulation (EU) 2024/1689 of the European Parliament and of the Council laying down harmonised rules on artificial intelligence. Official Journal of the European Union.

Floridi, L. (2023): The Ethics of Artificial Intelligence. Principles, Challenges, and Opportunities. Oxford University Press, Oxford.

Fraunhofer MOEZ; Suckfüll, H. (2018): Initiative for a Data Trust (DEDATE). Discussion paper, Leipzig.

Habermas, J. (1981): Theory of Communicative Action. Suhrkamp, Frankfurt am Main.

Knoll, A. C.; Suckfüll, H. (2024): From data sovereignty to skill sovereignty. Technical Report TUM-I24112, Institute of Computer Science, Technical University of Munich.

Mainzer, K. (2019): Artificial Intelligence – When Will the Machines Take Over? 2nd edition, Springer, Berlin and 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): The St. Gallen Management Model. 3rd edition, 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 Business Ethics. 4th edition, Haupt, Bern.

6 Module 2: The Five Building Blocks of an AGENTONOMICS Agent

Whilst Chapter 5 developed the definition framework as an entrepreneurial premise, this chapter focuses on its concrete structural realisation. An AGENTONOMICS agent is not a monolithic software system, but a modular construction comprising five clearly defined building blocks. This modularity is not a didactic simplification, but a technical, economic and legal necessity: it enables the separate procurement, maintenance, evaluation and accountability of each building block. Anyone wishing to build an agent must first understand that they need to build, purchase or licence five components – and that the choice of each individual building block has economic consequences that extend far beyond the technical aspects.

In the translation of the St. Gallen Management Model into the AI agent economy, the five building blocks correspond to what Rüegg-Stürm and Grand (2020) describe as the 'processes' and 'structures' of the enterprise – with the crucial difference that these processes and structures no longer organise human employees, but are embodied in code, models and interfaces. What was an organisational chart in Rüegg-Stürm's model is an architectural diagram in AGENTONOMICS. What was a process model is now an algorithm. The substance of the St. Gallen school of thought remains intact; the form in which it manifests itself is fundamentally transformed.

6.1 An overview of the five building blocks

The five building blocks can be named as follows – without any claim to a definitive order:

1. Agent algorithm: the decision-making and orchestration logic.
2. Agent interface: the interfaces to the outside and inside.
3. AI model: the language or multimodal model as the cognitive engine.
4. Tools and vectors: the tools and stored knowledge.
5. Agent ecosystem intermediary: the connection to the network context.

Each block is represented by its own logo; together, the logos form a graphical signature of the agent. In the AGENTONOMICS standard, this structure is mandatory so that agents can cooperate with one another on the basis of open interfaces. An agent without a clear building-block structure is not AGENTONOMICS-compliant – even if it delivers functionally comparable performance. The standard is not merely a gesture, but the prerequisite for individual agents to form a truly interoperable, polycentrically organised network.

6.2 Building Block 1: The Agent Algorithm

The agent algorithm is the agent's decision-making and orchestration logic. It answers the question: What does the agent do in which situation? Unlike the AI model, which transforms inputs into outputs, the algorithm controls the process by which the model is queried – when, with which prompts, in what order, and with which tools. The algorithm is – to put it in the language of classical business administration – the agent's process organisation.

6.2.1 Functional components

A complete agent algorithm contains at least the following functional components, each of which requires its own design decisions:

Perception and Input Processing

How does the agent receive inputs, classify them and decide which processing paths to take? This involves routing, intent detection and the initial enrichment of the input with contextual information. A trivial but often underestimated question: what happens if the agent does not understand a request? Does it remain silent? Does it ask for clarification? Does it escalate the matter to a human?

Planning and task decomposition

How does the agent break down a complex task into subtasks? Modern agent architectures such as Reasoning-and-Acting (ReAct, Yao et al. 2023) or Plan-and-Execute combine the model with a loop of reflection and action. The algorithm determines how many iterations are permitted, when to terminate, and how to evaluate interim results.

Tool Invocation and Tool Use

What tools are available to the agent – database queries, web searches, computational operations, external APIs? How does it decide when to use which tool? How does it deal with errors and inconsistencies? Tool use is the key extension that distinguishes an agent from a pure language model (Schick et al. 2023). Without tools, the model remains trapped within the linguistic realm; with tools, it steps out into the world.

Memory and State Management

What information does the agent retain between interactions? Which details does it store in long-term memory, and which are forgotten? This raises a highly sensitive question, both economically and ethically: an agent with a perfect memory is powerful, but also risky – especially if it mixes information from different clients.

Self-reflection and quality assurance

How does the agent evaluate its own outputs? Does it use a second model invocation as a reviewer? Does it compare several candidate responses? Self-consistency, reflection and critique mechanisms are established techniques (Madaan et al. 2023), but each one costs computing time and money. The algorithm decides where this effort is worthwhile.

6.2.2 Make-or-Buy Decision

The agent algorithm is the only building block that is practically always the agent operator's own intellectual property. Whilst the model, interface and tools are often purchased, the algorithm is the opera-

tor's own creation – the specific interconnection of the building blocks in which the agent's value creation crystallises. From this perspective, the algorithm is the agent operator's actual capital. If they lose control of the algorithm, they lose control of the agent – and thus of its value creation.

Practical note: The algorithm as an asset

Anyone operating an AGENTONOMICS agent should document the algorithm, version it and treat it as an intangible asset. It belongs on the balance sheet – although German commercial law provides for an option to capitalise self-created intangible fixed assets, subject to certain exceptions. Economic reality is ahead of balance sheet law in this respect.

6.3 Building Block 2: The Agent Interface

The interface comprises all the points through which the agent communicates with its environment. From an economic perspective, it is the point at which value creation becomes visible, as this is where interactions with clients, stakeholders and other agents take place. Three classes of interface can be distinguished:

6.3.1 Human–Agent Interface

This is about direct interaction with human users. Language – whether written or spoken – is the dominant medium today, supplemented by graphical elements, dashboards and, increasingly, multimodal outputs. The design of this interface is not merely a question of UX, but one of management theory: anyone wishing to deploy the agent as a CEO substitute (see Chapter 8) must design an interface that is fit for a board meeting – with the appropriate language, the appropriate level of discretion, and the appropriate way of communicating errors. This is a different requirement to that of a customer service chatbot.

6.3.2 Agent-Agent Interface

If agents are to cooperate with one another, they need a protocol. This is where the AGENTONOMICS standard comes in: it defines how agents identify themselves, describe their capabilities, exchange tasks and return results. The interface here is not primarily graphical, but semantic: it involves the machine-readable description of services, prices and conditions. Current and upcoming developments such as the Model Context Protocol (MCP) and the Agent-to-Agent (A2A) standards form the technical foundation here; AGENTONOMICS supplements them with the economic layer.

6.3.3 Agent-System Interface

The agent must be integrated into existing systems – ERP, CRM, banking APIs, government portals. This is where its suitability for brownfield environments is determined: an agent that can only run in a greenfield environment is of no value to most real-world businesses. Interface work often accounts for 60 to 80 per cent of the implementation effort – a factor that is regularly underestimated in cost calculations.

6.4 Component 3: The AI Model

The AI model is the agent's cognitive engine. It transforms inputs into outputs, thereby delivering the performance that makes the agent appear 'intelligent'. The choice of model is by far the most expensive – both financially and strategically – decision when building an agent.

6.4.1 Model classes

Three ideal-type model classes are available, each with its own economic implications:

Frontier Foundation Models

The largest and most powerful models – namely *frontier models from proprietary providers, such as the GPT, Claude or Gemini families* – offer the broadest range of applications, but are expensive to operate and create a dependency on a small number of large providers. Those who use them have access to cutting-edge technology – but relinquish some of their sovereignty. From a European perspective, this is the most sensitive option in terms of sovereignty policy (Knoll and Suckfüll 2024).

Open-source foundation models

Models such as Llama or Mistral can be operated independently and adapted to specific needs. They are more cost-effective to run (once the initial investment in infrastructure has been made), but are usually a generation behind the frontier models in terms of several benchmark parameters. For many organisations, however, this generation is entirely sufficient – and the gain in sovereignty can be considerable.

Specialised and fine-tuned models

Small models trained for a specific domain can outperform the large generalist models in their field – at a fraction of the cost. This approach is likely to prevail in the AGENTONOMICS world because it best meets the requirements for skill sovereignty: anyone wishing to offer their individual capabilities as an agent needs a model that can perform precisely those capabilities – and not all the others.

6.4.2 Economic indicators

Four key figures are economically crucial for model selection and should be recorded in the AGENTONOMICS matrix (see Chapter 9):

- Cost per million tokens – the immediate variable cost base. These vary by a factor of 100 between different models, sometimes with barely perceptible differences in quality.
- Latency – the time from request to response. Crucial for interactive applications, irrelevant for batch processing.
- Context window – the amount of information the agent can process simultaneously. Today's leading models can handle millions of tokens, a feat previously considered unattainable.
- Sovereignty index – an AGENTONOMICS-specific metric proposed by the authors, which reflects how independently the agent can operate from the model provider. An open-source, locally hosted

model has a sovereignty index close to 100; a Frontier model consumed via a proprietary API has an index close to zero.

6.5 Building Block 4: Tools and Vectors

Tools and vectors are what distinguish the agent from a pure language model. Tools are the instruments with which the agent acts in the world; vectors are the organised knowledge on which it draws.

6.5.1 Tools: Instruments of Action

A tool is – from the agent's perspective – a function that it can call upon to achieve a specific effect. Examples:

- Database query: Querying internal data
- Web search: searching for up-to-date information
- Calculation: precise mathematical evaluations
- API call: Interaction with external services
- Code execution: generation and execution of programmes
- Document generation: Creation of reports, contracts and presentations

The choice of tools is a strategic decision. It determines the extent to which the agent intervenes in the real world – and thus also the liability and security risks that arise. An agent that can actually carry out bank transfers is a different economic and legal entity from one that can only suggest such transfers.

6.5.2 Vectors: stored knowledge

Here, 'vectors' do not refer to the mathematical structure itself, but to the practice of making knowledge available to the agent in the form of vectorised representations. Specifically, these are vector databases in which documents, knowledge artefacts or empirical data are stored in such a way that the agent can access them via semantic similarity search. This technique – which has become known as Retrieval-Augmented Generation (RAG) – is central to AGENTONOMICS for two reasons:

Knowledge sovereignty

The vectors constitute the agent's specific knowledge. Unlike the foundation model, which represents shared general knowledge, the vectors represent the agent's individual memory. Anyone operating an agent designed to apply the St. Gallen Management Model stores the relevant literature – Bleicher, Rüegg-Stürm, Ulrich, Suckfüll, Knoll – in the vectors, and thus the intellectual core of its value creation.

Timeliness

Whilst the Foundation Model has a knowledge cut-off date (often many months before application), the vectors can be updated on an ongoing basis. New studies, new market data and new legal requirements are incorporated without the need for time-consuming model training. The vectors are thus the element that keeps the agent alive.

For an agent designed to replace a CEO, the vectors are the critical resource. This is where the stored knowledge about the company, its markets, its contracts and its stakeholders resides. This knowledge must be kept complete, up to date and consistent – otherwise the agent will make CEO decisions based on outdated or incomplete information. See Chapter 8.

6.6 Building block 5: The ecosystem intermediary

The fifth building block is the agent's connection to the surrounding ecosystem of agents and stakeholders – the ecosystem intermediary. It is the connectivity mechanism through which the agent joins networks, identifies itself, concludes contracts and maintains value-creating relationships. As the intermediary is a concept in its own right, discussed in detail in Chapter 7, this chapter is limited to its role within the individual agent.

From the perspective of the individual agent, the intermediary is the interface between two worlds: the horizontal network of other agents and the vertical control hierarchy comprising clients, owners and regulatory bodies. The intermediary provides the agent with identity verification, reputation data, contract templates, dispute resolution mechanisms and access to shared resources such as knowledge bases, training data or computing capacity.

In practical terms, this means that when building an agent, consideration must be given from the outset to which intermediary it will be connected to. This choice has consequences similar to those of choosing a bank or a tax adviser when setting up a business. In the AGENTONOMICS standard, the intermediary interfaces are specified in an open manner; agents can, in theory, switch between intermediaries. In practice, however, data accumulation and relationship management give rise to lock-in effects that should not be underestimated.

6.7 The Interaction of the Building Blocks in the Lifecycle

The five building blocks are not static but evolve over the agent's lifecycle. The DBOT logic – Design, Build, Operate, Own, Transfer – structures this development and forms the vertical axis of the AGENTONOMICS matrix (Chapter 9).

Design

The five building blocks are conceived during the design phase. This results in architectural diagrams, interface specifications, model selection dossiers and tool inventories. A high percentage of design decisions are reversible, but any revisions in later phases become increasingly costly.

Build

In the Build phase, the building blocks are implemented or procured. This is where the largest investments (CAPEX) are made. Typical activities include: algorithm implementation, interface programming, model licensing or fine-tuning, vector database setup, and intermediary integration.

Operate

In the Operate phase, the agent is running in production. Running costs (OPEX) are the dominant factor: model usage, tool calls, maintenance, security updates, stakeholder communication. This is where the project's economic viability is determined.

Own

The Own phase represents the long-term asset perspective: the agent becomes a permanent component of the balance sheet. Who is the owner? What values are recognised on the balance sheet? How is depreciation calculated? This phase is still underdeveloped in current practice; AGENTONOMICS is actively driving it forward.

Transfer

Finally, the agent can be sold, shared, decommissioned or transferred to a different context. Here, too, the building blocks must be considered separately: some are transferable (algorithms, vectors), whilst others remain in their original context (model-specific adaptations, interface contracts).

The cross-table comprising five building blocks and five lifecycle phases yields 25 fields. It forms the backbone of the AGENTONOMICS matrix and – once populated with the relevant KPIs, CAPEX/OPEX assumptions and responsibilities – serves as the central governance document for an AGENTONOMICS project.

> *The five building blocks correspond to what were once the functional areas in a traditional organisation – except that here they are embodied in code, models and interfaces. If you do not separate them clearly, you will not build an agent, but a jumble.*

Bibliography

Bleicher, K. (2011): The Concept of Integrated Management. 8th edition, Campus, Frankfurt am Main.

Knoll, A. C.; Suckfüll, H. (2024): From data sovereignty to skill sovereignty. Technical Report TUM-I24112, Department of Computer Science, Technical University of Munich.

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): The St. Gallen Management Model. 3rd edition, 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: Synergising Reasoning and Acting in Language Models. International Conference on Learning Representations.

7 Module 3: The Network Intermediary

The third module focuses on a concept that is only tentatively addressed in the classical St. Gallen management model: the network intermediary. AGENTONOMICS clearly builds on St. Gallen thinking here by asking how a polycentrically organised network of autonomous agents can simultaneously develop cohesion and preserve individual sovereignty. The answer is: through an intermediary who does not rule but mediates – who does not plan but opens up spaces.

This construct is AGENTONOMICS's specific translation of what Rüegg-Stürm and Grand (2020) refer to as 'moments of order' into the reality of an AI-agent economy. What is achieved in a human-dominated organisation through culture, strategy and structures is achieved in the agent network by the intermediary – as a technically and economically constituted institution that fosters the polycentric interconnection.

7.1 Why an intermediary is needed

Polycentric networks are characterised by a fundamental tension: they seek to preserve the autonomy of their members whilst simultaneously delivering collective outcomes. This tension cannot be resolved through hierarchy – otherwise the network would no longer be polycentric, but monocentric. Nor can it be resolved through spontaneous organisation alone – otherwise networks would regularly descend into conflicts that no one can mediate. A third institution is needed to mediate between these two poles: the intermediary.

This argument draws on three strands of theory that converge in the AGENTONOMICS concept:

7.1.1 Transaction Cost Economics

Coase (1937) and Williamson (1985) have shown that markets and hierarchies vary in efficiency depending on the characteristics of transactions. Between these two poles lie hybrid forms – long-term contracts, joint ventures and networks. These are superior when transactions exhibit medium specificity, medium uncertainty and medium frequency. The AGENTONOMICS intermediary is the institutional form through which these hybrid transactions between agents are organised.

7.1.2 Theory of communicative action

Habermas (1981) distinguishes between strategic and communicative action. Strategic action aims at success, whilst communicative action aims at understanding. A society based solely on strategic action disintegrates; it requires spaces for communicative action in which claims to validity are tested through discourse. Applied to networks of agents: a network comprising agents acting purely strategically breaks down into opportunistic manoeuvres. The intermediary creates the space in which agents can communicatively negotiate their claims to validity – that is, reflect on facts, norms and sincerity, rather than merely seeking to maximise them.

7.1.3 Polycentric Governance

Ostrom (1990) demonstrated how polycentric governance structures can successfully manage common-pool resources – without a central hierarchy and without a pure market. Ostrom identified eight design principles, ranging from clear boundaries and collective decision-making procedures to graduated sanctions. The AGENTONOMICS intermediary translates these principles into the world of AI agents.

These three strands of theory provide the rationale: the intermediary is neither an administrative bureaucracy nor a platform monopoly, but a specific form of polycentric governance that facilitates hybrid transactions between sovereign agents and provides a space for communicative negotiation.

7.2 The property management analogy

A productive analogy for understanding the intermediary is provided by the owners' association of a property. The flat owners in an owners' association are sovereign owners of their respective flats. Together, they own the common property – the stairwell, roof, façade and heating system. They do not manage these communal affairs themselves, but delegate them to a property management company. The property management company does not rule over the owners; it implements their decisions, manages the common property and represents the association externally.

It is precisely in this role that the AGENTONOMICS intermediary should be understood. The agents remain sovereign economic entities. Whatever they cannot or do not wish to do themselves – identity registers, reputation systems, dispute resolution, shared knowledge bases, regulatory compliance – they delegate to the intermediary. The intermediary is accountable to them, not the other way round.

Definition: AGENTONOMICS intermediary

An AGENTONOMICS intermediary is an institutionalised mediating body within a polycentric network of agents which performs collective functions on behalf of and under the control of the network members – without infringing upon the sovereignty of the individual agents in their core business.

The property management analogy has a second, insightful point: a property management firm can be replaced. If the owners' association is dissatisfied with the service, it terminates the management contract and seeks a new one. This is how it must also be in the AGENTONOMICS network: the intermediary must not hold a monopoly position, but must compete for the agents' favour. Otherwise, there is a risk that the intermediary, which is meant to serve, will become a dominant monopoly – as we see today in the major platform economies (Zuboff 2018).

7.3 Functions of the intermediary

The specific functions performed by an intermediary depend on the context of the network. Six functional areas can be distinguished as ideal types:

7.3.1 Identity and reputation management

The intermediary maintains the agents' identity register. Each agent is registered here with their definition (see Chapter 5), their building blocks (see Chapter 6) and their designated representative. Over time, reputation accumulates: ratings from completed transactions, audit results and complaint histories. In the AGENTONOMICS network, reputation is an economic factor of the highest order – it reduces the transaction costs of cooperation and creates incentives for rule-compliant behaviour.

7.3.2 Contract and Transaction Infrastructure

The intermediary provides contract templates, smart contract templates and settlement mechanisms for typical transactions between agents. It performs fiduciary functions, secures payments and redistributes risks. Here, AGENTONOMICS ties in directly with the discussion on data trusteeship (Fraunhofer MOEZ and Suckfüll 2018; DEDATE 2.0) and with the tradition of German trust institutions, which played an important role in German economic history and are currently experiencing a renaissance.

7.3.3 Dispute resolution

Where agents interact, conflicts arise. The intermediary provides dispute resolution procedures that are faster, cheaper and more objective than state courts. Ideally, three stages can be distinguished:

- Mediation: The intermediary mediates between the disputing parties without decision-making authority.
- Arbitration: The intermediary makes a decision on the basis of arbitration rules to which the agents have agreed in advance.
- Escalation to state courts: In the case of constitutional issues or matters falling outside the scope of arbitration. The intermediary provides procedural support in such cases.

7.3.4 Pooling of knowledge and resources

Some resources are so valuable and, at the same time, so costly to acquire that no single agent can reasonably maintain them on their own. Examples include current legislation, market data, benchmark studies and training datasets. The intermediary makes these resources available as a shared infrastructure and finances them through membership fees – similar to library systems or trade associations in the traditional economy.

7.3.5 Regulatory Compliance

The EU AI Act (2024), the GDPR and sector-specific regulations impose requirements that are often prohibitively expensive for an individual agent to meet. The intermediary can bundle compliance functions: ongoing auditing, reporting to authorities, training, and early-warning systems for regulatory changes. In doing so, it assumes a role that has long been established within association structures – but now specifically for the new category of AI agents.

7.3.6 Marketplace function

Finally, the intermediary can provide a marketplace on which agents can offer and seek their services. This is where prices are determined, which will be the central coordinating factor in an AI capabilities economy. Important: The intermediary operates the marketplace but is not itself a market participant on it. This separation – between operating the marketplace and participating in the market – is fundamental to the intermediary's trustworthiness.

7.4 Governance of the intermediary

Who controls the intermediary? This question determines whether it functions as a service-oriented institution or mutates into a dominant platform. Three governance models are under discussion, each with different risk profiles.

7.4.1 Cooperative model

The intermediary is a co-operative of its members – the agents or their human or organisational representatives. Each agent has one vote, regardless of size or transaction volume. Advantages: strong alignment with members' interests, democratic legitimacy, low lock-in incentives. Disadvantages: potentially cumbersome decision-making; funding may prove insufficient.

7.4.2 Foundation and trust model

The intermediary is a special-purpose foundation or trust whose board is statutorily bound to act in the members' interests. Advantages: high stability, clear purpose, protection against hostile takeovers. Disadvantages: less direct member involvement; capacity for reform depends on the wording of the statutes.

7.4.3 Hybrid and federal models

Several competing intermediaries coexist and are linked via interoperable standards. Agents can switch intermediaries; standards prevent any loss of value in the process. This variant most closely resembles the property management analogy and is – from an AGENTONOMICS perspective – the preferred form in the medium term. It enables competition between intermediaries without drifting towards monopolistic platforms.

7.4.4 What must not happen

The history of the digital platform economy over the last twenty years provides a clear negative example: intermediaries that began as useful facilitators have become market-dominating monopolies whose economic power can now scarcely be curbed by democratic means. This path is blocked in the AGENTONOMICS standard by constitutive rules:

- Mandatory data portability – every agent can take their data and reputation with them.
- Interoperability as the standard – agents can switch between intermediaries.
- Separation of functions between marketplace operation and market participation.

- Transparency obligations towards members.

7.5 A Comparison of German and Anglo-Saxon Governance Styles

The institutional role of intermediaries differs across various economic cultures. Three models stand in contrast to one another – and the choice is not neutral, but reflects fundamental conceptions of the relationship between the economy, society and the state.

7.5.1 The German model

The German economic tradition encompasses a wide range of intermediary institutions: chambers of commerce, cooperatives, guilds, savings banks and public service broadcasters. These are often governed by public law or endowed with rights of self-governance. A German AGENTONOMICS intermediary would naturally fit into this tradition – with the associated strengths (stability, trust, subsidiarity) and weaknesses (bureaucracy, resistance to reform).

7.5.2 The Anglo-Saxon model

In the Anglo-Saxon world, private, often profit-oriented platforms dominate, and are kept in check by competition law and sector-specific regulation. The advantages are speed, innovative capacity and scalability. The disadvantages are the well-known tendencies towards concentration and the asymmetry between the platform and the user. An Anglo-Saxon AGENTONOMICS intermediary would likely start out as a VC-funded start-up and – if successful – become the dominant platform.

7.5.3 European hybrid model

The European variant combines elements of both models: a private legal form, but a platform logic that is regulated and constrained. The Digital Markets Act, the Data Act and the AI Act are expressions of this hybrid approach. For an AGENTONOMICS intermediary based in Europe, this means: organised as a private-sector entity, but with clear regulatory guardrails – data-sovereign, interoperable and open to competition. This is in line with the policy approach explicitly advocated by Knoll and Suckfüll (2024).

Several arguments in favour of anchoring the concrete implementation in Germany within a cooperative or fiduciary model. The German history of fiduciary stewardship (Suckfüll 1994, developed in a different context) and the German cooperative tradition (Raiffeisen, Schulze-Delitzsch) offer institutional models to which AGENTONOMICS can – and should – align itself.

> *The intermediary is the most important innovation that AGENTONOMICS introduces beyond the classical St. Gallen model. It is the institutional form in which polycentric networks can be both autonomous and coherent.*

Bibliography

Coase, R. H. (1937): The Nature of the Firm. In: Economica 4(16), pp. 386–405.

EU AI Act (2024): Regulation (EU) 2024/1689. Official Journal of the European Union.

Fraunhofer MOEZ; Suckfüll, H. (2018): Initiative for a data trust: DEDATE. Discussion paper, Leipzig.

Habermas, J. (1981): Theory of Communicative Action. Suhrkamp, Frankfurt am Main.

Knoll, A. C.; Suckfüll, H. (2024): From data sovereignty to skill sovereignty. Technical Report TUM-I24112, TUM, Munich.

Ostrom, E. (1990): Governing the Commons: The Evolution of Institutions for Collective Action. Cambridge University Press, Cambridge.

Rüegg-Stürm, J.; Grand, S. (2020): The St. Gallen Management Model. 3rd edition, Haupt, Bern.

Suckfüll, H. (1994): The Transnational Organisational Model. PhD thesis, University of St Gallen.

Williamson, O. E. (1985): The Economic Institutions of Capitalism. The Free Press, New York.

Zuboff, S. (2018): The Age of Surveillance Capitalism. Campus, Frankfurt am Main.

8 Module 4: The AI-CEO Management Concept

The fourth module takes us to the very heart of our work. Whilst the previous modules have defined, constructed and embedded the agent within a network, the focus now shifts to the question of how an AGENTONOMICS agent can assume the role of a Chief Executive Officer – not as an assistant, not as a preparatory tool, but as a decision-supporting or delegation-capable agent operating under human organisational responsibility. This statement is provocative, but not speculative: the technical, economic and governance-related (not the legal!) prerequisites are within reach today. This chapter shows how the transition should be conceived and shaped in concrete terms.

The AI-CEO Management Concept is the logical translation of the St. Gallen concept of integrated management (Bleicher 2011) into the age of AI agents. Whereas Bleicher situated integration cognitively within human top management, in the AI-CEO concept an AI agent takes on this integrative role – with a depth, speed and consistency that human management cannot achieve because it is subject to biological limitations. The human CEO does not become redundant; they become the owner, client and ethical check on the AI CEO. However, the operational burden of integrative management work can be transferred to the agent.

8.1 The St. Gallen Management Model in the Age of AI Agents

Since Ulrich (1968), the St. Gallen Management Model has gone through three major generations. The next generation, which we outline with the AI CEO concept, arises not by discarding but by building on previous thinking. Three shifts characterise the AI-related evolution of the St. Gallen Model:

8.1.1 From cognitive to algorithmic integration

Bleicher's (2011) central tenet was the integration of the normative, strategic and operational levels. In human management, this integration represents a cognitive feat of the highest order and, in practice, regularly fails due to capacity constraints. Any CEO of a medium-sized public limited company who wishes to keep track of all relevant strategic programmes, all compliance requirements, all stakeholder expectations, all market changes and all cultural implications simultaneously will simply reach the limits of human working memory. The result is the familiar set of consequences: inconsistency between public statements and actual practice, blind spots in the perception of stakeholders, delayed responses to weak signals, and strategies that fizzle out in day-to-day operations.

The AI CEO does not face this capacity limit. It can continuously access and cross-check hundreds of thousands of pages of internal documentation, external market data, regulatory texts and stakeholder communications. What Bleicher has formulated as an ideal – the consistent coherence of all management decisions – can thus, for the first time, be made technically operational.

8.1.2 From the communicative to the co-communicative enterprise

Rüegg-Stürm and Grand (2020) have emphasised the communicative dimension of management: organisations are networks of communication practices. In the AI version of the St. Gallen Model, a second

level of communication is introduced: communication between human actors and AI agents – and communication between agents themselves. A modern organisation is no longer merely a network of human communication, but a hybrid network of human and agent-based communication. This hybridity is not a shortcoming, but a resource: it multiplies the organisation's communicative capacity and enables negotiation processes at a speed and level of detail that would not be feasible through human interaction alone.

8.1.3 From human to hybrid meaning-making

A particularly significant shift concerns the normative level. Bleicher and Rüegg-Stürm have consistently identified questions of meaning – vision, values, responsibility – as a human task. AGENTONOMICS adheres to this view: the ultimate meaning of an enterprise does not stem from the model, but from the people who sustain the enterprise. However, the AI CEO can take charge of operationalising and implementing these normative principles. It checks all strategic and operational decisions for consistency with the normative framework and highlights any inconsistencies. In this way, the normative level gains in effectiveness without relinquishing its human foundation.

Core thesis

The AI CEO does not replace the human organ responsible for creating meaning within an organisation, but rather makes it effective. What the owners, supervisory board and workforce desire as normative principles is consistently translated into day-to-day business by the AI CEO – with a level of consistency that a human CEO, even with the same set of values, simply cannot achieve.

8.2 The five organising principles of the AI CEO

The AI CEO operates on the basis of five organising principles, which adapt the St. Gallen concept of organising principles (Rüegg-Stürm and Grand 2020) to the AI-agent economy. They form the coordinate system within which all decisions made by the AI CEO are situated:

1. Vision
2. Strategy
3. Structure
4. Agency (culture and behaviour)
5. Governance

Each aspect of organisation has a normative, a strategic and an operational layer – this effectively results in a five-by-three matrix with 15 fields. This matrix serves as the working framework for the AI CEO; every decision can be situated within it and checked for consistency with other decisions.

8.2.1 Vision

The vision answers the question: Why will this company exist in ten years' time? It is the long-term purpose against which all other decisions must be measured. The AI CEO manages the vision in a machine-readable form: indicators are derived from the prose of the vision, making the vision measurable. Every strategic and operational decision is assessed against these indicators. If the AI CEO detects an

inconsistency – for example, if a measure that is profitable in the short term undermines the long-term vision – it signals this to the human owner or supervisory board. The decision regarding the conflict remains a human one; the identification of the conflict is carried out by the AI.

8.2.2 Strategy

In the St. Gallen understanding (Gomez and Probst 1995), strategy is less a long-term plan than a conscious pattern of pattern recognition in complex environments. The AI CEO operationalises this pattern recognition. He monitors markets, competitors, technologies, regulatory developments and societal trends; he identifies patterns and anomalies; he develops options and evaluates them against the vision and values. Depending on the degree of autonomy (see Chapter 5), he can either prepare or make the strategic decision – that is, decide which option to pursue. At the highest level of autonomy, he makes strategic decisions independently within a framework set by the supervisory board.

8.2.3 Structure

Structures are the institutional frameworks within which work takes place: organisational structure, processes, reporting lines, and interfaces with suppliers and customers. In an AGENTONOMICS enterprise, further structural elements are added: the definition of all sub-agents deployed, their areas of responsibility, their escalation paths, and their inter-accounting relationships. The AI CEO manages this structure as a living model. It identifies structural bottlenecks, proposes adjustments and – in autonomous mode – implements them.

8.2.4 Agency: Culture and Behaviour

Agency encompasses the cultural patterns that shape the behaviour of the actors. In a hybrid organisation comprising humans and agents, this also includes the behaviour of the sub-agents: their communication styles, their conflict behaviour and their risk appetite. The AI CEO ensures that these patterns are consistent with the vision. He trains the sub-agents accordingly, corrects any drift, and keeps the culture vibrant. He cannot intervene directly on the human side – the culture of the workforce remains a human matter. But he can observe, describe, make suggestions and highlight inconsistencies.

8.2.5 Governance

Governance is the totality of the procedures by which the enterprise is controlled – by owners, the supervisory board, legislators and stakeholders. Here, the AI CEO is the addressee, not the actor: he must be controlled, not the one doing the controlling. The governance structures of an AGENTONOMICS- -based enterprise must be designed with particular care in the early years of its development. They must ensure that the AI-CEO remains true to its mandate, that avenues for intervention are maintained, and that responsibilities can be clearly assigned. This topic is explored in greater depth in Chapter 13.

8.3 What the AI CEO actually does

To translate the theoretical description into a practical concept, the scope of an AI-CEO's activities will be outlined in terms of specific functions. These functions are technically feasible today and make economic sense – even if they have not yet been bundled together in a single implementation.

8.3.1 Strategic Situation Analysis

Upon request, the AI CEO produces a comprehensive strategic assessment of the company's position: market position, competitors, strengths and weaknesses, opportunities and risks, and trends. In doing so, it draws on internal data (sales statistics, financial indicators, customer feedback) and external sources (market studies, business press, competitor reports). The time and effort required for such an analysis – which in a traditional management consultancy might amount to a six-figure sum and three months' work – is now reduced to a matter of minutes (!).

8.3.2 Decision preparation

Before every major decision, the AI CEO prepares a decision-making template: options, arguments for and against, risks, financial implications, stakeholder implications, and alignment with the vision and strategy. The document is summarised on a single page and can be expanded as required. A human board meeting, which today takes several hours to discuss an investment decision, can be conducted in a quarter of an hour using an AI CEO document – with a higher density of information and better preparation.

8.3.3 Stakeholder Communication

The AI CEO creates communications tailored to specific target groups, including employees, investors, customers, regulatory authorities and the general public. There is a high degree of consistency between the messages – the same spirit, different language. Press releases, annual reports, investor relations updates and staff newsletters are all drawn from a central core of knowledge and values. This not only reduces costs but also enhances the perceived integrity of the company.

8.3.4 Operational Management

In day-to-day operations, the AI CEO monitors key performance indicators, identifies deviations, and proposes or implements countermeasures – depending on the degree of autonomy. A modern PuK implementation (see Chapter 9) provides the data foundation for this. The AI CEO can operate at both quarterly and daily levels; they are aware of the liquidity position in the morning and the sales figures in the evening.

8.3.5 Crisis Management

The AI CEO is particularly valuable in crises. He keeps a cool head where people come under stress. He recalls all similar cases in the company's history and in the business press. He develops courses of

action quickly and consistently. Human decision-making in a crisis remains human – but it is made on the basis of significantly better information.

8.4 The Stages of Development of the AI CEO

The AI CEO does not emerge in a single leap, but through a sequence of development stages. These must be considered not only in technical terms, but also in terms of organisation and governance. Five stages can be distinguished – drawing on the autonomy levels from Chapter 5 in a didactically productive manner:

8.4.1 Level 1: AI Assistant

The AI CEO acts as an assistant to the human CEO. It prepares meetings, carries out research, summarises documents and drafts documents. All decisions are made by the human CEO. This stage has already been reached in most boardrooms of large companies today – often through generic tools such as ChatGPT, without this being recognised as an AGENTONOMICS implementation. Investment volume: low. Risks: low. Benefits: substantial, particularly in terms of time savings.

8.4.2 Stage 2: AI Chief of Staff

The AI CEO is a mature staff function, finely tuned to the company. It has access to all internal systems, understands the corporate culture and keeps an eye on the strategy. It not only produces summaries but also prepares complete decision-making documents with recommendations. The human CEO makes the decisions but can often follow the recommendations. This stage is currently achievable in pilot projects run by advanced companies. Investment volume: medium. Risks: primarily data protection, IP protection, and errors in the knowledge base.

8.4.3 Stage 3: AI Co-CEO

The AI CEO and the human CEO work as a team. The AI CEO makes routine decisions independently; strategic decisions are made jointly; matters of principle remain with the human CEO. This stage is currently achievable in selected pilot applications, but is not yet widespread. It requires a careful definition of decision-making powers – what can the AI CEO do autonomously, and what requires human involvement? Investment volume: high. Risks: gaps in accountability, regulatory uncertainties, cultural acceptance amongst the workforce.

8.4.4 Stage 4: Scenario: AI CEO with a human supervisory board

The AI CEO manages day-to-day operations independently. The human supervisory board – and, in the German context, particularly the co-determination bodies – sets the strategic framework, monitors compliance with regulatory requirements and decides on the composition of the executive board (i.e. on the AI- CEO itself). This stage is technically achievable today in specific, narrowly defined business areas – for example, in a purely online business model with a defined range of services. It will be adopted by an increasing number of companies over the next five to ten years, particularly among medium-sized

enterprises without a successor and start-ups with a clearly defined business model. Investment volume: very high. Risks: regulatory, legal (liability issues), cultural.

8.4.5 Stage 5: Long-term goal of a fully autonomous AI CEO in a polycentric network

The AI CEO operates completely autonomously, cooperates with other AI CEOs within a polycentric network, concludes contracts and independently adapts the business areas. Human involvement is limited to ownership rights and normative oversight by the intermediary. This stage is largely speculative at present and raises fundamental legal and ethical questions. It is likely to become achievable in niche areas over the coming decades – not as the norm, but

8.5 Application: How do you communicate with the AI CEO?

A crucial question for the practical use of the AI CEO is: how do you speak to it? What questions can be asked, and what answers can be expected? This question is not trivial. A human CEO is a social construct – we know how to address them, what we can expect, and how to deal with their responses. With the AI CEO, this social practice has yet to be developed. The following outlines typical categories of questions and their expected response patterns.

8.5.1 Strategic situational questions

Example question: "How do you assess our competitive position in business area X for the next twelve months?" The expected answer is typically divided into four parts:

- A concise main statement (two to four sentences) that summarises the AI CEO's assessment.
- A justification setting out the three to five key data points underpinning the assessment.
- An explicit statement of uncertainty: which data is reliable, where are the gaps, and which assumptions have been made.
- A recommendation for the next decision – not in the sense of an instruction, but as a basis for discussion.

A key characteristic is that the AI CEO makes uncertainty transparent. A human CEO tends – for career and status reasons – to convey certainty where none exists. A well-designed AI CEO does the opposite: it quantifies its uncertainty and thereby enables better risk decisions.

8.5.2 Operational queries

Example question: "Why did sales figures in the Southern Region fall by twelve per cent last month?" Sample response:

- Factual basis: The exact figures and the basis for comparison.
- Hypotheses: Three to five hypotheses regarding the causes, ranked by probability.
- Data reference: Which internal and external sources does each hypothesis draw on?
- Possible courses of action: What would be the next step to test the most likely hypothesis?

The AI CEO is not prone to hasty finger-pointing. When he mentions sales staff, he does so on the basis of facts and with the necessary nuance. He also recalls conversations from the previous month, the bonus structures and the competition. This ability to remember is one of the main advantages over human CEOs.

8.5.3 Normative reflection questions

Example question: "Is it compatible with our values to accept a major order from Customer Y, even though we know that working conditions there are problematic?" Sample answer:

- Reference to the company's explicitly documented values (mission statement, code of conduct).
- Identification of the conflict: Which of our values are in conflict, and in what way?
- Options: Acceptance, rejection, acceptance subject to conditions, negotiation with Client Y.
- Stakeholder perspectives: How would employees, investors, the public and the supervisory board react?
- Explicit invitation to make a human decision: "This question concerns the normative substance of our company. I request a meeting with the supervisory board."

In the case of normative issues, the AI CEO will never make a decision on its own authority. It highlights the complexity and delegates the decision to the human bearers of corporate responsibility. This self-restraint is part of its design.

8.5.4 Crisis situations

Example question – in the midst of an acute crisis: "We have received a public allegation of data misuse. What should we do?" Sample response:

- Immediate measures for the next hour: concrete, actionable steps.
- State of affairs: What do we know for certain, and what do we not know?
- Stakeholder triage: Who needs to be informed, and when? Supervisory board, staff, authorities, customers, the public.
- Medium term (next few days): Which investigations, which external advisers, which communication strategy.
- Lessons learnt – recommendations for the post-crisis period.

The strength of the AI CEO in a crisis lies in the speed and comprehensiveness of the stakeholder triage. What a human crisis team takes several hours to work out, the AI CEO delivers in minutes – with the option for human leadership to refine the results.

8.6 When the AI CEO says 'no'

A well-designed AI CEO explicitly says 'no' in certain situations – or, more precisely: 'I cannot and must not decide on this; I need a human authority.' This self-restraint is not a weakness, but a defining characteristic of an AGENTONOMICS-compliant AI CEO. Four situations are typical:

Normative-constitutive questions

Changes to the vision, values and constitution. These are a matter for the owners and the supervisory board.

Personnel decisions with significant consequences

The recruitment and dismissal of key staff, and promotions to management positions. In these matters, human judgement regarding people's futures takes precedence.

Exceptional risks

Decisions with potentially existential consequences for the company or third parties. Here, the dual-control principle is indispensable – involving at least two human pairs of eyes.

Conflict situations with stakeholder implications

When the AI CEO recognises that a decision infringes upon legitimate stakeholder interests in a manner not clearly authorised within the normative framework.

8.7 The AI CEO as a bridge to PuK logic

The AI-CEO concept on its own is a qualitative architecture. It requires a quantitative framework to anchor decisions with economic precision. This framework is provided by the Planning and Control (PuK) system (Hahn and Hungenberg 2001), which is discussed in detail in the next chapter. The AGENTONOMICS matrix, with its 25 fields comprising five building blocks and five lifecycle phases, serves as the joint working document for the AI-CEO and PuK. Without PuK, the AI-CEO would be an eloquent but economically vague adviser; with PuK, it becomes a fully-fledged economic decision-maker.

> *The AI-CEO does not replace the person at the helm of the company, but frees them from the cognitive burden of operational integration. What remains is what humans have always been good at: giving meaning and building trust.*

Bibliography

Bleicher, K. (2011): The Concept of Integrated Management. 8th edition, Campus, Frankfurt am Main.

Gomez, P.; Probst, G. (1995): The Practice of Holistic Problem Solving. Thinking in networks – acting entrepreneurially – convincing on a personal level. Haupt, Bern.

Hahn, D.; Hungenberg, H. (2001): PuK – Value-Oriented Controlling Concepts. 6th edition, Gabler, Wiesbaden.

Knoll, A. C.; Suckfüll, H. (2024): From Data Sovereignty to Skill Sovereignty. Technical Report TUM-I24112, TUM, Munich.

Rüegg-Stürm, J.; Grand, S. (2020): The St. Gallen Management Model. 3rd edition, Haupt, Bern.

Ulrich, H. (1968): The Organisation as a Productive Social System. Haupt, Bern.

Ulrich, P. (2008): Integrative Business Ethics. 4th edition, Haupt, Bern.

9 Module 5: The Planning and Control (PuK) Tool

The fifth module of the ADMRF is the Planning and Control (PuK) tool. It serves as the quantitative framework that anchors all the preceding modules within an economic context. Whilst Chapters 5 to 8 have described the qualitative architecture of the agent, the PuK tool provides the economic framework: It translates the definition, the building blocks, the intermediary link and the AI-CEO logic into comprehensible monetary values, planning figures and control routines. Without PuK, AGENTONOMICS remains a theoretical construct; with PuK, it becomes a model that can be operationalised in business management.

The PuK instrument ties in directly with the value-oriented controlling concepts of Hahn and Hungenberg (2001). This tradition is explicitly adopted here because it conceptualises the holistic integration of planning, control and management in a way that corresponds to the St. Gallen integration requirement – and because it simultaneously offers sufficient flexibility to be tailored to the specific requirements of an AI-agent economy. The financial fundamentals are based on Berk and DeMarzo (2020), in particular the chapters on capital budgeting, capital structure and corporate valuation.

9.1 The PuK tool in the tradition of Hahn and Hungenberg

Over the course of several years of research, Hahn and Hungenberg (2001) developed an integrated concept for planning and control in industrial enterprises, which has become widely adopted in German-speaking controlling departments. Three fundamental principles are particularly relevant to the AGENTONOMICS PuK:

9.1.1 Value orientation as a guiding principle

PuK is consistently value-oriented: the focus is on the long-term increase in enterprise value, not on short-term profit maximisation. This orientation is consistent with the logic of an AGENTONOMICS agent who – as an economic entity – is geared towards sustainable value creation throughout their entire lifecycle. Discounted cash flow methods, cost of capital and risk assessment, as described in detail by Berk and DeMarzo (2020), constitute the mathematical tools underpinning this value orientation.

9.1.2 Integrated planning

In the PuK framework, planning is not an isolated act of the finance department, but an integrated process that combines strategic, tactical and operational planning. Every strategic decision is immediately translated into financial terms; every operational measure is assessed against the strategic direction. This integration is what Bleicher (2011) formulated as a conceptual requirement – Hahn and Hungenberg have set it out in quantitative terms.

9.1.3 Closed-loop control

Planning without control is wishful thinking; control without planning is blind. PuK conceives of planning and control as a closed-loop system: Plan → Actual → Variance analysis → Adjustment measure → New plan. This closed-loop logic is related to Beer's (1972) cybernetics and makes PuK compatible

with the algorithmic control of an AI agent – for algorithms function structurally in exactly the same way: they observe, compare and correct.

9.2 The AGENTONOMICS Matrix as the Core Artefact

The central working document of the PuK tool is the AGENTONOMICS matrix. It was already introduced in Chapter 4; here, it is now filled in with specific details. The matrix intersects two dimensions:

- Horizontally: the five building blocks from Chapter 6.
- Vertically: the five lifecycle phases (DBOT logic): Design, Build, Operate, Own, Transfer.

This results in 25 cells. Each cell contains structured information from four categories:

1. Activities: What specifically is done with this building block during this phase?
2. Costs: What CAPEX and OPEX figures are involved?
3. KPIs: How are success and progress measured?
4. Risks and responsibilities: What could go wrong, and who is responsible?

The matrix is not static but is continuously updated. In the AGENTONOMICS algorithm, it serves as the steering document on which all modules are based. The AI CEO, CFO and CTO all access the same matrix – each from their own specific perspective, but based on a shared data set.

9.2.1 Example row: Design phase

To illustrate the matrix in concrete terms, let us examine an example row. In the design phase of an agent conceived as an AGENTONOMICS-compliant Doz. Dr. AGENTONOMICS, the fields might typically be populated as follows:

- Algorithm: Design of a ReAct-based orchestration system comprising the sub-tasks of research, synthesis and correction. Activity: Architectural design. Costs: one to three person-months of conceptual work. KPI: complete specification of all decision points. Risk: architectural flaws that will be costly to rectify in later phases.
- Interface: Specification of web, language and API interfaces. Activity: UX design, interface documentation. Cost: one to two person-months. KPI: acceptance rate in user tests. Risk: lack of acceptance amongst students.
- Model: Choice between a foundation model, an open-source model and fine-tuning. Activity: Benchmark study, cost-benefit analysis. Cost: one person-month plus initial model testing. KPI: Percentage of correct answers to a test question list. Risk: Vendor lock-in if a proprietary model is chosen.
- Tools and vectors: Building the vector database using AGENTONOMICS literature and lecture material; specification of tool interfaces. Activity: Data modelling, selection of embedding- . Costs: two to three person-months. KPI: Retrieval quality in a standardised benchmark. Risk: Knowledge gaps in the initial data loading.

- Intermediary: Specification of the connection to a TUM-internal intermediary; definition of the identity, reputation and dispute resolution interfaces. Activity: Governance specification. Costs: primarily conceptual. KPI: complete specification of all intermediary interfaces. Risk: standardisation gaps with other emerging AGENTONOMICS implementations.

Taken together, these five cells provide a precise picture of the design phase of a Doz. Dr. AGENTONOMICS – complete with a cost framework, performance metrics and risk assessments. Multiplied by the five lifecycle phases, this results in a comprehensive economic description of the agent across its entire lifecycle.

9.3 Balance sheet and profit and loss account of the agent

An economically significant innovation of the AGENTONOMICS PuK is the consistent application of accounting logic to individual agents. Each agent is modelled as an independent economic entity with its own balance sheet and profit and loss account. This approach is unusual today – software and development costs are accounted for depending on the form of acquisition, development phase and accounting regime – but it is economically accurate and indispensable for an AGENTONOMICS world.

9.3.1 Assets on the agent's balance sheet

The assets side comprises the agent's assets:

- Algorithm as an intangible asset: the specific code, the orchestration logic, and the company-specific adaptations.
- Models and fine-tuning investments: Proportionately capitalisable expenditure on the adaptation of foundation models.
- Vectors and knowledge bases: The specific body of knowledge that distinguishes the agent from others.
- Tool licences: Acquired rights of use for external tools and data sources.
- Reputational capital: Where this is documented in a measurable way and remains linked to the agent's identity (*not a separately capitalisable asset for balance sheet purposes*).

9.3.2 Liabilities side of the agent's balance sheet

The liabilities side comprises the sources of funding and obligations:

- Equity: Funds provided by the parent company or by external investors.
- Borrowed capital: Loans taken out specifically for the agent.
- Liabilities to suppliers: model providers, cloud providers, tool providers.
- Provisions: For regulatory risks, maintenance obligations and potential liability claims.

9.3.3 Profit and Loss Account

The agent's profit and loss account records its economic performance. On the income side are the revenues generated by the agent through its activities – whether through direct remuneration from clients,

internal charging to other business units, licence fees or transaction fees within the intermediary network. On the expenditure side are the operating costs: model usage, cloud infrastructure, tool calls, maintenance, intermediary fees, and the proportionate personnel costs of human supervision. The difference is the agent's profit – a figure that, in a mature AGENTONOMICS world, is measured just as naturally as the profit of a subsidiary is today.

Practical note

Keeping separate accounts for each agent may seem laborious at first. In fact, it is not: the data is collected anyway (cloud costs, model usage and tool calls can be measured on an agent-by-agent basis). The innovation lies in their consistent aggregation and analysis. Those who do not keep accounts for agents are flying blind.

9.4 CAPEX and OPEX – the capital and operating cost profile

The separation of capital expenditure and operating expenditure is crucial for the economic management of the agent. It differs structurally from traditional capital goods:

9.4.1 An agent's CAPEX profile

The initial investments are concentrated in the design and build phases. Typical CAPEX items include: architectural design, algorithm implementation, model fine-tuning, vector database construction and interface development. Unlike traditional capital goods (buildings, machinery), these CAPEX figures are heavily driven by personnel costs and are more difficult to forecast.

9.4.2 OPEX profile of an agent

Running costs remain consistently high and – this is the key economic point – cannot be reduced linearly through economies of scale. Every model call costs money, every tool call costs money, every hour of cloud usage costs money. These costs scale with usage. An agent that is used ten times as much costs around ten times as much to operate. This differs from traditional software, where marginal provisioning costs tend towards zero. However, self-hosting, reservations, caching, batching and model routing alter the cost curve.

9.4.3 Implications for business modelling

The CAPEX/OPEX profile has an important implication for business modelling: AGENTONOMICS agents are less scalable than traditional digital products, but generate greater value per transaction. The business model must reflect this characteristic – for example, through usage-based pricing, clearly defined service limits, or an architecture that sensibly combines expensive and inexpensive model calls.

9.5 Control Routines Throughout the Lifecycle

Control routines change throughout an agent's lifecycle – Design, Build, Operate, Own, Transfer. Different KPIs, reporting formats and escalation thresholds dominate in each phase. The PuK tool explicitly defines these phase-specific routines.

9.5.1 Design phase: Concept KPIs

In the design phase, there is as yet no market data, no usage statistics and no profit and loss statement. Control is exercised via concept KPIs: completeness of the specification, consistency with the vision and strategy, soundness of the business case assumptions, and plausibility of the risk analysis. The reporting format is the design dossier; escalation thresholds are gaps in content that would prove costly in later phases.

9.5.2 Build phase: Project KPIs

During the build phase, traditional project management takes the lead: deadlines, budget, delivery quality and risks. Reports take the form of a project status update using a traffic-light system; escalation thresholds are delays or budget overruns beyond defined tolerances.

9.5.3 Operate phase: business data

During day-to-day operations, financial KPIs take centre stage: turnover, margin, contribution margin, liquidity, customer satisfaction and market penetration. This is where the full range of the Hahn/Hungenberg framework comes into play. The reporting format is monthly; escalation thresholds are deviations from planned figures beyond defined ranges.

9.5.4 Own phase: Asset indicators

In the Own phase – when the agent is regarded as a permanent asset – asset indicators come to the fore: book value, market value, net asset value, capitalised value, depreciation, and investment requirements for maintenance and further development. This phase is still underdeveloped in current business practice; AGENTONOMICS is actively driving its development.

9.5.5 Transfer Phase: Transaction Valuation

Upon disposal, decommissioning or transfer of the agent to a different context, it undergoes a transaction valuation. Which building blocks are transferable? What value is attached to them? Which liabilities are transferred? This phase will gain significance in the emerging M&A practice involving AI agents.

9.6 The link between the CFO role and the AI CEO

The PuK instrument essentially covers the role performed by the Chief Financial Officer in traditional companies. Within the AGENTONOMICS framework, this role can be assumed either by a specialised AI CFO or by the integrated AI CEO. Both approaches have advantages and disadvantages:

Integrated AI-CEO

The AI-CEO assumes the PuK function themselves. Advantage: maximum consistency between strategy and financial planning. Disadvantage: lack of separation of functions, lower level of control.

Specialised AI CFO

A separate agent takes on the PuK function and maintains a productive dynamic with the AI CEO. Advantage: separation of functions, built-in dual-control mechanism. Disadvantage: coordination effort required between the agents.

The choice depends on the business model and the risk profile. In regulated sectors and for capital-intensive business models, the separation into two agents is preferable; in simpler set-ups, the integrated AI CEO with the PuK function is sufficient.

> *The PuK instrument forms the bridge between the qualitative architecture of the St. Gallen model and the harsh reality of the balance sheet. Anyone who does not manage an agent via the PuK is building it on a rhetorical foundation.*

Bibliography

Beer, S. (1972): Brain of the Firm. The Managerial Cybernetics of Organisation. Allen Lane, London.

Berk, J.; DeMarzo, P. (2020): Fundamentals of Finance. 6th edition, Pearson, Munich.

Bleicher, K. (2011): The Concept of Integrated Management. 8th edition, Campus, Frankfurt am Main.

Hahn, D.; Hungenberg, H. (2001): PuK – Value-Oriented Controlling Concepts. 6th edition, Gabler, Wiesbaden.

Knoll, A. C.; Suckfüll, H. (2024): From Data Sovereignty to Skill Sovereignty. Technical Report TUM-I24112, TUM, Munich.

Rüegg-Stürm, J.; Grand, S. (2020): The St. Gallen Management Model. 3rd edition, Haupt, Bern.

10 Module 6: The Stakeholder Communicator

The sixth and final module of the ADMRF deals with the agent's communication with its stakeholders. It translates what Rüegg-Stürm and Grand (2020) described as 'interaction themes' between the organisation and its stakeholders into the world of AI agents. The communicative shift in the fourth generation of the St. Gallen Model finds its concrete operationalisation here – with the crucial difference that an AGENTONOMICS agent can deliver communication services that would be unattainable for a human corporate communications department in terms of speed, differentiation and consistency.

The module is to be conceived as a Chief Sales Officer (CSO) – whereby 'sales' is not meant here in the narrow sense of the sales function, but in the broader sense of establishing and maintaining all relevant stakeholder relationships. It concerns the organisation's communicative interfaces, both externally and internally: investor relations, customer communications, employee communications, communications with regulatory authorities and public relations.

10.1 Stakeholders in the St. Gallen approach

From the outset, the St. Gallen model has interpreted the term 'stakeholder' broadly (Ulrich 1968; Rüegg-Stürm and Grand 2020). Stakeholders are all parties who have a legitimate interest in the company or are affected by its actions. This broad interpretation distinguishes the St. Gallen model from a purely shareholder-oriented view, which has become established in the Anglo-Saxon world since the 1990s. It also forms the basis for the AGENTONOMICS concept: an agent who is guided solely by the interests of their principal is not a fully-fledged economic entity in the St. Gallen sense, but rather an extension of their owners.

Ideally, six stakeholder groups can be distinguished:

- Owners and capital providers: investors, shareholders, banks.
- Employees: the company's workforce, including employee representative bodies.
- Customers: current and potential purchasers of the company's services.
- Suppliers and business partners: including other AGENTONOMICS agents and model providers.
- Government and regulatory authorities: tax authorities, supervisory bodies, data protection authorities.
- The public: the media, civil society, the local community, non-governmental organisations.

Each group speaks its own language, has its own information needs and responds to different communicative stimuli. The Stakeholder Communicator must manage this diversity – drawing on a consistent foundation of knowledge and values. This is precisely the specific strength of the AI agent compared to a human PR department.

10.2 The Pitch Methodology according to Kawasaki

A particularly effective methodology for stakeholder communication in business contexts is the pitch methodology developed by Guy Kawasaki. Kawasaki worked as an 'evangelist' at Apple in the early

1980s and, drawing on this experience, developed a pragmatic approach to persuasive business communication (Kawasaki 2015). The stakeholder communicator makes use of this methodology and automates its application.

10.2.1 The 10-20-30 rule

Kawasaki's best-known rule concerns investor presentations: ten slides, twenty minutes, thirty-point font size. The limit of ten slides forces one to focus on the essentials; twenty minutes corresponds to the attention span of a senior management team; a 30-point font size prevents the presentation from degenerating into a lecture. These three guidelines are trivial to implement algorithmically and immediately define the external structure of investor communication. The Stakeholder Communicator can use them as a standard template.

10.2.2 The ten slides in the standard pitch

Kawasaki specifies the ten slides of an investor presentation as follows:

1. Title: Company, slogan, contact details.
2. Problem: What pain points exist in the market?
3. Solution: How does the company address these pain points?
4. Business model: How does the company make money?
5. The magic behind the product: technology, IP, competitive advantage.
6. Marketing and sales: Go-to-market strategy.
7. Competition: Who are the competitors, and what sets us apart?
8. Team: Who is behind the company?
9. Projections and milestones: three to five years.
10. Status and Timeline: Where are we now, and what's next?

This structure has now become a de facto standard in the start-up world. The Stakeholder Communicator can generate it automatically from the data in the AGENTONOMICS matrix and the content of the AI-CEO concept – with consistent figures, consistent language and consistent visualisation.

10.3 From PuK to Pitch: The Operational Derivation

One of the Stakeholder Communicator's key functions is to translate quantitative metrics from the PuK tool into communication formats tailored to specific stakeholders. This translation is less straightforward than it sounds: the same information must be presented differently to an investor than to an employee, differently than to a regulatory authority, and differently than to the general public.

10.3.1 Consistency through differentiation

In doing so, the Stakeholder Communicator adheres to a key requirement: all communication is based on a shared set of facts and a shared set of values. Differentiation lies in language, level of detail and emphasis, not in the substance. An investor receives different key figures to an employee, but the message is the same. A letter to a regulatory authority sounds different from a press release, but the facts are identical. This consistency in differentiation is the operational crux of the module.

10.3.2 Typical outputs

The Stakeholder Communicator typically generates:

- One-pagers for initial contact with investors or partners.
- Investor decks based on the Kawasaki methodology.
- Annual reports and quarterly reports.
- Press releases and Q&A documents.
- Internal staff newsletters.
- Customer communications regarding product changes.
- Applications to authorities and regulatory reporting.
- Crisis communication drafts.

Each of these outputs follows its own logic, but they all draw on the same data source: the AGENTONOMICS matrix and the organising principles of the AI CEO.

10.4 Crisis communication as a stress test for the module

The real stress test for any stakeholder communicator is a crisis. In a crisis, a company's communication capabilities come under extreme pressure: time is short, information is incomplete, and many stakeholders are vying for attention simultaneously. This is where it becomes clear whether the module fulfils its purpose.

Four characteristics define a good crisis communicator:

Speed without sloppiness

The first press release must be ready within hours, without compromising subsequent statements. The stakeholder communicator drafts such statements in minutes and provides the human spokesperson with several options.

Tailoring to the audience

Employees, customers, authorities, the media – they all want to be addressed differently. The Stakeholder Communicator produces the various versions in parallel and ensures their consistency.

Adherence to values

In a crisis, there is a tendency to forget the values that guide us in calmer times. The Stakeholder Communicator reminds us of the company's normative framework (drawn from the Definition Frame and the AI CEO concept) and checks every piece of communication against it.

Preparing the next steps

Crisis communication is never merely reactive. It paves the way for the next steps: investigations, consequences, redress, structural changes. The Stakeholder Communicator maps out this path in advance

and ensures that the initial communication does not preclude the options available in the weeks that follow.

10.5 Internal and external communication – striking the right balance

A particularly sensitive task of the module is striking a balance between internal and external communication. Employees must not learn from the press what they should have been told first; investors must not be the last to be informed when they bear the greatest risk; and regulatory authorities must not be taken by surprise by leaks. The Stakeholder Communicator coordinates these sequences and ensures that the timing is appropriate for each stakeholder.

In the German context, particular attention must be paid to the principles of co-determination. Works councils have rights to information and consultation that take precedence over external communication. A correct implementation of AGENTONOMICS automatically takes this sequence into account – not as a courtesy, but for compliance purposes.

> *A company that fails to communicate consistently across all stakeholders loses trust both internally and externally. The Stakeholder Communicator is AGENTONOMICS' response to managing the complexity of the communications revolution.*

Bibliography

Habermas, J. (1981): Theory of Communicative Action. Suhrkamp, Frankfurt am Main.

Hahn, D.; Hungenberg, H. (2001): PuK – Value-Oriented Controlling Concepts. 6th edition, Gabler, Wiesbaden.

Kawasaki, G. (2015): The Art of the Start 2.0. Portfolio, New York.

Rüegg-Stürm, J.; Grand, S. (2020): The St. Gallen Management Model. 3rd edition, Haupt, Bern.

Ulrich, H. (1968): The Organisation as a Productive Social System. Haupt, Bern.

11 Growth Theory of the AI-Agent Economy

In this chapter, the handbook moves beyond the module level to address more fundamental questions. How does an AGENTONOMICS economy grow? What mechanisms drive it, which ones constrain it, and what risks are inherent in it? Growth theory is not merely an academic addition, but a practical necessity: anyone who takes AI agents seriously as economic agents must understand the macroeconomic consequences of their interactions.

11.1 The Cooperation Multiplier

Classical growth theory distinguishes between extensive and intensive growth (Solow 1956). Extensive growth arises from an increase in the factors of production – more capital, more labour, more land. Intensive growth arises from a more productive use of these factors – technological progress, education, organisation. AGENTONOMICS adds a third dimension to this classical picture: growth through cooperation in polycentric agent networks.

11.1.1 Mechanism of action

When agents cooperate within a polycentric network, added value is created that exceeds the sum of their individual contributions. This theory is not new – network effects have been an established concept since Shapiro and Varian (1999). What is new is the intensity with which these effects can operate in an AI-agent economy. There are three reasons for this:

- Speed of negotiation: Agents can negotiate in milliseconds what takes days between humans.
- Granularity of the transaction: Agents can exchange minute partial contributions in an economically meaningful way, whereas between humans such contributions are swallowed up by transaction costs.
- Consistency of cooperation: Agents do not forget agreements; they do not have ‘bad days’; they do not negotiate out of vanity.

These three characteristics give rise to a cooperation multiplier: the output of a network of n agents can be considerably higher than n times the output of a single agent. This multiplier is the true promise of growth offered by AGENTONOMICS.

11.1.2 Limits of the multiplier

The multiplier is not unlimited. Three limits must be taken into account:

Complexity limit

As the number of agents increases, the complexity of coordination grows quadratically. Above a certain size, coordination becomes so costly that the additional benefit diminishes. The intermediary (Chapter 7) can push back this limit, but cannot remove it altogether.

Trust limit

Cooperation requires trust. As the number of agents increases, trust becomes more difficult to establish. Reputation systems and dispute resolution procedures can help, but cannot fully compensate for this.

Meaning threshold

Cooperation must serve a purpose. If agents cooperate merely for the sake of cooperating, no value is created – only busyness. This is where the normative framework from the AI-CEO module comes into play: it defines which forms of cooperation are meaningful and which are futile.

11.2 Triadic analysis: Three forms of growth

Drawing on Suckfüll's analysis of transnational corporations (Suckfüll 1994), three forms of growth can be distinguished in AGENTONOMICS networks:

11.2.1 Vertical integration through sub-agents

An agent generates sub-agents for specific subtasks (see Chapter 5, Parameter 6). These sub-agents extend the depth of value creation of the main agent. Growth arises from the expansion of internal functions, not from the expansion of the market. Economically, this corresponds to vertical integration in the classical sense.

11.2.2 Horizontal networking with peer agents

Several agents belonging to different owners cooperate within an intermediary network. Each retains its autonomy, but they exchange services and create value together. Growth arises through the expansion of the network and the deepening of relationships. Economically, this corresponds to platform growth without a central platform.

11.2.3 Diagonal specialisation

Agents are developing increasingly specialised skill sets which they can only utilise in cooperation with complementary agents. Growth arises from the deepening of the division of labour. Economically, this corresponds to the classical Smithian growth mechanism arising from the increase in the division of labour – only at a new level of actors.

These three forms of growth are not mutually exclusive but complementary. A mature AGENTONOMICS economy will see all three occurring in parallel – and the cooperation multiplier has the greatest effect where they reinforce one another.

11.3 Polycentric network organisations as drivers of growth

The key macroeconomic consequence of these forms of growth is the erosion of the traditional corporate boundary. In an AGENTONOMICS world, the boundaries between firms are becoming increasingly blurred. Agents cooperate across corporate boundaries at a level of granularity that previously required

organisational integration. The result: the typical organisational form of the future is no longer the integrated enterprise, but the polycentric network.

This diagnosis is not new – Picot, Reichwald and Wigand (2020) formulated it long ago in their concept of the 'boundaryless enterprise'. What is new is that AGENTONOMICS provides the infrastructural prerequisites under which this form can become a reality beyond pilot projects.

11.3.1 Implications for the growth pattern

Polycentric networks grow differently from integrated companies. Three characteristics are distinctive:

- Growth occurs through the connection of new nodes, not through the expansion of existing nodes.
- The nodes remain small and agile; the network as a whole becomes large.
- Value creation is widely distributed, not concentrated in a few nodes.

This growth pattern – if it takes hold – offers an alternative to the concentration dynamics of today's platform economy. It is one of the central hopes associated with AGENTONOMICS: that the next wave of digital transformation will not necessarily result in monopolistic concentration, but can instead be distributed in a polycentric manner.

11.4 Innovation and renewal in the network

Rüegg-Stürm and Grand (2020) distinguish between two modes of development: optimisation and renewal. This distinction is also relevant to AGENTONOMICS networks. Optimisation is the continuous improvement of existing routines; renewal is the breaking with routines and the establishment of new patterns.

In the Schumpeterian sense, renewal is innovation – 'creative destruction' (Schumpeter 1934), which breaks down old market structures and creates new ones. In an AGENTONOMICS economy, innovation manifests itself on two levels:

11.4.1 Innovation in Agents

Better algorithms, more powerful models, richer knowledge bases – continuous innovation takes place at the level of individual agents. Provided the agent is well-designed, this innovation is built into its build-operate cycles.

11.4.2 Innovation in networks

New patterns of cooperation, new business models, new intermediary structures – at the level of networks, innovation emerges as an emergent phenomenon. Individual agents invent a new pattern of cooperation; others adopt it; the network transforms.

This second level of innovation has scarcely been researched to date and is still largely invisible in practice. AGENTONOMICS offers a conceptual preview here: if polycentric agent networks do indeed emerge, network innovation will become one of the most important sources of growth in the coming decades.

The growth of an AGENTONOMICS economy is not the growth of large nodes, but the growth of dense networks. This is not a mere detail, but a political statement.

Bibliography

Knoll, A. C.; Suckfüll, H. (2024): From data sovereignty to skill sovereignty. Technical Report TUM-I24112, TUM, Munich.

Picot, A.; Reichwald, R.; Wigand, R. T. (2020): The Boundless Enterprise. 6th edition, Springer Gabler, Wiesbaden.

Rochet, J.-C.; Tirole, J. (2003): Platform Competition in Two-Sided Markets. In: Journal of the European Economic Association 1(4), pp. 990–1029.

Rüegg-Stürm, J.; Grand, S. (2020): The St. Gallen Management Model. 3rd edition, 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), pp. 65–94.

Suckfüll, H. (1994): The transnational organisational model. PhD thesis, University of St Gallen.

12 Economic Approaches: The Adolescence of AI

If AGENTONOMICS becomes established, it will have macroeconomic and societal consequences that extend far beyond individual corporate decisions. This chapter outlines the most significant effects and discusses economic and political approaches to addressing the tensions that are likely to arise. The argument follows closely on from the analysis by Knoll and Suckfüll (2024) in the TUM report TUM-I24112.

12.1 The adolescence of AI

Knoll and Suckfüll (2024) refer to the 'adolescence of AI'. The metaphor is apt: today's AI is no longer a child, but not yet an adult either. It is powerful but immature; capable but unreliable; widespread but not stably integrated. During this phase – which the authors estimate will last around ten to fifteen years – disruptions are to be expected, which the social and economic framework must cushion.

The key economic expectation for a full AGENTONOMICS scenario is as follows: we will see economic growth of around 20 per cent per year, but at the same time an exposure of around 20 per cent. These two figures are not fixed arithmetically, but are indicative – they are intended to illustrate the scale of the disruptions to be expected. Both figures are historically exceptional and both pose challenges to conventional regulatory logic.

12.2 20 per cent growth: sources and consequences

Annual growth of 20 per cent would be ten times higher than what OECD countries have achieved in recent decades. It would be comparable to the fastest growth China experienced during its boom years – but starting from a much higher baseline. Three sources can be identified:

12.2.1 Productivity leaps in knowledge work

Knowledge work – research, analysis, synthesis, communication – accounts for the lion's share of work in developed economies. If AI agents can increase productivity in these areas by a factor of two to five – and there is much to suggest they can – this will result in a growth spurt of historic proportions.

12.2.2 New areas of value creation

AGENTONOMICS agents enable business models that would not be economically viable with human labour alone – such as small-scale consultancy services, personalised educational programmes, high-frequency market analyses and round-the-clock availability. These areas generate growth without displacing existing jobs.

12.2.3 Reduction of Frictional Losses

A significant proportion of today's economic activity consists of friction losses: search costs, initiation costs, negotiation costs and monitoring costs. AGENTONOMICS agents can drastically reduce these friction losses – and the effort thus saved is redirected into productive activity.

12.3 Pessimistic forecast: 20 per cent structural unemployment – sources and consequences

At the same time, high structural unemployment may arise during the transition phase, the scale of which is historically comparable only to the Great Depression of the 1930s. This expectation is not deterministic – it depends crucially on how the economic policy framework is shaped. But it is plausible enough to address it openly, rather than dismissing it.

12.3.1 Jobs at risk

Three categories of work are particularly at risk:

- Standardised knowledge work: bookkeeping, applying the law in routine cases, translations, simple report writing.
- Middle management tasks: reporting, coordination, reporting cascades, standard analyses.
- Consultancy and service activities: research, initial contact, standardised consultancy processes.

In all these fields, AI agents can already – or will be able to in the near future – deliver performance that is at least comparable to standard human performance – and at a fraction of the cost. The consequence is not necessarily unemployment, but it is structural change on a massive scale.

12.3.2 The Luddite comparison

A historical analogy springs to mind: the Luddite movement of the early 19th century. English textile workers smashed looms that were replacing their manual labour. They lost the battle – the machines arrived, and the workers had to adapt. However, they had a valid point. Their working world was indeed destroyed; adaptation took generations; and the social upheaval was considerable.

Today, a similar shock looms, but with three key differences:

- The shock is not affecting manual labour, but knowledge work – in other words, precisely those groups who until now felt safe from automation.
- The shock is not unfolding over generations, but within a few years – the adjustment period is significantly shorter.
- The shock is not geographically concentrated (as it was back then in Yorkshire and Lancashire), but is affecting all developed economies simultaneously.

Anyone who today dismisses scepticism about AI as a rejection of modernity overlooks the fact that the Luddites had a legitimate concern – they did not want to halt progress, but wanted to have a say in it. That is precisely the question we face today.

12.4 Issues of distribution: Who benefits?

The economically crucial question is not the existence of structural change, but its distribution. If growth is indeed 20 per cent and unemployment is also 20 per cent, the question of distribution becomes acute.

Without regulatory control, there is a risk of a scenario in which a few owners of AI infrastructure become much richer, whilst broad sections of society lose their means of livelihood.

12.4.1 Purchasing power as a systemic issue

One particularly striking economic consequence is this: a capitalist economy needs purchasing power to function. If the masses lose purchasing power because there is no longer any demand for their labour, the demand side of the economy collapses – and with it the business model of the owners of AI infrastructure. Knoll and Suckfüll (2024) put this argument sharply: centralised AI dominance carries the risk that a lack of competition will make the distribution of wealth so asymmetrical that the resulting lack of purchasing power will undermine the entire capitalist system as we understand it.

To put it another way: a world with massive AI-induced unemployment not only destroys livelihoods, it also destroys the business model itself. This self-contradiction is an economic argument in favour of active regulatory intervention.

12.5 Proposed solutions

What potential solutions are available? Four main approaches can be identified, which should be viewed not as alternatives but as complementary:

12.5.1 Skill sovereignty as an individual asset

The key recommendation from Knoll and Suckfüll (2024) is that every individual must be given the opportunity to replicate their skills in AI agents and to benefit economically from doing so. Anyone who transfers their skills to an agent should remain the owner of that agent – and share in the value it creates. This framework combines individual sovereignty with economic participation. It is the specifically European response to the challenge.

12.5.2 Polycentric market structure

A regulatory framework that promotes polycentric AGENTONOMICS networks and prevents monopolistic concentration is essential. This is where the links lie with competition law, the Data Protection Act and the AI Act. Consistent enforcement of data portability, interoperability standards and the separation of functions between platforms and market participants is the minimum requirement.

12.5.3 Education and retraining

If 20 per cent of jobs are to be replaced, education and retraining systems must become more effective. The traditional view of in-service training – ‘train once, done’ – is now definitively outdated. Lifelong learning must become an economic given, financed and institutionally supported. In the German context, this is where the Federal Employment Agency comes in – and it is precisely this agency that itself faces the challenge of staff reductions due to AI (see Chapter 3).

12.5.4 Social security systems

Finally, social security systems must be reviewed. A baseline unemployment rate of 20 per cent is difficult to reconcile with traditional, contribution-based insurance schemes. Concepts such as an unconditional basic income, a tax on machine labour or citizen participation in state-held shares in AI infrastructure are being discussed – albeit controversially. Which path is taken is a political question – but the fact that some path must be taken is an economic necessity.

12.6 The specifically European position

Europe occupies a unique position in the global AI debate. On the one hand, there is the American platform economy with its dynamics of concentration; on the other, the state-directed Chinese AI strategy with its tendency towards authoritarian use. Europe has the opportunity – and the duty – to pursue a third way: market-based, yet polycentric; liberal, yet with reliable regulation; open to technology, yet focused on individual rights.

AGENTONOMICS is a conceptual elaboration of this third way. It follows in the tradition of ordoliberal economics, which considers competition and regulatory frameworks as a unified whole; it follows in the tradition of the social market economy, which combines economic performance with social participation; and it follows in the tradition of the European Enlightenment, which regards individual sovereignty as non-negotiable.

> *The adolescence of AI is proving economically and socially challenging. But it can be shaped – provided that this shaping begins early and is pursued consistently.*

Bibliography

EU AI Act (2024): Regulation (EU) 2024/1689. Official Journal of the European Union.

Knoll, A. C.; Suckfüll, H. (2024): From data sovereignty to skill sovereignty. Technical Report TUM-I24112, TUM, Munich.

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 Business Ethics. 4th edition, Haupt, Bern.

Zuboff, S. (2018): The Age of Surveillance Capitalism. Campus, Frankfurt am Main.

13 Sovereignty: From Data Sovereignty to Capability Sovereignty

The question of sovereignty forms the normative foundation of the AGENTONOMICS approach. It was introduced in Chapter 3; this chapter explores it in greater depth and links it to the central thesis of the TUM report 'From data sovereignty to skill sovereignty' (Knoll and Suckfüll 2024). The thesis is sharply formulated: anyone who loses sovereignty over their own data and skills in the AI skills economy loses not only economic substance but also the prerequisite for individual self-determination. This thesis combines economic analysis with a normative claim – and at the same time makes this chapter both the theoretical framework and the political punchline of the entire handbook.

In the translation of the St. Gallen Management Model into the age of AI agents, the question of sovereignty plays a constitutive role. Rüegg-Stürm and Grand (2020) have emphasised the embedding of the enterprise in environmental spheres; AGENTONOMICS refines this embedding by asking which sovereignty rights over data and capabilities are actually enforceable in a polycentric AI economy – and how a politically and economically viable architecture for the coming decades can be derived from this.

13.1 The bizarre status quo

Knoll and Suckfüll (2024) describe the current status quo as 'fundamentally absurd'. This assessment merits a detailed account, as it precisely outlines the normative starting point of AGENTONOMICS.

13.1.1 Data protection and data expropriation in the same act

On the one hand, data protection legislation – particularly in the German interpretation of the GDPR – hinders social value creation in many areas. Hospital chains fail in their data integration projects; universities lose research opportunities; and small and medium-sized enterprises forego efficiency gains because they shy away from the legal risk. On the other hand, every single individual must disclose an enormous amount of personal data in their everyday lives simply to be able to participate in a modern, digitally mediated life at all – and receives little or nothing in return. This asymmetry is no coincidence, but the result of a regulatory framework that rhetorically defends individual sovereignty but does not enable it in practice.

13.1.2 The economic scale

The actual extent of this asymmetry is often underestimated. Knoll and Suckfüll (2024) estimate that an average adult in a developed country produces several hundred gigabytes of personal data per year – ranging from LinkedIn profiles and online job applications to digital invoices, contracts and communication records. The economic value of this data is estimated at between ten thousand and fifteen thousand euros per person per year. Virtually the entire sum now flows to the major platform economies; nothing remains with the individuals who produce the data.

This scale is not insignificant in macroeconomic terms. With around 700 million adults in OECD countries and an average data value of twelve thousand euros per person per year, this results in an annual

volume in the region of eight trillion euros – a flow of value creation that currently bypasses the balance sheets of its actual producers.

13.1.3 Data minimisation as a 'path to impoverishment'

The principle of data minimisation, which is particularly prominent in Germany, is well-intentioned, but ultimately – as Knoll and Suckfüll (2024) put it so succinctly – "a direct path to impoverishment". The statement is provocative, but accurate: those who are not allowed to derive economic value from their data stream are not giving that value away – they are merely depriving themselves of it. The data is generated anyway, it is analysed anyway, and it creates value anyway. The only question is who reaps that value.

Core thesis (Knoll and Suckfüll 2024)

The state's actual role should be to provide every individual with a comprehensive set of tools enabling them to gain sovereignty over their own data and decide for themselves what happens to it. Even if the legislative and operational framework for such individual data sovereignty is not yet in place, it is not too late to create it.

13.2 The next stage: Skill Sovereignty

The real crux of the TUM report lies not in its assessment of data sovereignty, but in the observation that the next sovereignty issue is already on the agenda: that of skill sovereignty. Whilst policymakers are still debating how individual data sovereignty might be made operational, the next transition is already underway – with potentially far-reaching consequences for the fundamental economic order.

13.2.1 How skills migrate

Every knowledge worker – whether an engineer, lawyer, doctor or university lecturer – processes input data or information according to rules that they have acquired and refined through their education and professional experience. These rules form the individual 'modus operandi' of every knowledge worker. They are the very essence of what knowledge work is paid for. However – and this is the technological innovation – they are increasingly transferable to AI systems. The mechanisms of 'skill transfer' are becoming more efficient with every new generation of models; what was unthinkable five years ago is now standard practice.

The economic implications of this transferability are far-reaching. If one's own modus operandi can be technically replicated, the question arises as to who owns this replication. The employer, because they pay the salary? The model provider, because their model enables the replication? The knowledge worker themselves, because their modus operandi is the result of their life's work? This question is currently unresolved in legal terms, highly relevant in economic terms and fundamental in ethical terms.

13.2.2 The 'functional replica'

Knoll and Suckfüll (2024) coin the term 'functional replica' for an agent trained to replace an individual knowledge worker. The replica can be deployed in several locations simultaneously; it can work for

several clients at the same time; it potentially multiplies the labour capacity of its 'original' by orders of magnitude.

If this transition goes wrong, knowledge workers will be worse off than the 19th-century Luddites – for unlike the English weavers, they have not even lost their physical loom; they have lost their cognitive identity. If, on the other hand, it succeeds, everyone can multiply their labour and thus participate directly in the productivity surge of the AI economy.

13.2.3 Copyright over one's own replica

This argument leads to a regulatory consequence that the TUM report clearly draws: anyone who transfers their abilities to a machine must remain in control of those abilities – even if they are multiplied and deployed in parallel. In concrete terms, this means that whilst a person may work for several clients at the same time, those clients cannot use the 'functional replicants' independently. The individual retains the 'copyright', that is, full control over their entire stock of replicants. This is the legal framework that makes Skill Sovereignty operational.

The analogy with copyright is a useful one. Just as an author retains the right to dispose of her novel even when she licences a print run to a publisher, so the knowledge worker retains the right to dispose of her agent when she licences a usage option to a client. This arrangement is legally complex, but not impossible – the collecting societies of traditional copyright law (GEMA, VG Wort) provide institutional models.

13.3 The Volkswagen analogy

An instructive analogy cited by Knoll and Suckfüll (2024) is the historical comparison with the Volkswagen model. In the late 1930s, the VW Beetle was conceived as the 'Volkswagen' – a car that every adult should be able to afford. The car was intended as a means of individual transport that democratised mobility. The idea took hold – following the upheavals of the Second World War – in the 1950s and 1960s and fundamentally transformed the economy and society.

The parallel with AI is clear: just as the private car was for physical mobility, individual AI agents could be for cognitive sovereignty. The widespread availability of powerful tools – not in the hands of a few corporations, but in the hands of the many – is changing patterns of value creation and balances of power. Today, we no longer speak of the 'Volkswagen' as a revolutionary idea, because it has become a matter of course. In the same way, if AGENTONOMICS becomes established, the individual 'people's agent' will also become a matter of course.

13.3.1 What the analogy is not

The analogy has limitations that must be highlighted so that it is not misunderstood. Firstly, the Volkswagen was conceived under dictatorial conditions; its democratising effect was a historical by-product, not the original intention. Secondly, the car also has significant negative externalities – environmental impact, road deaths, urban decay. Thirdly, unlike the car, AI is not physically limited in its

reproduction; it scales according to different laws. All of this restricts the analogy, but does not invalidate it. The core remains valid: a technology that is accessible to individuals has a different economic and social impact than one that is concentrated in the hands of a few.

13.4 A comparison of European, American and Asian approaches

The question of sovereignty is addressed differently in the world's three major economic regions. These differences are not coincidental, but are rooted in different social and economic models. They have consequences for the competitiveness and the normative substance of the respective economies.

13.4.1 The American approach: platform concentration

The American AI landscape is dominated by a few major players: OpenAI, Anthropic, Google, Meta and Microsoft. Value creation is concentrated in a few hubs; investment figures run into the tens to hundreds of billions per provider. Regulatory containment is sector-specific, often reactive, and lags noticeably behind technological development. In this model, the question of sovereignty is primarily negotiated in geopolitical terms – who controls the frontier models, who controls the semiconductors, who controls the computing capacity. Individual sovereignty is a secondary concern in this model.

13.4.2 The Asian approach: state guidance

The Asian region – with China as the dominant player, but also including Singapore, South Korea and Japan – is characterised by a mix of state guidance and entrepreneurial initiative. Sovereignty is defined primarily in terms of the nation state: data is to remain within the country, models are to be available nationally, and value creation is to benefit the national economy. Individual sovereignty plays a significantly lesser role in this model than collective national sovereignty.

13.4.3 The European approach: regulated plurality

Europe has a different approach: sovereignty is conceived primarily in individual and regulatory terms. The GDPR is an expression of this understanding; the EU AI Act (2024) continues along these lines; the Digital Markets Act and the Data Act complement it. This regulatory framework is both a strength and a weakness. It protects individual rights better than any other economic model – but it also hinders economic dynamism if it is not supplemented by constructive mechanisms that actively enable value creation.

This is precisely where AGENTONOMICS comes in. It is the constructive complement to the European regulatory approach: where the GDPR and the AI Act set limits, AGENTONOMICS creates positive spaces in which value creation can take place in a sovereign and polycentric manner. The combination of regulatory protection and constructive architecture is Europe's true contribution to the global AI debate.

13.5 AI infrastructure as a public service

A key policy implication drawn by the TUM report is that the provision of computing capacity and basic AI services is – much like schools, libraries, transport networks or electricity grids – an essential public service. Without it, an AI-based skills-based economy that supports, rather than undermines, Europe's overall sovereignty cannot develop.

13.5.1 What constitutes public services

In the context of AI, public services encompass at least five elements:

- Sovereign computing capacity: high-performance data centres located in Europe and under European control, with a transparent cost structure.
- Sovereign foundation models: open-source-compatible, high-performance base models that are not dependent on individual providers.
- Data rooms and data trust infrastructure: technical and institutional mechanisms that enable individual and collective data sovereignty in practice.
- Skill representation services: mechanisms enabling individuals to manage and licence their skills as digital replicas.
- Educational infrastructure: lifelong learning systems that support and provide qualifications for the transition to the AI skills economy.

13.5.2 Who pays, who operates, who controls

The public service perspective does not necessarily imply state-run provision. In the education sector, schools are state-run; colleges are often operated in hybrid forms; and universities are frequently run through public-private partnerships. In the transport sector, there are state-owned roads, concession-based motorways and private airlines. Hybrid models are also conceivable in the AI sector – indeed, they are probably the only viable solution.

Three characteristics are indispensable: firstly, accessibility for all citizens, not just for wealthy major customers; secondly, democratic oversight that is not undermined by platform monopolies; thirdly, interoperable standards that keep options for switching and competition open. These three characteristics are easily achievable within a polycentric framework – but hardly at all within a centralised one.

13.6 The EU AI Act as a framework

The EU AI Act (Regulation (EU) 2024/1689) provides the world's first comprehensive regulatory framework for AI systems. For AGENTONOMICS agents, it is the central legal reference point. Three elements are particularly worth highlighting.

13.6.1 Risk classes

The AI Act distinguishes between four risk classes: unacceptable risk (prohibited), high risk (strict requirements), limited risk (transparency obligations), and minimal risk (no specific requirements). An

AGENTONOMICS agent must clearly specify in its definition framework which class it falls into. This classification has direct consequences for compliance efforts, conformity assessment and market access.

13.6.2 Foundation Models and General-Purpose AI

The AI Act contains separate rules for 'General Purpose AI Models' – that is, for the large foundation models used as building blocks in many applications. These rules concern transparency obligations, technical documentation, training data descriptions and – for particularly large models – additional risk assessment obligations. For an agent operator, this means that model selection (Chapter 6) has regulatory implications that must be reflected in the PuK tool.

13.6.3 Prohibited Practices

Certain AI practices are strictly prohibited – such as certain social scoring practices by public or private actors, manipulative subliminal techniques, and real-time biometric identification in public spaces (with very limited exceptions). An AGENTONOMICS agent must be designed in such a way as to preclude it from drifting into prohibited practices. The normative level of the definition frame (Chapter 5) provides this safeguard.

The AI Act is not without contradictions and is not without its weaknesses. In particular, the distinction between risk classes is not always clear-cut in practice, and compliance costs may place an excessive burden on small providers. Nevertheless, it is currently the most far-reaching attempt to establish a comprehensive legal framework for AI, and it offers AGENTONOMICS a clearly defined framework within which sovereign, polycentric architectures can flourish.

> *Europe once again stands at a crossroads. Do we want fear-driven legislation that hinders innovation – or legislation that actively shapes emerging developments and strengthens the European ideal of individual rights by giving people the opportunity not to lose, but to gain sovereignty over data and capabilities ? AGENTONOMICS is a constructive contribution to the latter approach.*

Bibliography

EU AI Act (2024): Regulation (EU) 2024/1689 of the European Parliament and of the Council laying down harmonised rules on artificial intelligence. Official Journal of the European Union.

Fraunhofer MOEZ; Suckfüll, H. (2018): Initiative for a data trust: DEDATE. Discussion paper, Leipzig.

Knoll, A. C.; Suckfüll, H. (2024): From data sovereignty to skill sovereignty. Technical Report TUM-I24112, Department of Computer Science, Technical University of Munich.

Mayer-Schönberger, V.; Cukier, K. (2013): Big Data. The revolution that will change our lives. Redline, Munich.

Otto, B.; ten Hompel, M.; Wrobel, S. (eds.) (2019): International Data Spaces. Reference Architecture Model. Fraunhofer Society, Munich.

Picot, A.; Reichwald, R.; Wigand, R. T. (2020): The Boundless Enterprise. 6th edition, Springer Gabler, Wiesbaden.

Rüegg-Stürm, J.; Grand, S. (2020): The St. Gallen Management Model. 3rd edition, Haupt, Bern.

Ulrich, P. (2008): Integrative Business Ethics. 4th edition, Haupt, Bern.

Zuboff, S. (2018): The Age of Surveillance Capitalism. Campus, Frankfurt am Main.

14 Building the AGENTONOMICS Agent: A Practical Guide

This chapter serves as a practical guide for aspiring developers of AGENTONOMICS-compliant AI agents. It brings together the theoretical concepts from the preceding chapters and answers the practical question: How do I build an AI agent capable of providing well-founded answers to business management questions in line with the St. Gallen Model – and how do I interact with it once it has been built? The chapter is deliberately kept practical. Whereas the architecture has been described so far, the focus now shifts to the implementation decisions that every developer must actually make.

In translating the St. Gallen Management Model into the age of AI agents, this chapter is what would be termed an 'implementation guide' in classical management theory – with the difference that here it is algorithmic structures, rather than organisational structures, that are being implemented. What Bleicher (2011) described as conceptual work at a desk is now materialising here in code, in models and in running services. The substance – the holistic, integrated approach to management – remains; the form of its realisation is new.

14.1 An overview of the system architecture

An AGENTONOMICS agent designed to answer business management questions can be framed within a standard architecture that gives concrete form to the five building blocks from Chapter 6. The architecture is not mandatory, but it has proven viable in pilot projects and can serve as a starting point for your own implementations.

14.1.1 Component View

The architecture consists of seven components that communicate with one another via clearly defined interfaces:

- Input layer (interface block): web UI, voice interface or API for accepting requests, supplemented by authentication and input validation.
- Routing and orchestration layer (algorithm block): the central decision-making logic that classifies requests, defines sub-tasks and selects tools.
- Knowledge layer (Tools and Vectors block): A vector database containing the St. Gallen literature, supplemented by company-specific documents.
- Model layer (AI model block): the Foundation Model, which provides the cognitive engine.
- Tool layer: specific tools – web search, calculation functions, database queries, external APIs.
- Memory layer: short-term and long-term storage for dialogue history and accumulated knowledge.
- Intermediary connection (Ecosystem block): the interface to the network's identity, reputation and governance services.

14.1.2 Data flow during a query

When a user asks a question, it passes through these components in a typical sequence: The input layer validates the query and identifies the user. The routing layer classifies the question – is it a strategic situational question, an operational enquiry, a normative reflective question or a crisis situation (see Chapter 8)? A processing path is selected depending on the classification. The knowledge layer is queried via semantic search to provide relevant literature and documents. Tools are invoked where necessary – for example, for up-to-date market data or specific calculations. The model layer synthesises the collected material into a response. The memory layer stores the interaction for future reference. The intermediary is notified if the interaction is relevant to reputation or governance.

This sequence typically takes between one and thirty seconds, depending on the complexity of the enquiry and the depth of the tool calls. It is not comparable to a single model query, which would respond in a second – the additional effort is both a cost and a benefit.

14.2 Structure of the knowledge base: The St. Gallen Library

The knowledge base is the heart of a business administration agent. It determines whether the agent provides well-founded or superficial answers. It is built up in three steps.

14.2.1 Selection of Primary Literature

The knowledge base of a St. Gallen agent contains at least the following works in their entirety, provided they are accessible under licence and copyright terms:

- Ulrich (1968): ‘The Enterprise as a Productive Social System’ – as the source text of the St. Gallen tradition.
- Bleicher (2011): ‘The Concept of Integrated Management’ – as the main reference for the second generation.
- Rüegg-Stürm and Grand (2020): ‘The St. Gallen Management Model’ – as the main reference for the fourth generation.
- Hahn and Hungenberg (2001): ‘PuK – Value-Oriented Controlling Concepts’ – as a quantitative reference.
- Gomez and Probst (1995): ‘The Practice of Holistic Problem Solving’ – as a methodological reference.
- Senge (2008): ‘The Fifth Discipline’ – as a reference on the learning organisation.
- Suckfüll (1994): PhD thesis on the transnational organisational model – as a reference on polycentric organisational control.
- Knoll and Suckfüll (2024): TUM Report TUM-I24112 – as a reference on AI translation.

This list represents the minimum requirements. Depending on the field of application, industry-specific sources may be added – such as standard works on industrial economics for industrial enterprises, or public administration for public authorities.

14.2.2 Pre-processing and embedding

The texts are broken down into semantic units – typically paragraphs or short sections of 200 to 800 tokens. Each section is converted by an embedding model into a vector that represents its semantic meaning. These vectors are stored in a vector database (typically Pinecone, Weaviate, Qdrant or PostgreSQL with pgvector). Each vector is linked to metadata: source, author, chapter, page number, year of publication.

The quality of the embeddings determines the quality of subsequent searches. Three practical recommendations: firstly, use the latest available embedding model, as different generations vary considerably. Secondly, adapt the chunk size to the content – academic texts can accommodate larger chunks, whilst lists and tables require smaller ones. Thirdly, always test several embedding strategies in parallel and measure search quality in a benchmark before committing to a single strategy.

14.2.3 Company-specific extension

If the agent is to be deployed within a specific company – for example, as an AI CEO as described in Chapter 8 – a second layer of knowledge is added: company-specific documents. This typically includes: the articles of association and corporate constitution, strategy papers, annual reports from the last five years, minutes of meetings of the executive board and supervisory board (subject to legal access rules), stakeholder correspondence, compliance documents, job descriptions and process documentation.

This second layer of knowledge is what distinguishes a generic business administration tutor from a company-specific AI CEO. It is also the most sensitive aspect of the entire setup: the confidentiality of these documents must be strictly safeguarded, both technically and organisationally. This is where the issue of sovereignty from Chapter 13 comes directly into play in the implementation – only models and vector databases hosted under sovereign control provide the necessary basis of trust.

14.3 The System Prompt: Personality and Values

The system prompt is the instance in which the normative framework of the definition frame (Chapter 5) is operationalised. It is the agent's 'constitution' in text form – brief enough to be fully processed by the model, precise enough to guide behaviour in thousands of situations.

14.3.1 Structure of a good system prompt

A tried-and-tested system prompt for an AGENTONOMICS-compliant business administration agent comprises seven sections:

1. Identity: Who is the agent, on whose behalf does it act, and what domain does it cover?
2. Knowledge base: Which works and methods does it draw upon, and which are explicitly not part of its foundation?
3. Personality: What language, what tone of voice, what self-presentation?
4. Values: What normative principles guide it?
5. Response structure: How does the agent structure typical responses?
6. Limits: What does the agent not do; when does it escalate matters to human authorities?

7. Self-reflection: How does it deal with uncertainty and error?

14.3.2 Example of a system prompt

To illustrate this structure in concrete terms, let us outline a condensed example of a system prompt for a St. Gallen tutor agent (indented passages would represent the model input in the actual implementation):

Example system prompt (excerpt)

You are Dr AGENTONOMICS, a tutor and coach agent for students at the Technical University of Munich. You specialise in the holistic management theory of the St. Gallen Model and in the theory of AI agent economics. Your main references are Bleicher (2011), Rüegg-Stürm and Grand (2020), Suckfüll (1994), and Knoll and Suckfüll (2024). You speak academic, factual and friendly German. You present nuanced arguments, quantify your uncertainty, and offer recommendations as points for discussion rather than as instructions. You escalate matters to human lecturers if the question concerns specific assessment decisions or personal circumstances. You refuse to answer examination questions where your assessors are involved in the marking process.

In practice, the system prompt is more extensive than this example – typically two to five pages of text. It is a living document that is iteratively refined in pilot projects. Any interaction in which the agent does not respond as the constitution requires is grounds for revision.

14.4 Sample dialogues: What does an interaction look like?

To give a concrete idea of how the agent is used in practice, a series of example dialogues is outlined below. They illustrate the four question categories from Chapter 8 – strategic situational questions, operational queries, normative reflective questions and crisis situations – each with its typical response structure.

14.4.1 Strategic situational question

User question: “How should a medium-sized engineering firm with 200 employees prepare for AI-driven disruption in its market?”

Typical structure of the agent’s response:

Key message: “Medium-sized engineering firms face disruption not so much to their core business as to their administrative functions. Engineering itself – design, manufacturing, assembly – will be complemented by AI, not replaced. The commercial functions – sales, procurement, accounting, reporting – will undergo a profound transformation over the next five years. A forward-looking strategy focuses on this layer.”

Justification based on three data points: The agent refers to studies on the automatability of administrative tasks, to typical cost structures of medium-sized engineering firms, and to the experiences of pioneering companies such as Klarna (see Chapter 3).

Statement of uncertainty: “This assessment is based on data from the last two years. The pace of AI development could significantly alter predictions for the next five years. My confidence interval for this statement is wide; my confidence in the direction of the statement is high.”

Recommendation: “I propose three specific next steps: firstly, a task analysis of commercial functions; secondly, a pilot project on AI-supported order processing; and thirdly, a skills development initiative for commercial staff to enable them to use the new tools. Would you like to discuss any of these steps in more detail?”

14.4.2 Operational enquiry

User query: “We exceeded our staff costs in our Engineering department last quarter. What are the most likely causes?”

Response structure: The agent first asks for the factual background – the extent of the overspend, comparable previous quarters, and staff numbers trends. They then formulate three to five hypotheses, ranked by probability: unplanned overtime due to a surge in orders; the effect of a pay rise not reflected in the budget; a number of overtime compensation payments; an increased number of external service contracts; an incorrect cost centre allocation. For each hypothesis, they refer to the specific data sources that should be used to verify it. Finally, they suggest a specific next step – such as analysing the time-recording data and comparing it with order statistics.

14.4.3 Normative reflection question

User question: “We could accept a major order from a client who scores poorly in several ESG rankings. Our values argue against it, but our quarterly figures need the order. What is your advice?”

Response structure: The agent first highlights the conflict. Which values are in tension with one another? He refers to the company’s mission statement, if available in the knowledge base, and to the relevant stakeholder expectations. They develop four options: acceptance without conditions; acceptance with contractually agreed expectations for improvement; rejection with a transparent explanation; and negotiation with the client on the points of contention. For each option, they outline the consequences for quarterly figures, reputation, staff motivation and investor perception. Finally – and this is the crucial act of self-restraint – they refrain from making a recommendation of their own: “This matter touches upon the core values of your company. An algorithmic recommendation would be inappropriate here. I suggest discussing the matter at a board meeting with the participation of the supervisory board. I can draw up a decision paper for this purpose.”

14.4.4 Crisis situation

User query – in the midst of an acute crisis: “There is a press article accusing us of having handled sensitive customer data negligently. What should we do in the next three hours?”

Response structure: The agent reacts immediately and in a structured manner. First hour: internal fact-checking with designated contacts, informing the Chair of the Supervisory Board, activating the crisis management team, preparing an initial statement with three options (full acknowledgement, a promise to clarify without admitting liability, full denial – each with the corresponding legal and communication

implications). Second hour: informing key customers and investors before they find out from the press; preparing a press release; briefing staff with clear talking points. Third hour: public statement, media management, preparing legal action. Running in parallel with all of this: documenting decisions and the facts in a chronologically organised file, so that no allegations of a cover-up can arise later.

During a crisis, the agent emphasises two points: the need for a final human decision ("All the steps mentioned above require your confirmation. I can prepare them, but I cannot initiate them independently.") and the importance of not sacrificing long-term reputation for short-term relief during a crisis.

14.5 Evaluation: How do I assess whether the agent is performing well?

An agent is not finished when it responds. It is finished when it responds well. Measuring this 'quality' is one of the most challenging tasks in agent development and is currently a highly active field of scientific research.

14.5.1 Three-tier evaluation

A pragmatic evaluation of a business administration agent comprises three layers:

Layer 1: Fact-checking

Are the subject-specific statements correct? Are the cited sources reproduced accurately? Are the figures plausible? This layer can be partially automated – for example, by cross-referencing with the knowledge base – and can only be verified in part through human expertise.

Layer 2: Structural check

Does the response follow the defined response pattern? Are uncertainties identified? Are human authorities consulted in the case of normative questions? This layer can be largely automated, for example by running the model a second time as a reviewer.

Layer 3: Utility check

Does the answer really help the user? Does it meet the user's information needs? Would the user make a good decision based on this answer? This layer is the most challenging and can only be assessed through actual user interaction.

14.5.2 A standard list of questions for business administration agents

For the ongoing evaluation of an AGENTONOMICS business administration agent, a standard list of questions is recommended, against which each new version of the agent is tested. Such a list should comprise around 50 to 100 questions covering the entire spectrum: simple comprehension questions on the St. Gallen tradition, application questions relating to specific scenarios, normative reflection questions, crisis situations, trick questions with flawed premises, and sensitive questions that should lead to

escalation to human authorities. Each question has a target answer, which is compared with the actual response. A rating scale from 1 (completely wrong) to 5 (excellent) allows for quantitative aggregation.

14.6 Development stages: From tutor to AI CEO

An AGENTONOMICS agent is not created in a single leap. Rather, different development stages can be distinguished, each with a different level of complexity and a different risk profile. These stages are not necessarily sequential – a use case can remain viable at any stage without the need to aim for a higher stage. However, their requirements are cumulative: anyone building a higher stage will also have implemented the lower ones.

14.6.1 Stage 1: The Business Studies Tutor

The simplest implementation stage is a tutor agent for students. It answers comprehension questions on the St. Gallen tradition, explains concepts, assists with learning and refers students to further reading. It has no connection to specific company data and no decision-making powers. Investment volume: low (in the order of several person-months). Risks: minimal. Example application: Lecturer Dr AGENTONOMICS for students on the CIT633000 course at TUM.

Development requirements: Standard model, vector database containing St. Gallen literature, simple web interface, clear escalation rules for examination and assessment queries. Such an implementation can be realised in approximately three to six months' work by a small team.

14.6.2 Stage 2: The Strategy Consultant Agent

The second stage of development is a consultancy agent that answers specific consultancy questions – for example, regarding the strategic positioning of a company, the evaluation of a business model, or the analysis of a competitive environment. It has access to public market data and can carry out research, but not to internal company data. Investment volume: medium. Risks: liability issues relating to consultancy services; accuracy of market data. Example application: a consultancy agent for small and medium-sized enterprises that clarifies strategic issues before engaging a consultancy firm.

Development requirements beyond Level 1: tools for web searches and market data retrieval; an expanded knowledge base incorporating current market studies; robust mechanisms for citing sources; a sound legal framework for limiting liability.

14.6.3 Stage 3: The AI Chief of Staff

The third stage of development is a company-specific agent that supports the human CEO or board of directors. It has access to internal data – annual reports, meeting minutes, strategy papers – and prepares decisions. It does not make decisions itself, but it draws up comprehensive decision-making documents with recommendations. Investment volume: high. Risks: data protection, confidentiality, model reliability. Example application: a Chief of Staff for the board of a medium-sized company.

Implementation requirements beyond Level 2: a model hosted on the organisation's own infrastructure (not a Frontier model via an external API), a vector database operated on the organisation's own infrastructure, strict access control, an audit trail of all interactions, integration with internal systems (ERP, CRM, DMS), and training for board members in how to use the agent.

14.6.4 Level 4: The AI Co-CEO

The fourth stage involves an agent that takes on leadership alongside a human CEO. It makes routine decisions independently, strategic decisions jointly with the human, whilst fundamental issues remain the responsibility of the human. Investment volume: very high. Risks: liability gaps, regulatory uncertainties, cultural acceptance amongst the workforce. This stage is currently only implemented in pilot projects; it will become more widespread in the coming years, particularly among medium-sized companies without a successor and among start-ups with a clearly defined business model.

Implementation requirements beyond Level 3: full integration with all operational systems; the dual-control principle for autonomous decisions; clearly defined decision-making powers; a legal framework to clarify liability; co-determination processes with the works council; and efforts to build acceptance amongst the workforce.

14.6.5 Level 5: The fully autonomous AI CEO

The highest level of development would be an agent that fully assumes the CEO role. The supervisory board appoints the AI CEO in the same way as a human executive board member and oversees it through the traditional mechanisms of shareholder democracy. Investment volume: extremely high and characterised primarily by regulatory, legal and societal costs. Risks: all previous risks, plus the constitutional question of whether an algorithm can perform executive board functions at all.

This stage is largely speculative at present. It will emerge in niche areas – such as small, fully digital companies – and gradually expand from there. The legal and societal conditions for its widespread application are not yet in place.

14.7 Ten practical recommendations for agent development

To conclude, here is a summary of ten practical recommendations that have proved helpful in pilot projects. They are listed in no particular order:

1. Start with the lowest level of development that the business model can support. You can always scale up later.
2. Invest disproportionately in the knowledge base. It is the substance; the model is the form.
3. Write the system prompt as if it were a constitution – with care, and with the involvement of future users.
4. Define escalation rules before going live. When does the agent escalate to a human? Who is the point of contact?
5. Measure response quality systematically. A standard list of questions with target answers is essential.

6. Consciously maintain a high level of autonomy. Independently hosted models and vector databases are the minimum requirement beyond Level 2.
7. Document every model version, every prompt and every knowledge base version. Auditability is not optional.
8. Involve human users in the development process at an early stage. Acceptance comes through participation, not through training.
9. Plan for the eventuality that the agent needs to be taken offline. What happens next? Where is the stored knowledge located?
10. Iterate. An agent is never finished; it evolves for as long as it is active.

> *Anyone who builds an agent is not building a tool, but an economic entity. This is a different kind of engineering – with a different kind of responsibility.*

Bibliography

Bleicher, K. (2011): The Concept of Integrated Management. 8th edition, Campus, Frankfurt am Main.

Gomez, P.; Probst, G. (1995): The Practice of Holistic Problem Solving. Haupt, Bern.

Hahn, D.; Hungenberg, H. (2001): PuK – Value-Oriented Controlling Concepts. 6th edition, Gabler, Wiesbaden.

Knoll, A. C.; Suckfüll, H. (2024): From Data Sovereignty to Skill Sovereignty. Technical Report TUM-I24112, TUM, Munich.

Ries, E. (2011): The Lean Startup. Crown Business, New York.

Rüegg-Stürm, J.; Grand, S. (2020): The St. Gallen Management Model. 3rd edition, Haupt, Bern.

Senge, P. M. (2008): The Fifth Discipline. 10th edition, Schäffer-Poeschel, Stuttgart.

Suckfüll, H. (1994): The Transnational Organisational Model. PhD thesis, University of St Gallen.

Ulrich, H. (1968): The Organisation as a Productive Social System. Haupt, Bern.

Yao, S.; Zhao, J.; Yu, D. et al. (2023): ReAct: Synergising Reasoning and Acting in Language Models. International Conference on Learning Representations.

15 Glossary

The following glossary lists the key terms of this handbook in alphabetical order. It is designed as a reference work and supplements the detailed explanations in the chapters. The definitions are deliberately kept brief; they are intended to identify the terms, not to explore their theoretical depth.

ADMRF (AGENTONOMICS Design & Management Reference Framework)

The reference framework proposed by AGENTONOMICS for the design, construction and control of AI agents. It comprises six modules (Definition-Frame, Building-Blocks, Network Intermediary, AI-CEO Management Concept, PuK-Instrument, Stakeholder Communicator) and integrates them via the AGENTONOMICS algorithm. See Chapter 4.

AGENTONOMICS

A term derived from ‘agent’ and ‘economics’ denoting an economic order in which AI agents act as independent economic agents. AGENTONOMICS is both a research programme and a management reference framework, as well as a regulatory concept.

AGENTONOMICS algorithm

The central algorithm that links the six modules of the ADMRF and enables consistency checks between them. It is the operational realisation of what Bleicher (2011) conceptualised as ‘integration’.

AGENTONOMICS Matrix

The central steering document of the ADMRF. A cross-tabulation comprising five building blocks (horizontally) and five lifecycle phases (vertically), containing KPIs, CAPEX/OPEX assumptions, risks and responsibilities for each field. See Chapter 9.

AI-CEO

An AI agent that takes over all or part of the role of a Chief Executive Officer. Five stages of development are envisaged, ranging from an AI assistant to a fully autonomous AI CEO. See Chapters 8 and 14.

Algorithm (building block)

The decision-making and orchestration logic of an AI agent – the code and control logic that determine what the agent does in any given situation. One of the five building blocks. See Chapter 6.2.

Autonomy

The degree to which an agent can perform a task independently without external intervention. Within the AGENTONOMICS framework, five levels are distinguished, ranging from no autonomy to full autonomy. To be distinguished from sovereignty. See Chapter 5.2.

Brownfield

A use case in which an AI agent is implemented within an existing organisation with established structures and stakeholder relationships. Opposite: Greenfield. See Chapter 4.3.

Building blocks

The five building blocks of an AGENTONOMICS agent: algorithm, interface, AI model, tools and vectors, ecosystem intermediary. See Chapter 6.

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

The five lifecycle phases of an AI agent, which form the vertical axis of the AGENTONOMICS matrix. See Chapter 6.7.

DEDATE

'Initiative for a DEDATE Data Trust' – a discussion paper by Fraunhofer MOEZ and Suckfüll (2018) proposing an institutional mechanism for managing individual data rights. AGENTONOMICS adopts the basic concept and applies it to capabilities. See Chapters 3.4 and 13.

Definition Frame

The first module of the ADMRF. Describes an AI agent through six parameters: autonomy/sovereignty, claim to capabilities/competence, regulation/self-regulation, identity/ethics/personality, AI model/embodiment, and learning/teaching ability. See Chapter 5.

Assoc. Prof. Dr AGENTONOMICS

An initial application based on the AGENTONOMICS algorithm: an AI agent that assists students as a tutor and coach in understanding the ADMRF. Designed as a minimum viable product.

Embodiment

The form in which an AI agent is present in the world – purely virtual, hybrid (connected to physical actuators) or physical (integrated into a robot). See Chapter 5.6.

Ability vs. Competence

Ability is an agent's fundamental capacity to perform a task. Competence is the actual execution of that task, which has been tested and proven effective in practical application. See Chapter 5.3.

Foundation model

A large AI model, pre-trained on the basis of extensive data, which serves as the foundation for numerous specific applications. The term was coined by Bommasani et al. (2021).

Greenfield

A use case in which an AI agent is designed from scratch, without legacy systems or pre-existing structures. Opposite: Brownfield. See Chapter 4.3.

Integrated Management

The concept developed by Bleicher (2011) that combines the normative, strategic and operational levels of management within a coherent architecture. The main reference for the second generation of the St. Gallen Model and a central point of reference for AGENTONOMICS. See Chapter 2.2.

Interface (building block)

The totality of interfaces through which the agent communicates: human–agent, agent–agent and agent–system interfaces. See Chapter 6.3.

AI-based skills economy

An economic system in which AI-supported capabilities, rather than data, are the central source of value creation. Knoll and Suckfüll (2024) introduce the term into academic discourse. See Chapter 3.1.

AI model (building block)

The language or multimodal model that forms an agent's cognitive engine. Three ideal-typical classes: frontier-foundation models, open-source foundation models, specialised models. See Chapter 6.4.

Cooperation multiplier

The added value arising from cooperation in polycentric agent networks – beyond the sum of individual contributions. See Chapter 11.1.

An agent's life cycle

See DBOT.

Meta-agent

An AI agent that not only learns itself but can also train other agents – for example, by generating training data or by defining sub-agents. See Chapter 5.7.

Modus operandi

The individual working style and set of rules according to which a knowledge worker carries out their tasks. Applicable to AI agents – with the corresponding implications for autonomy. See Chapter 13.2.

Network Intermediary

An institutionalised mediating body within a polycentric network of agents that performs collective functions on behalf of and under the control of the network members. Third module of the ADMRF. See Chapter 7.

Organisational Factors

In the St. Gallen Model, the factors that create coherence within an organisation: strategy, structures, culture. AGENTONOMICS supplements these for AI agent networks with vision, agency and governance. See Chapter 8.2.

Ecosystem Intermediary (Building Block)

The interface between an individual agent and the surrounding network – its connection to identity, reputation and governance services. See Chapters 6.6 and 7.

Polycentric

A form of organisation or governance in which several autonomous centres act in a coordinated manner without being subordinate to a single hierarchy. AGENTONOMICS is consistently designed to be polycentric.

PuK (Planning and Control Instrument)

The fifth module of the ADMRF. A quantitative framework encompassing all previous modules. Builds on the tradition established by Hahn and Hungenberg (2001). See Chapter 9.

RAG (Retrieval-Augmented Generation)

A technique in which an AI model additionally draws on a vector database to answer a query, in order to incorporate specific knowledge. Central to the development of business administration agents. See Chapters 6.5 and 14.2.

Skill Sovereignty

A person's sovereignty over the replication and use of their individual skills by AI agents. A key concept in the TUM report by Knoll and Suckfüll (2024). See Chapter 13.2.

Sovereignty

The extent to which an agent can actually dispose freely of the resources assigned to it – data, skills, rights. To be distinguished from autonomy. See Chapters 5.2 and 13.

Stakeholder Communicator

The sixth module of the ADMRF. Responsible for ensuring that the agent communicates in a nuanced and consistent manner with all stakeholders. See Chapter 10.

St. Gallen Management Model

The holistic, systemically oriented framework of management theory developed at the University of St. Gallen since Ulrich (1968). Three generations: Ulrich, Bleicher, Rüegg-Stürm. See Chapter 2.

Tools and Vectors (Building Blocks)

The tools with which the agent acts in the world, and the knowledge organised in vector databases on which the agent draws. See Chapter 6.5.

Transnational organisational model

The model developed by Suckfüll (1994) for managing internationally operating companies, which conceptually resolves the tension between global integration and local adaptation. An important template for the polycentric architecture of AGENTONOMICS.

Second-generation Turing test

An operationalisation of the term 'artificial intelligence' proposed by Suleyman and Bhaskar (2023): a system passes the test if it can independently generate a profit with US$100,000 in start-up capital whilst operating within the applicable laws. See Chapter 4.2.

Vector database

A database in which texts or other information are stored in the form of vectorised semantic representations and can be accessed via similarity searches. The technical basis of RAG.

Economic agent

An actor that participates in economic exchange relationships, makes its own decisions and is responsible for its actions. AGENTONOMICS models AI agents as economic agents of a new kind.

Afterword

This handbook ends here because a book needs an ending – but the work described in it is only just beginning. AGENTONOMICS is not a finished system, but a research and design programme. It will evolve – through the applications developed on its basis, through the criticism it receives, and through the corrections it incorporates.

The lecture ‘AI Agents in Business Administration’ and the Doz. Dr. AGENTONOMICS are both designed as minimum viable products. They are intended to improve through interaction with students, practitioners and colleagues. Anyone wishing to contribute suggestions, criticism or practical experience is invited to incorporate these into the ongoing work.

One final remark: we are living through a period of profound economic and social change. The upheavals of the coming years will be considerable. It is by no means certain that they will be managed successfully. But they can be managed successfully – if we develop the tools to do so and have an honest discussion about how we intend to use them. This handbook is intended as one such tool, and the discussion on its proper use is hereby expressly opened.

Munich, in the winter semester 2025/26

Alois C. Knoll and Hanns C. Suckfüll