

AGENTONOMICS

人工智能代理经济学

关于开发和控制
经济型人工智能代理

孔安诺教授，汉斯·C·苏克富尔博士

(Prof. Dr Alois C. Knoll, Dr. Hanns C. Suckfüll)

机器人学、人工智能与实时系统教研室
慕尼黑工业大学计算、信息与技术学院
慕尼黑工业大学

2026 年第 1 版 · 2025/26 冬季学期

作者前言

本手册是伴随 2025/26 冬季学期在慕尼黑工业大学首次开设的课程《工商管理中的 AI 代理——Agentic AI in Business Administration》而编写的。它融合了两个在工商管理教学中鲜少共同探讨的知识传统：以圣加仑管理模型为代表的整体性、系统性管理理论，以及关于自主、可学习软件智能体的计算机科学。

这种结合并非出于学术上的自我目的。当前，经济现实正在以一种令传统管理学理论难以招架的速度发生变化。生成式语言模型和具备工具能力的智能代理正进入此前被视为纯属人类专属的活动领域。知识工作、决策准备、谈判引导、客户沟通——在各个领域，正在涌现出不再仅仅是工具、而是作为行动主体发挥作用的系统。由此，企业经济学的研究对象本身也在发生转变。

我们提出“AGENTONOMICS”这一参考框架，该框架将人工智能代理系统系统地视为经济主体，并将其嵌入到一种管理架构中——该架构历史性地源于圣加仑的“综合管理”概念（Bleicher 2011），并针对多中心代理网络（Suckfüll 1994）的特定需求进行了扩展。其基础算法——AGENTONOMICS 算法——及其首次应用“Doz. Dr. AGENTONOMICS”，均被设计为“最小可行产品”：它们将通过具体的应用案例进行迭代优化。

本手册旨在实现双重目标。首先，它旨在为学生提供一套精准且植根于成熟管理学文献的概念工具。其次，旨在使读者最终能够自主设计并技术实现人工智能代理，从而依据圣加仑模型对企业管理问题给出有据可依的解答。因此，本手册结尾附有关于代理构建的详细且可实际操作的指南（第 14 章）。

我们感谢那些通过提出批判性问题帮助完善论述的学生和同事，也感谢来自工业界和学术界的项目合作伙伴，他们的实践见解深刻影响了本手册中的案例分析。

慕尼黑，2025/26 冬季学期

孔安诺 和 汉斯·C·苏克富尔

目录

作者前言	2
1 《企业管理学中的 AI 代理——导论与动机》	6
1.1 本课程三个目标方向	6
1.2 目标受众与教学方法	7
1.3 在人工智能主导的工作环境中所需的规范性能力	8
1.4 AGENTONOMICS 算法作为教学实验	9
1.5 手册的结构	10
2 圣加仑管理模型作为理论基础	12
2.1 渊源：汉斯·乌尔里希的系统导向企业学	12
2.2 第二代：布莱希尔的“综合管理”理念	13
2.3 第四代：吕格-施图姆与格兰德	14
2.4 互补要素：网络化思维、学习型组织、控制论	16
2.5 哪些内容应保留，哪些需要拓展	17
3 《AGENTONOMICS》的兴起——历史经济背景分析	20
3.1 三种经济：信息、数据、能力	20
3.2 实体经济中转变的迹象	22
3.3 主权问题：从数据主权到能力主权	23
3.4 从数据托管讨论到“代理经济学”（AGENTONOMICS）	24
3.5 多中心人工智能与通用人工智能——战略界定	25
3.6 人工智能代理作为新型经济主体	25
3.7 从科学理论角度对“代理经济学”的定位	26
4 “代理经济学”设计与管理参考框架（ADMRF）	30
4.1 目标与理论背景	30
4.2 以第二代图灵测试为目标	31
4.3 应用领域：绿地项目与棕地项目	31
4.4 六个模块概述	32
4.5 各模块之间的关联	34
4.6 AGENTONOMICS 矩阵作为浓缩的控制工具	35
4.7 首次应用：“AGENTONOMICS 副教授”	35
5 模块 1：AGENTONOMICS 定义框架	37
5.1 定义框架的功能	37
5.2 参数 1：自主性与主权	37
5.3 参数 2：能力与胜任力要求	39
5.4 参数 3：监管与自我监管	40

5.5 参数 4：身份、伦理与人格.....	40
5.6 参数 5：人工智能模型与具身性.....	42
5.7 参数 6：学习与教学——元代理.....	43
5.8 一致性检验与董事会视角	43
6 模块 2：AGENTONOMICS 代理的五大构建模块	46
6.1 五大构建模块概述.....	46
6.2 构建模块 1：代理算法.....	47
6.3 构建模块 2：代理接口.....	48
6.4 构建模块 3：人工智能模型.....	49
6.5 构成要素 4：工具与向量.....	51
6.6 模块 5：生态系统协调机构.....	52
6.7 生命周期中各构建模块的协同作用.....	52
7 模块 3：网络协调机构	55
7.1 为什么需要协调机构.....	55
7.2 物业管理类比.....	56
7.3 协调机构的作用	57
7.4 协调机构的治理.....	58
7.5 德式与盎格鲁-撒克逊式治理风格的比较.....	59
8 模块 4：AI-CEO 管理理念.....	62
8.1 人工智能代理时代下的圣加仑管理模型	62
8.2 AI-CEO 的五大治理维度.....	63
8.3 AI-CEO 的具体职责.....	65
8.4 AI-CEO 的发展阶段.....	66
8.5 应用：如何与 AI-CEO 沟通？	68
8.6 人工智能 CEO 何时会说“不”	70
8.7 AI-CEO 作为通向 PuK 逻辑的桥梁	70
9 模块 5：PuK 工具	72
9.1 承袭哈恩（Hahn）和洪根贝格（Hungenberg）传统的 PuK 工具	72
9.2 AGENTONOMICS 矩阵作为核心工具.....	73
9.3 代理人的资产负债表和损益表.....	74
9.4 CAPEX 与 OPEX——投资与运营成本概况	75
9.5 生命周期中的控制流程	76
9.6 首席财务官（CFO）职能与人工智能首席执行官（AI-CEO）之间的联系	77
10 模块 6：利益相关者沟通者	79
10.1 圣加仑学派对利益相关者的理解	79
10.2 川崎式路演方法论	80

10.3 从 PuK 到 Pitch：操作层面的推导	81
10.4 危机沟通作为本模块的压力测试	82
10.5 内部与外部沟通——独特的平衡	82
11 人工智能代理经济体的增长理论	84
11.1 合作乘数	84
11.2 三元分析：三种增长形式	85
11.3 多中心网络组织作为增长引擎	86
11.4 网络中的创新与更新	87
12 经济解决方案：人工智能的青春期	89
12.1 人工智能的青春期	89
12.2 20%的增长：成因与后果	89
12.3 悲观预测：20%的结构性失业率——成因与后果	90
12.4 分配问题：谁将受益？	91
12.5 解决方案	91
12.6 欧洲的独特立场	92
13 主权：从数据主权到能力主权	94
13.1 荒诞的现状	94
13.2 下一阶段：技能主权	95
13.3 大众汽车的类比	97
13.4 欧洲、美国和亚洲各方的做法比较	98
13.5 人工智能基础设施作为公共服务	99
13.6 《欧盟人工智能法案》作为框架	100
14 构建 AGENTONOMICS 智能代理：操作指南	102
14.1 系统架构概述	102
14.2 知识库的构建：圣加仑图书馆	103
14.3 系统提示：个性与价值观	104
14.4 对话示例：交互过程是怎样的？	106
14.5 评估：如何衡量智能代理的优劣？	108
14.6 发展阶段：从辅导员到 AI 首席执行官	109
14.7 智能代理构建的十条实用建议	110
15 术语表	112
后记	118

1 管理学中的 AI 代理——引言与动机

几十年来，企业管理学的研究与教学主要将组织、市场和人类行为者作为研究对象，而随着自主人工智能系统的出现，新的行为者登上了舞台：人工智能代理。它们不再仅仅发挥辅助作用，而是越来越能够独立、自适应且有目标地行动。这一转变要求对工商管理术语体系进行根本性的扩展——而非彻底取代。课程“工商管理中的 AI 代理”及本手册正是以此为切入点。

可以观察到，尽管经典管理模型（Ulrich 1968；Bleicher 2011；Rüegg-Stürm 和 Grand 2020）仍然提供了可靠的基础，但一旦将自主、具有学习能力且以经济方式运作的人工智能系统整合到价值创造过程中，这些模型的解释力便会减弱。因此，本手册的核心问题是：

核心问题

当人工智能代理作为独立行动的经济主体出现时，应如何从企业管理角度对其进行思考、设计和控制——这又将对整体管理模型的架构产生哪些影响？

这一核心问题构成了本课程所开发、并在下文中逐步构建的“AGENTONOMICS”管理模型的出发点。该名称源自“代理”（Agent）和“经济学”（Economics）两个概念，旨在表明：不应将人工智能代理视为孤立的技术产物，而应将其视为一个独立经济秩序的组成部分。

AGENTONOMICS 宣称是一种最先进的管理模型，它借助 AGENTONOMICS 算法运用人工智能技术，为人类决策者提供在深度和速度上均优于人类的管理能力。作为第一步，AGENTONOMICS 将以讲座的形式向慕尼黑工业大学（TUM）的学生进行介绍，这些学生——无论是作为雇员还是自雇人士——都在为职业生涯做准备。

1.1 讲座的三个目标方向

本讲座旨在实现三个相互关联的目标，这些目标在手册的结构中得到了体现：

术语与概念导向

旨在使学生能够将人工智能代理不仅视为技术产物，更将其理解为经济行为主体。这包括将“代理经济学”（AGENTONOMICS）准确地置于企业管理学、国民经济学以及相关学科——

——特别是经济信息学、经济伦理学和法学——的语境中（参见 Picot 等，2020；Mainzer，2019）。

应用与设计能力

基于一个整体性的企业管理控制模型——AGENTONOMICS 设计与管理参考框架（ADMRF）——学生将学习如何系统地描述和设定 AI 代理的参数，将其作为商业模式或组织形式的一部分进行设计、技术实现，并实现经济上的成功管理。第 4 章阐述了 ADMRF 的整体架构；第 5 至 10 章则深入探讨了其六个模块。

对未来经济结构的反思

本课程明确面向那些正在为进入由人工智能主导的工作与经济世界做准备的学生——无论是作为员工、创业者，还是作为新型组织与市场形态的设计者。“AI 能力经济”这一概念由 Knoll 和 Suckfüll（2024）首次引入学术讨论，将在本手册中进行深入探讨，并作为微观经济模型加以阐述。

本课程的总体目标是在传统管理方法与创新的人工智能发展机遇之间架起一座桥梁。在此过程中，本课程不仅致力于知识传授，更始终致力于营造一个反思与实验的空间。

1.2 目标受众与教学方法

本课程有意识地面向跨学科受众。特别是面向

- 计算机科学、
- 工商管理、
- 经济学、
- 以及相关工程学和社会科学专业的学生。

本课程的教学方法采用双重视角转换，旨在培养整体思维。对于主要接受过技术教育的学生，本课程将帮助他们了解工商管理领域的思维方式、决策逻辑和管控工具。对于具有工商管理背景的学生，人工智能不会被视为“黑箱”，而是被介绍为一个可设计、参数化且具有经济效益的系统。

从中长期来看，本课程还将定位为企业管理层以及其他具有社会影响力的利益相关者群体的培训工具，这些群体希望为成功转型至以人工智能为主导的社会——即“人工智能能力经济”——做出贡献。从长远来看，本课程旨在进一步将 AGENTONOMICS 算法发展为企业实践管理的工具，并结合实际案例逐步对其进行优化。

1.3 人工智能主导的工作环境中的规范性元能力

本课程基于以下认识：要在人工智能主导的工作环境中成功立足，决定性因素并非首先是专业技术知识，而是更高层次的规范性元能力。参照自洪堡以来的教育理论思考，并结合对知识社会的当代诊断（Nida-Rümelin 和 Weidenfeld, 2018），这些能力可归纳为四大类：

体现在创业行为中的创造力

这里所指的是超越现有框架思考可能性，并将其转化为可行行动方案的能力。这种创造力并非艺术家的浪漫灵感，而是将现有资源理性组合以创造新解决方案的务实能力（熊彼特，1934）。

韧性与积极的世界观

在技术变革日新月异的世界中，能够建设性地应对不确定性已成为一项关键的经济素养。此处的韧性并非指个人特质，而是指一种可以通过具体实践——例如情景思维、学习型组织、受控实验——加以培养的能力（森格，2008）。

伦理道德的基本取向

本课程将基于康德命令的伦理取向，视为经济判断力的有机组成部分，而非附带的空谈。一个由代表委托人行事的代理人运作的经济体系，必须依赖于稳定的规范性预期——否则其交易成本将崩溃（Ulrich 2008；康德 1785）。

全面的世界知识与通识知识

在过去，人们通常称之为“通识教育”的内容，在当今这个由专业化与人工智能支持的深度研究所主导的世界中，再次成为了一种战略资源。想要评估人工智能产出所产生的影响，就必须具备从历史、哲学、自然科学和经济学等广泛视角进行思考的能力。

在此背景下，本课程将结合来自实践和多学科知识领域（甚至超出狭义企业经济学范畴）的具体、时下案例，以及日常政治事件进行探讨。

1.4 AGENTONOMICS 算法作为教学实验

本课程的一个核心特征在于其实验性。人工智能不仅是课程主题，更被积极融入教学与学习过程。从长远来看，学生将获得一位可个性化的 AI 代理——“Doz. Dr. AGENTONOMICS”作为辅助。它既担任导师、教练和讨论伙伴，同时也为学生之间的网络互动奠定基础。

“Doz. Dr. AGENTONOMICS”通过一个整体的循环过程——即 AGENTONOMICS 设计与管理参考框架（ADMRF）——来阐释该课程的六个模块。该代理应被视为基于 AGENTONOMICS 算法的首个应用。该应用使学生和从业者都能系统地、基于扎实的商业管理理论来思考、设计、实施人工智能代理，并在人工智能能力经济中成功地对其进行管理。

ADMRF 由六个模块组成，整合了企业管理的不同视角——包括首席执行官（CEO）、首席财务官（CFO）、首席技术官（CTO）、首席战略官（CSO）以及董事会整体的视角。一个核心的人工智能模型将这些视角有机结合，从而为人工智能代理的开发和管理提供最先进的决策支持。该人工智能支持的 ADMRF 明确定位为传统管理参考框架的进一步发展。从历史角度来看，它可被视为圣加仑整合管理理论（Bleicher, 2011）以及规划与控制会计模型 PuK（Hahn 和 Hungenberg, 2001）的逻辑延续。第 2 章将详细阐述这些关联。

特别是“综合管理”概念中将视角划分为规范性、战略性和操作性三个维度的划分，已被引入人工智能代理网络组织的语境中，并在 ADMRF 的六个模块中得到了体现。具体而言，多中心——即人工智能代理自主交互的网络——的控制机制在很大程度上基于跨国组织模型（Suckfüll, 1994）的组织理念，该模型是在跨国企业的应用领域中发展起来的。人工智能代理与外部利益相关者之间协调机构职能的设计也从中受益。在德国，关于构建一种能够促进数据共

享的协调机构功能的讨论——该讨论主要由《德国数据信托》（DEDATE；弗劳恩霍夫 MOEZ 和 Suckfüll 2018）的发布所引发——已被纳入 AGENTONOMICS 算法之中。

1.5 本手册的结构

本手册的设计旨在让读者在阅读完毕后，不仅能够理解 AGENTONOMICS 模型，还能亲自构建人工智能代理，从而依据圣加仑管理模型对企业管理问题给出有据可依的解答。这种兼具教科书与构建指南的双重功能体现在手册的结构中：

- 第 1 至 3 章从理论层面阐述了 AGENTONOMICS。这些章节介绍了该项目，重构了圣加仑管理模型的关键发展阶段，并从历史经济角度阐释了 AGENTONOMICS 的兴起。
- 第 4 至 10 章详细阐述了“AGENTONOMICS 设计与管理参考框架”（ADMRF）及其六个模块——从定义框架、五大构建模块和网络协调机构，到 CEO 理念、规划与控制核算以及利益相关者管理。
- 第 11 至 13 章致力于探讨增长理论、经济解决方案以及主权问题——特别是从数据主权向能力主权的过渡（Knoll 和 Suckfüll, 2024）。
- 最后，第 14 章提供了操作指南：如何构建一个符合 AGENTONOMICS 理论、能够解答圣加仑学派问题的人工智能代理？第 15 章的术语表汇总了最重要的概念。

每章末尾均附有独立的参考文献列表，正文中会引用这些文献。这样确保了各章即使单独阅读也通顺易懂。

参考文献

Bleicher, K. (2011): 《综合管理概念：愿景——使命——计划》。第 8 版，更新与扩充版，Campus 出版社，法兰克福。

弗劳恩霍夫 MOEZ；Suckfüll, H. (2018): 关于建立数据托管机构 DEDATE 的倡议。讨论文件，弗劳恩霍夫国际管理与知识经济中心，莱比锡。

Hahn, D.; Hungenberg, H. (2001): 《PuK——价值导向的控制理念》。第 6 版，修订扩充版，Gabler 出版社，威斯巴登。

康德, I. (1785) : 《道德形而上学奠基》。里加。

克诺尔, A. C.; 苏克富尔, H. (2024) : 《从数据主权到技能主权》。技术报告 TUM-I24112, 慕尼黑工业大学计算机科学研究所。

Mainzer, K. (2019): 《人工智能——机器何时接管世界? 》第 2 版, Springer 出版社, 柏林和海德堡。

尼达-吕梅林, J.; 魏登费尔德, N. (2018) : 《数字人文主义: 人工智能时代的伦理》。派珀出版社, 慕尼黑。

皮科特, A.; 赖希瓦尔德, R.; 维甘德, R. T. (2020) : 《无边界企业》。第 6 版, 斯普林格-加布勒出版社, 威斯巴登。

吕格-施图姆, J.; 格兰德, S. (2020) : 《圣加仑管理模型: 复杂世界中的管理》。第 3 版, 豪普特出版社, 伯尔尼。

舒姆佩特, J. A. (1934) : 《经济发展理论》。哈佛大学出版社, 马萨诸塞州剑桥。

森格, P. M. (2008) : 《第五项修炼: 学习型组织的艺术与实践》。第 10 版, 舍费尔-波舍尔出版社, 斯图加特。

苏克富尔, H. (1994) : 《跨国组织模型》。博士论文, 圣加仑大学。

乌尔里希, H. (1968) : 《企业作为生产性社会系统: 一般企业学的基础》。豪普特出版社, 伯尔尼。

乌尔里希, P. (2008) : 《整合性经济伦理学: 服务于生活的经济基础》。第 4 版, 豪普特出版社, 伯尔尼。

2 圣加仑管理模型作为理论基础

想要理解“代理经济学”（AGENTONOMICS），首先必须了解其所承袭的学术传统。本章重构了圣加仑管理模型的基本特征及其三个发展阶段。该模型是“代理经济学”参考框架的核心参照系，该框架既以此为导向，又在某些方面——基于合理依据——与其划清界限。因此，本章同时提供了后续章节中描述的人工智能代理所使用的术语库：若要构建一个能够依据圣加仑模型解答企业管理问题的代理，就必须精确掌握该模型。

2.1 渊源：汉斯·乌尔里希的系统导向企业学

圣加仑管理模型并非作为适合市场营销的咨询产品而诞生，而是一个科学项目。它与汉斯·乌尔里希的名字紧密相连，他在 1968 年出版的著作《企业作为生产性社会系统》（Ulrich 1968）中，将企业归入系统导向的理论传统之中。乌尔里希借鉴了一般系统理论（Bertalanffy 1968）和控制论（Wiener 1948；Beer 1972），并将这些理论的基本思想应用到企业研究中。

这一观点由三个基本假设所塑造，并贯穿该模型直至其最新版本：

企业作为开放的、目标导向的系统

企业与其环境之间存在着多种多样的互动关系。企业不能被简化为一种因果性的机械逻辑，而是依赖于意义赋予和目标导向。其可控性是有限的：与技术设备不同，社会系统无法被明确地“编程”。

复杂性是管理的基本体验

在圣加仑学派的理解中，管理并非解决界定明确的问题，而是应对复杂性。因此，经典的问题不是“什么才是正确的解决方案？”，而是“尽管我们无法完全洞悉世界，但如何才能保持行动能力？”。这种基本态度使圣加仑学派的思维至今仍与纯粹工具主义的管理理论有着鲜明的区别。

设计、引导与发展作为基本职能

乌尔里希将管理分为三项基本职能：构建制度框架、引导日常流程以及推动系统随时间发展。这一三元结构至今仍构成贯穿该模型各代理理论的隐性主轴。

定义：管理

在圣加仑学派的理解中，管理是指针对目标导向型社会系统所进行的塑造、引导和发展活动，旨在确保这些系统在高度复杂和不确定的条件下保持其生存能力和绩效能力。

2.2 第二代：布莱希尔的“综合管理”理念

克努特·布莱希尔 (Knut Bleicher) 在 20 世纪 90 年代至 21 世纪初，将圣加仑学派的理论遗产转化为广受瞩目且在国际上广为接受的“综合管理理论” (Bleicher 2011)。布莱希尔构建的理论框架是 AGENTONOMICS 理论体系最主要的理论依据，因此本文将对此进行更详细的阐述。

2.2.1 三个层面：规范层面、战略层面、运营层面

其核心在于区分三个管理层面，这些层面并非像楼层那样层层叠加，而是如同同心圆般相互交织：

规范性管理

规范性管理回答了企业最根本的意义问题：我们存在的目的是什么？哪些价值观指引着我们？我们对利益相关者承担着怎样的责任？其成果体现在企业宪章、企业政策和企业文化中。规范性层面确立了制定战略和运营决策的规则框架。

战略管理

战略管理将规范性设定转化为长期计划。其核心在于市场定位、竞争优势、业务领域、组织结构及管理体系。根据圣加仑学派的理解，战略首先不是一个优化问题，而是在复杂环境中识别模式，并在多种发展路径中进行有意识选择的问题 (Gomez 和 Probst, 1995)。

运营管理

运营管理负责确保日常业务的高效和有效实施。其经典领域包括任务分配、流程管理、成本核算、流动性控制和员工管理。重要提示：运营层面同样植根于规范和战略之中；它并非一种价值中立的执行行为。

2.2.2 三个方面：活动、结构、行为

横跨这三个层面，布莱希尔区分了管理发挥作用的三个方面：

- 活动——即企业所做的事情：计划、任务、项目、措施。
- 结构——即开展行动的制度框架：组织结构、管理体系、章程。
- 行为——行为者实际上遵循的文化模式：价值观、规范、沟通风格、领导行为。

著名的“布莱希尔矩阵”正是通过这三个层面与三个方面的交叉而形成的。它形成了九个领域，管理决策可定位于其中——从公司章程（规范性/结构）到战略规划（战略性/活动），再到狭义上的员工管理（操作性/行为）。

2.2.3 整合的概念

布莱希尔的核心概念是“整合”。这里所指的并非各部门之间的官僚式协调，而是所有层面和方面之间内部一致的相互关联。若战略未植根于企业文化，必将失败；若运营措施与愿景相悖，则会削弱管理效力。因此，整合管理意味着：矩阵中的所有九个领域之间都存在内在的、连贯的关联。

这种对连贯性的要求，正是 AGENTONOMICS 切入的切入点。布莱希尔的模型要求高层管理人员在认知层面实现连贯性，而 AGENTONOMICS 则提供了一种算法，能够从技术层面验证并支持这种连贯性。这并非对布莱希尔要求的替代，而是对其要求的强化。

2.3 第四代：吕格-施图姆与格兰德

随着 2002 年首次发布的“新”版本以及 2017/2020 年推出的第四代模型，约翰内斯·吕格-施图姆（Johannes Rüegg-Stürm）将圣加仑模型提升到了网络化组织、高沟通密集型经济的高

度 (Rüegg-Stürm 和 Grand 2020) 。这一第四代理论塑造了该模型当今的教学体系，对“代理经济学” (AGENTONOMICS) 也具有直接的现实意义。

2.3.1 六个基本类别

Rüegg-Stürm 和 Grand 依据六个基本类别对该模型进行了结构化划分，这些类别共同构成了参考框架：

环境领域

企业嵌入于社会、自然、技术和经济这四大环境领域之中。这些环境领域并非脱离语境的背景，而是积极的参与者。它们提供资源和期望，并持续检验企业行为的正当性。

利益相关方

取代单方面的股东导向的，是多元化的利益相关方群体：客户、员工、投资者、供应商、竞争对手、政府、公众、非政府组织。成功的管理在于能够有效协调这些利益相关方之间有时相互矛盾的期望。

互动议题

在企业与利益相关者之间，会出现引发关系摩擦的议题：价格、质量、风险、价值观、关切事项。这些议题的清单并非一成不变，而是处于动态变化之中。企业管理者必须能够持续洞察当前哪些议题最为突出。

秩序要素

秩序要素是新圣加仑模型的真正核心——也是 AGENTONOMICS 的关键切入点。它们涵盖战略、结构和文化，构成了日常价值创造发生的框架。AGENTONOMICS 将这一概念转化为针对 AI 代理人的五个核心组织要素：愿景、战略、结构、能动性（文化）和治理（参见第 8 章）。

流程

第四代模型将流程分为三类：管理流程（用于指导企业）；业务流程（直接创造价值）；支持流程（为上述流程创造条件）。从这一视角来看，流程并非结构的附属品，而是独立的设计对象。

发展模式

企业的发展并非遵循线性因果链，而是通过两种交替的模式进行：优化（改进现有常规流程）和革新（打破现有常规流程并建立新流程）。这两种模式均不可或缺；如何平衡二者是管理工作的核心任务。

2.3.2 沟通转向

与布莱希尔（Bleicher）相比，第四代理论的一个概念性转变在于对沟通的强调明显更加突出。基于“沟通行动”理论（哈贝马斯，1981），企业被理解为由沟通实践构成的网络。战略首先不是一个计划，而是一种共识；文化不是一成不变的价值观体系，而是对何为恰当的持续协商。

这一“沟通转向”对 AGENTONOMICS 具有双重意义。首先，因为作为人工智能代理的语言模型，其运作层面恰恰是新圣加仑模型所确定的核心作用层面：即沟通。其次，因为多中心代理网络必须通过沟通协商来建立其凝聚力——这一点将在第 7 章中深入探讨。

2.4 互补性构成要素：网络化思维、学习型组织、控制论

圣加仑模型并非一成不变。它汲取了多条思想脉络，这些脉络对“代理经济学”（AGENTONOMICS）同样具有重要意义。

2.4.1 网络化思维

戈麦斯和普罗布斯特（1995）在其“网络化思维”方法论中开发了一系列工具，可用于系统地解析复杂问题领域：影响矩阵、效应图、情景分析、调控空间。在当今的人工智能时代，这些工具非但没有过时，反而比以往任何时候都更具现实意义：它们恰恰提供了那种结构化的形式，使人工智能代理能够以此为人类委托人梳理复杂的实际情况。

2.4.2 学习型组织

彼得·圣吉（Peter Senge, 2008）将“学习型组织”的概念引入了管理学的主流。该概念由五大要素构成：个人精进、思维模型、共同愿景、团队学习、系统思维。将其应用到人工智能代理网络中，一个显而易见的事实是：一个由学习型代理构成的运营层组织，必须确保其在更高层面的反思能力——否则，这些代理的优化将陷入局部死胡同。这种反思层面正是 AI-CEO 的主要职责之一（第 8 章）。

2.4.3 管理控制论

斯塔福德·比尔（Stafford Beer）的“可存活系统模型”（Beer 1972）提出了一种由五个子系统组成的可存活系统的控制论架构。该架构对人工智能代理网络的设计具有重要意义：它阐明了系统为保持可存活性必须至少具备哪些功能——从运营执行、协调与运营管理，到战略和规范性控制。“AGENTONOMICS”在其模块结构中隐含地采用了这一架构。

2.5 哪些内容得以保留，哪些需要拓展

有三个要点需要指出，这些方面既体现了 AGENTONOMICS 对圣加仑学派的传承，也展现了其超越之处：

继承：整体架构

规范-战略-运营的三元结构、对利益相关者的关注、秩序要素与过程的区分，以及沟通的基本态度，均被 AGENTONOMICS 完全继承。ADMRF（第 4 章）在其架构上是一个适用于 AI 代理网络的圣加仑模型。

在继承基础上调整重点：多中心性

早在跨国组织模型（Suckfüll 1994）中，圣加仑框架就已表明，国际运营的企业既不能采用纯粹的集中式管理，也不能采用纯粹的分散式管理，而是需要一种多中心架构。这一见解在 AGENTONOMICS 中被延伸至人工智能代理层面：代理网络既非纯粹的等级制，也非无序的，而是呈现多中心化的组织结构（第 7 章）。

扩展：算法一致性检验

传统圣加仑模型将一致性能力归因于管理层的认知层面，而“代理经济学”则在此基础上增添了一层算法机制。“代理经济学”算法通过技术手段检验规范层面、战略层面与运营层面之间的一致性，并针对发现的不一致之处提出处理建议。这并非取代圣加仑模型，而是将其发展为一种具备人工智能能力的变体。

扩展：AI 代理主体

最后——这也是最根本的扩展——AGENTONOMICS 在该模型中引入了一种新的行为主体：作为经济主体的 AI 代理。圣加仑模型迄今为止一直隐含地以人类和组织行为体为前提。随着人工智能代理的出现，一种融合了技术产物与经济行为逻辑的混合体登上了舞台。这一扩展带来的影响将在第 5 章和第 6 章中展开论述。

圣加仑管理模型并未过时——恰恰相反：它是一种具有极强兼容性的管理理论，能够真正应对多中心人工智能代理网络的复杂性。只需将其逻辑推演到底即可。

参考文献

Beer, S. (1972): 《企业的脑：组织的 managerial cybernetics》。Allen Lane, 伦敦。

Bertalanffy, L. von (1968): 《一般系统理论：基础、发展与应用》。Braziller 出版社，纽约。

Bleicher, K. (2011): 《综合管理概念：愿景——使命——计划》。第 8 版，更新与扩充版，Campus 出版社，法兰克福。

戈麦斯, P.; 普罗布斯特, G. (1995) : 《整体性问题解决与实践：网络化思维——企业家式行动——个人说服力》。豪普特出版社，伯尔尼。

哈贝马斯, J. (1981) : 《沟通行动理论》。两卷本，苏尔坎普出版社，法兰克福。

吕格-施图姆, J.; 格兰德, S. (2020) : 《圣加仑管理模型：复杂世界中的管理》。第 3 版，Haupt 出版社，伯尔尼。

森格, P. M. (2008) : 《第五项修炼：学习型组织的艺术与实践》。第 10 版，舍费尔-波舍尔出版社，斯图加特。

苏克富尔, H. (1994) : 《跨国组织模型》。博士论文, 圣加仑大学。

乌尔里希, H. (1968) : 《企业作为生产性社会系统》。《一般企业学基础》。Haupt 出版社, 伯尔尼。

乌尔里希, H. (1984) : 《管理》。豪普特出版社, 伯尔尼。

维纳, N. (1948) : 《控制论: 论动物与机器中的控制与通信》。麻省理工学院出版社, 马萨诸塞州剑桥。

3 “代理经济学” (AGENTONOMICS) 的兴起 ——历史经济视角下的定位

为了清晰地定位“代理经济学”，有必要审视数字经济形态的发展历程。当前关于人工智能代理的讨论并非孤立存在，而是多次连续经济范式转变的结果。从信息经济到数字经济，再到人工智能能力经济的过渡，展现出一种模式，这种模式让我们理解了当下的逻辑——并为“代理经济学”的介入开辟了论述空间（Knoll 和 Suckfüll 2024）。

3.1 三种经济形态：信息、数据、能力

3.1.1 20 世纪 90 年代的信息经济

在 20 世纪 90 年代的信息经济中，现有的商业模式主要被转移到数字渠道上。企业最初将互联网作为额外的沟通和销售渠道。尽管涌现出许多创新的商业模式（Shapiro 和 Varian, 1999），但许多早期的尝试被证明是不可持续的。网络泡沫以及 2000 年开始的互联网崩盘标志着一个初步的学习过程：并非每项技术创新都会自动带来可持续的价值创造。

关于这一阶段的理论描述主要见于网络经济学（Shapiro 和 Varian, 1999）和平台经济学（Rochet 和 Tirole, 2003）的相关研究。其核心概念——网络效应、锁定效应、标准竞争——至今仍是描述数字经济商业模式的术语体系。然而，这些概念基于“信息”这一商品在很大程度上可自由交换的假设——这一假设在随后的数字经济中必须予以修正。

3.1.2 数字经济

数字经济是从信息经济逐步发展而来的。在此背景下，数据日益被视为一种战略资源和潜在的资产（Mayer-Schönberger 和 Cukier, 2013）。价值创造的重心已从单纯提供信息，转向对数据的系统性收集、分析和货币化。大型数字平台建立了基于规模效应、网络效应和锁定机制的商业模式（祖博夫, 2018）。

尽管取得了经济上的成功，数字经济仍暴露出三种基本的结构性张力，这些张力对于理解“代理经济学”（AGENTONOMICS）的形成至关重要：

权力关系不对称

对数据的支配权集中在少数全球参与者手中。提供数据的个人和组织，在合理分享由此产生的价值方面，机会十分有限。这一趋势目前正呈指数级加剧。正如诺尔和苏克富尔（2024）所强调的，发达国家的一名普通成年人每年产生的个人数据估计达数百千兆字节——从领英（LinkedIn）个人资料、在线求职申请，到电子账单和合同。这些数据的经济价值估计为每人每年 10,000 至 15,000 欧元——而这笔价值几乎完全由大型平台所掌控。

失灵的数据市场

尽管数据已被公认为具有经济价值，但在国内和国际层面均未能建立起透明、公平且可扩展的数据交换市场。权利归属、估值和报酬等问题在很大程度上仍未得到解决（Otto 等，2019）。由此导致了一种悖论：数据虽具有高经济价值，却无法由个人进行交易——其结果是数据生产者事实上被剥夺了所有权。

监管碎片化

政府干预——尤其是通过数据保护和数字监管——主要旨在控制风险和实施监管。然而，这些措施并未建立起独立、富有成效的市场机制，反而在某些情况下导致了市场壁垒和创新障碍。正如克诺尔和苏克富尔（2024）所尖锐指出的：如果在德国尤为突出的“数据节约”原则没有被个人数据主权的富有成效的机制所制衡，那么它就是“通向贫乏的捷径”。

3.1.3 人工智能能力经济

随着高性能生成式人工智能系统的突破——尤其是自 2022 年起大型语言模型发布以来（Brown 等人，2020；OpenAI，2023；Bommasani 等人，2021）——经济价值的创造重心再次发生转移。数据并未失去其重要性，但正日益成为一种中间产品：只有通过人工智能模型，数据才能转化为与行动相关的技能。

这一发展可被描述为向“AI 能力经济”的过渡。在此经济模式下，核心不再是孤立的数据集，而是 AI 支持的能力，例如：

- 决策支持、
- 模式识别、

- 预测与规划,
- 与市场及组织的自主交互,
- 综合复杂知识体系。

在此过程中，具有经济意义的不仅是技术性能，更是能够有针对性、持久且创造价值地运用这些人工智能能力。克诺尔和苏克菲尔（2024）一针见血地提到了第二次主权飞跃：继（德国在很大程度上错失的）数据主权飞跃之后，现在即将迎来能力主权（skill sovereignty）的飞跃。

3.2 实体经济中转型的迹象

从数据经济向能力经济的转型，早已在德国经济媒体的日常报道中有所体现。克诺尔和苏克菲尔（2024）列举了两个具有代表性的例子，本文将对此进行进一步探讨：

3.2.1 联邦就业局案例

联邦就业局执行董事安德烈娅·纳勒斯曾公开指出，在未来十年内，该局 11.4 万名员工中将有 3.5 万人退休，而传统意义上的接班人却无法补充。结果：未来该机构的服务水平只能通过自动化——即运用人工智能——来保障。这已不再是效率顾问的愿望，而是一个大型公共行政机构对现状的冷静评估。

3.2.2 Klarna 案例

瑞典金融服务提供商 Klarna 目前已停止招聘新员工（工程师除外）。在一年内，通过应用人工智能，员工人数从 5,000 人缩减至 3,800 人，长期目标是将员工规模控制在 2,000 人——而在此期间，总工作量却在持续增加。这并非传统意义上的人员裁减，而是一种结构性的职能转变。尽管这一激进举措后来被证明有些过头，但从大趋势来看，这一趋势依然存在。

这两个案例生动地说明了“人工智能能力经济”的具体含义：决定性因素不再是人类员工的人数，而是具有经济价值的技能数量（无论是由人类提供还是由智能体复制）。这一转变在许多方面与 20 世纪 70 年代以来体力劳动的发展相似——但关键区别在于：在信息处理领域，计算机几乎不存在原则性限制，而工业机器人迄今仍无法企及受过训练的人类双手的灵巧程度（Knoll 和 Suckfüll, 2024）。

3.3 主权问题：从数据主权到能力主权

克诺尔和苏克富尔（2024）在其技术报告 TUM-I24112 中提出了一个对 AGENTONOMICS 具有构成性意义的论点，本文将对此进行详细阐述，因为它构成了整个参考框架的规范基础。

3.3.1 荒诞的现状

核心诊断如下：我们实际上正处于一种荒谬的境地。一方面，数据保护立法以及缺乏针对个人数据使用的操作机制，在许多方面阻碍了社会价值的创造；另一方面，为了参与现代数字化生活，每个人都必须披露海量的个人数据——却几乎得不到任何回报。

国家真正的职责，应该是为每个人提供一套全面的工具，使每个人都能获得对自己数据的主权，并自主决定如何处理这些数据。尽管实现这种个人数据主权的立法和操作基础尚不明确，但现在着手建立还为时不晚。

3.3.2 技能主权作为下一阶段

关键在于克诺尔和苏克富尔（2024）提出的观点：当政界仍在就如何使个人数据主权付诸实践展开辩论时，下一个主权问题——即“技能主权”（Skill sovereignty）——已经迫在眉睫。

核心论点（Knoll 和 Suckfüll 2024）

每位知识工作者都会根据自己在教育和职业生涯中习得并不断发展的规则来处理输入数据或信息。通过这种方式，每个人都会发展出自己的专业能力，形成自己的工作模式。这种个人的专业能力正越来越多地转移到计算机，即人工智能上。“技能转移”的机制正变得越来越高效。如果这种转移出现差错，知识型劳动者的处境将比 19 世纪的“砸机器者”还要糟糕。如果转移成功，每个人都能成倍地提升自己的工作能力。

关键在于：将能力移交给机器的人，应当继续掌握这些能力——即使这些能力被倍增并并行运用。具体来说，这意味着：虽然一个人可以同时为多个委托方工作，但这些委托方不能独立使用该人的“功能性复制体”。该人应当保留“版权”，即对其全部复制体资源的完全处置权。

3.3.3 政府的责任

由此推导出一个明确的监管要求：提供计算能力和人工智能基础服务——与学校、图书馆或交通网络类似——是公共服务中关乎生存的核心任务。没有这些服务，基于人工智能的能力经济就无法发展——至少无法发展出一种既能支撑欧洲整体主权、又不会削弱其主权的经济模式。

因此，欧洲再次站在了十字路口：我们是想要一套阻碍创新、由恐惧驱动的立法，还是想要一套能够积极塑造新兴发展趋势的立法——通过赋予人们机会，使其不仅不会失去对自身数据和能力的主权，反而能够赢得这种主权，从而强化欧洲关于个人权利的理念？AGENTONOMICS 为第二种选择做出了建设性的贡献。

3.4 从数据托管讨论到 AGENTONOMICS

AGENTONOMICS 的一个重要概念先驱，体现在关于数据托管模型的讨论中——这一讨论在政治层面尤为重要。数据托管旨在从法律上确保平衡三类典型利益相关方的利益：

1. 数据提供者——生产数据的主体。
2. 数据利用者——对数据进行经济利用的主体。
3. 数据处置权人——依法拥有数据处置权的参与者。

关于这一讨论的首批相关论述可见于《DEDATE 数据托管倡议》（弗劳恩霍夫 MOEZ 研究所与 Suckfüll, 2018 年），以及其进一步发展、目前尚未公开发布的《DEDATE 2.0》工作文件。使法律上明确可确定的数据处置权人能够在技术上自主地与他人就“其”数据进行协商，是数字经济有效运行的最终目标。

AGENTONOMICS 继承了这一基本理念，但将重点转移到了：动态的、基于人工智能的能力取代了静态的数据权利，成为价值创造的核心。AI 代理除了发挥自身的核心能力外，还作为自主且独立行动的 AI 能力协调机构，能够根据具体情境协商权利、利益和价值创造。这一过程所展现的复杂性和速度，是仅靠人类或机构解决方案几乎无法企及的。

3.5 多中心人工智能与通用人工智能（AGI）——战略界定

“AGENTONOMICS”的兴起与对中央集权式人工智能战略的批判性探讨密切相关，正如“只有规模才重要”这一口号所尖锐指出的那样。当前的全球竞争主要体现在试图开发尽可能全面、通用的通用人工智能（AGI）系统，其特点包括：

- 巨大的规模效应，
- 高资本密集度，
- 集中式治理结构，
- 寡头垄断至垄断式的市场地位。

与此相对，AGENTONOMICS 则采取基于网络的多中心方法。其关注点不在于单一的、包罗万象的人工智能，而在于众多专业化且相互协作的人工智能代理。价值的创造主要并非源于集中化，而是源于网络内部的协调、分工和情境化领导。借鉴万花筒的原理，AGENTONOMICS 代理可根据具体情境组合并整合资源：试想，西门子、宝马和德国电信将自身——以及它们之间的接口——定义为人工智能代理，并相互协作以完成共同任务。

这种方法不应被视为技术层面的对立方案，而应被视为一种源于经济和社会动机的替代方案——特别是在参与权、所有权结构和可持续价值创造等问题上。集中式人工智能的统治存在这样的风险：，由于缺乏竞争，财富分配将变得极度不均衡，以致于购买力的匮乏将动摇整个资本主义体系的根基。

3.6 人工智能代理作为新型经济主体

“AGENTONOMICS”概念正是从这一点切入。它将人工智能代理视为具有经济行为能力的参与者，从而在现有的人工智能应用基础上更进一步。这些代理不仅仅是工具或辅助系统，而是：

- 拥有独立的决策逻辑，
- 在既定框架内自主开展经济活动，
- 与其他智能体、组织及人类进行交互，
- 致力于实现可持续的价值创造和自我优化。

在经济学术语中，人工智能代理因此可被理解为一种新型的“经济主体”。它们参与交换关系，进行竞争或合作，并在稀缺条件下做出决策。这些相互作用的经济主体进而构成一个可独立描述的市场，该市场在微观经济层面具有自身独特的竞争条件、市场参与者、交换关系和价格形成特征。

其中，从数字经济向人工智能能力经济的转型阶段尤为引人关注。迄今为止，人工智能代理之间的纯粹交互仍仅是一种理想情景。因此，需要对人—人工智能代理—人工智能代理之间的关系及交互界面进行全面重新思考。“代理经济学”（AGENTONOMICS）正是这种重新思考的体现。

3.7 AGENTONOMICS 的科学理论定位

AGENTONOMICS 既不是短暂的时尚现象，也不是一个流行词。它也不是严格意义上的实证描述性理论。相反，它是一个理想理论模型，为具有经济行为能力的人工智能代理提供了一个概念性参考框架。理想理论模型的特点在于：

- 对复杂的现实现象进行抽象化，
- 孤立核心作用机制，
- 为分析和设计提供指导性知识。

“代理经济学”（AGENTONOMICS）遵循这一逻辑，将人工智能代理建模为经济理想类型。这些代理具备明确的能力、决策逻辑和激励机制，尽管实际实现可能与此有所偏差。人工智能代理的标准化通过规范化，将构建一个创造更多价值的生态系统：兼容的接口提高了代理之间产生高效互动的可能性。

3.7.1 位于企业经济学与经济学交叉点

在经济学领域，AGENTONOMICS 位于企业经济学与国民经济学的交叉点上。

企业管理视角

从企业管理学的角度来看，核心问题在于如何对经济实体进行控制、协调并使其致力于创造价值。AGENTONOMICS 将人工智能代理视为可控但并非完全可确定的行为主体，从而拓展了

这一视角。其核心参考点包括管理与控制模型、组织与网络结构、战略决策过程以及治理与激励体系。

宏观经济视角

从国民经济学的角度来看，AGENTONOMICS 与微观经济学紧密相关。人工智能代理被建模为市场参与者，它们参与交换关系，塑造竞争与合作局面，并参与价格形成过程。然而，与传统行为体不同的是，人工智能代理具有高度的可扩展性、可联网性和学习能力。这要求对微观经济模型进行扩展，特别是在网络经济和平台经济方面。

3.7.2 与相关论述的界定

AGENTONOMICS 有意识地与若干既有的学术讨论领域划清界限：

- 纯粹的人工智能研究侧重于算法、模型和技术性能，但很少涉及经济调控问题。
- 数字平台经济学分析市场支配性结构，但通常以人类或组织行为体为前提。
- 关于自动化和替代的讨论主要将人工智能视为人类劳动的替代品，却忽视了其作为独立行为主体的作用。

相比之下，“代理经济学”（AGENTONOMICS）将技术、经济和制度视角整合为一个独立的研究对象。跨学科性是其本质特征。主要涉及的学科包括计算机科学（自主系统开发）、企业管理学（调控与价值创造模型）、经济学（市场与竞争结构）以及法学（责任、所有权和治理问题）。

3.7.3 研究与设计计划

除了其理论定位之外，“AGENTONOMICS”还应被视为一项研究与设计计划。它不仅涉及分析，更明确致力于开发、测试和评估新的经济行为形式：

- 将人工智能代理设计为经济主体，
- 设计激励与治理结构，
- 开发实验性的商业和组织模型，
- 在真实或模拟环境中进行实证测试。

从这个意义上说，AGENTONOMICS 在理论、实验与实践之间架起了一座桥梁。人工智能代理的开发尤其适合以“最小可行产品”（MVP）的形式进行，正如 Ries（2011）在“精益创业”概念中所系统阐述的那样。“AGENTONOMICS 副教授博士”正是这样一个 MVP。

*AGENTONOMICS 是对一项早已发生的技术变革所作出的经济上自治的回应——
尽管这一变革尚未被管理学界所接纳。*

参考文献

- Bommasani, R.; Hudson, D. A.; Adeli, E. 等（2021）：《关于基础模型的机遇与风险》。
arXiv:2108.07258, 斯坦福大学 CRFM。
- Brown, T.; Mann, B.; Ryder, N. 等（2020）：《语言模型是少样本学习者》。载于：《神经信息处理系统进展》第 33 卷，第 1877–1901 页。
- 弗劳恩霍夫 MOEZ；苏克富尔，H.（2018）：关于建立数据托管机构 DEDATE 的倡议。讨论文件，弗劳恩霍夫国际管理与知识经济中心，莱比锡。
- Knoll, A. C.; Suckfüll, H. (2024): 从数据主权到技能主权。技术报告 TUM-I24112, 慕尼黑工业大学计算机科学研究所。
- Mayer-Schönberger, V.; Cukier, K. (2013): 《大数据：将改变我们生活的革命》。Redline 出版社，慕尼黑。
- OpenAI（2023）：《GPT-4 技术报告》。arXiv:2303.08774。
- 奥托，B.；滕·霍姆佩尔，M.；沃罗贝尔，S.（编）（2019）：《国际数据空间：参考架构模型》。弗劳恩霍夫协会，慕尼黑。
- 里斯，E.（2011）：《精益创业：当今企业家如何通过持续创新打造极具成功的企业》。Crown Business 出版社，纽约。
- 罗谢，J.-C.；蒂罗尔，J.（2003）：《双边市场中的平台竞争》。载于：《欧洲经济协会期刊》1(4)，第 990–1029 页。
- 夏皮罗，C.；瓦里安，H. R.（1999）：《信息法则：网络经济的战略指南》。哈佛商学院出版社，波士顿。

苏莱曼, M.; 巴斯卡尔, M. (2023) : 《即将到来的浪潮: 技术、权力与二十一世纪的最大困境》。克朗出版社, 纽约。

祖博夫, S. (2018) : 《监控资本主义时代》。坎普斯出版社, 法兰克福。

4 AGENTONOMICS 设计与管理参考框架 (ADMRF)

本章构成本手册的操作核心。它阐述了作为集成参考框架的 AGENTONOMICS 设计与管理参考框架 (ADMRF) 的整体逻辑。本章简要介绍了 ADMRF 的六个模块，并将其置于整体框架中进行定位；随后第 5 至 10 章将详细展开每个模块的内容。对于希望构建人工智能代理以解答商业管理及圣加仑学派相关问题的人士而言，ADMRF 同时提供了此类代理必须遵循的知识架构。

4.1 目标与理论背景

ADMRF 基于圣加仑管理模型——特别是 Bleicher (2011) 提出的“综合管理”概念以及 Rüegg-Stürm 和 Grand (2020) 提出的第四代模型。这六个模块将一个整体性整合管理参考框架中的规范层面、战略层面和运营层面有机结合。ADMRF 被视为由符合共同标准——即 AGENTONOMICS 标准——的人工智能代理组成的网络组织的管理参考框架。

针对多中心或准等级制网络的治理理念借鉴了博士论文《跨国组织模型》(Suckfüll, 1994) 中的观点。该模型之所以具有重要意义，是因为早在 20 世纪 90 年代，它就试图解决这样一个问题：一个由自主且对当地负责的单元组成的组织，如何同时实现全球凝聚力和本地适应性——这恰恰是当今人工智能代理网络必须解决的问题。

作为一项技术，人工智能通过两种方式为 ADMRF 所用。一方面，它用于优化计算各模块之间的相互关系以及利益相关者的参与；另一方面，它本身也是控制的对象，因为需要管理的经济主体正是人工智能代理。这种双重角色具有构成性意义：人工智能既是控制的主体，也是控制的工具。

下文介绍的六个模块并非孤立存在，而是被视为一个集成化发展与控制过程中相互构建的阶段。必须明确考虑其中的相互依赖性和交互作用，才能理解整体逻辑——即“AGENTONOMICS”算法——并识别其应用领域。

4.2 第二代图灵测试作为目标基准

人工智能代理的自主性目标——即在市场上取得经济成功，并在此过程中能够利用组织（即法人）的所有自由度——借鉴了苏莱曼和巴斯卡尔（2023）在《The Coming Wave》中提出的第二代图灵测试的要求。

苏莱曼的现代图灵测试（意译）

如果一个人工智能系统能够在没有其他具体指导的情况下，独立完成“在遵守所有现行法律的前提下，利用 10 万美元的初始资金，在零售网络平台上于数月内赚取 100 万美元”这一任务——包括商业构想、合同签订、市场营销和税务处理——则该系统通过了现代图灵测试。

AGENTONOMICS 人工智能代理是满足这些要求道路上的下一步。在此过程中，第二代图灵测试主要不是技术标准，而是商业标准：它检验的是经济行动能力。ADMRF 作为连接框架，将创业初始构想、管理逻辑和实际实施——直至成功运营——有机地结合在一起。

4.3 应用领域：绿地项目与棕地项目

ADMRF 同样适用于两种不同的应用场景，在企业实践中通常被称为“绿地”和“棕地”案例：

绿地项目：在空白领域开发 AI 代理

在此情境下，AI 代理是从零开始设计的。不存在既有资产、既定流程，也没有既定的利益相关方期望。ADMRF 指导从定义框架、构建模块到利益相关方管理的所有阶段。这种应用典型地见于开发基于代理的商业模式的初创企业。

棕地 (Brownfield)：现有组织的转型

在此场景中，现有组织——无论是传统企业、政府机构还是非政府组织——将逐步转型为人工智能代理，或转型为由此类代理组成的网络。ADMRF 作为路线图，引导现有系统向支持代理的结构转型。这种应用场景远比前者更为常见，但也更为复杂。

在这两种情况下，ADMRF 都作为连接企业初始构想、管理逻辑与实际实施的统合性原则发挥作用。换言之：它描述了人工智能代理如何从概念定义、模块化实现、经济控制，直至达到市场成熟度，并在此过程中进行迭代开发。

4.4 六个模块概述

ADMRF 的核心是六个模块，用传统企业管理的术语来说，它们体现了六种理想型的管理视角。本文将对它们进行简要介绍；第 5 章至第 10 章将详细展开阐述。

4.4.1 模块 1：定义框架——企业定位

第一个模块——AGENTONOMICS 定义框架——通过六个参数来描述智能体：自主性与主权、能力与胜任力要求、监管与自我监管、身份与伦理、AI 模型与（虚拟或现实）具身化，以及学习与教学能力（元智能体）。原则上，AGENTONOMICS 对人工智能代理的定义要求其能够作为可区分的经济主体，并能够与其他代理——在协调机构职能的组织和监督下——进行合作。所有权结构使其能够独立于其他代理，因此可以称之为多中心控制的组织。

该模块是应对企业挑战的一种工具，旨在定义一个符合 AGENTONOMICS 代理要求的 AI 代理。所有管理层视角（CEO、CFO、CTO、CSO）应共同作为董事会参与其中。参见第 5 章。

4.4.2 模块 2：五大构建模块——技术实现

第二个模块将人工智能代理分解为五个在法律和技术上可明确界定的构建模块：代理算法、代理接口、人工智能模型、工具与向量，以及代理生态系统协调机构。每个构建模块均由专属徽标标识；所有徽标组合在一起，在图形上构成了人工智能代理。对各个构建模块的技术实现，以及针对框架预先定义的 AI 代理对其组合的优化，旨在将代理的技术开发与商业运营开发这两项核心任务有机结合。

从商业管理角度来看，这一阶段可被视为业务开发流程与产品开发流程的结合——即遵循在定义框架中制定的愿景和使命而进行的流程。“5 个构建模块”模块充当了商业问题与技术问题之间的接口。它基本上体现了首席技术官（CTO）的视角。参见第 6 章。

4.4.3 模块 3：网络协调机构——代理与治理

第三个模块探讨了网络受托人（Network-Trustee）或网络协调机构。AI代理具备协作能力，能够作为在多中心组织网络中自主互动的单元，产生协同增值效应。网络凝聚力源于 AGENTONOMICS 代理体之间准民主式的协商过程。就此而言，网络受托人代表了参与其中的人工智能代理体之间——类比于“沟通行动理论”（哈贝马斯，1981）——无等级结构的讨论所产生的结果。

由网络通过共识制定的、具有传递性质的职能，将由所谓的“网络协调机构”负责具体执行。可以以某处房产的业主委员会为例，其权利和义务由物业管理公司负责具体执行。在设计单个 AGENTONOMICS 人工智能代理时，必须从一开始就考虑此类情境特定协调机构功能（如数据受托人、经纪代理、清算所等）的具体规范，这些功能负责协调权利、数据、能力和交易。参见第 7 章。

4.4.4 模块 4：AI-CEO 管理理念——将秩序要素 operationalisieren

第四个模块——管理工具或 AI-CEO——将愿景、战略、结构、能动性（文化）和治理等秩序要素 operationalisiert，作为在管理层或董事会层面为企业决策提供全面周到的支持。这是一个基于人工智能的整体管理概念，可视为布莱希尔（Bleicher，2011）提出的“整合管理”概念的进一步发展。该参考框架（亦被称为“圣加仑管理概念”）的重点集中在人工智能代理网络上。参见第 8 章。

4.4.5 模块 5：PuK 工具——定量框架

第五个模块是 PuK 工具（规划与控制工具）。它基于 Hahn 和 Hungenberg（2001）开发的 PuK 模型。该概念同样侧重于在网络中运作的人工智能代理，并借助人工智能模型加以整合。AGENTONOMICS 矩阵的系统性填充，横向采用列出的五个构建模块，纵向则采用设计/建设/运营/拥有/移交（DBOT 逻辑）的阶段，并根据哈恩和洪根贝格（2001）的规范以及财务管理的基本原理（伯克和德马佐，2020），为每个阶段分配相应的资产负债表和损益表。该模块承担了传统、以人为主导的组织中首席财务官（CFO）的职责。参见第 9 章。

4.4.6 模块 6：利益相关者沟通者——由 AI 驱动的对话

第六个模块是 AGENTONOMICS 利益相关者沟通者。它根据 AGENTONOMICS-PuK 工具或针对相应 AI 代理填充的 Excel 工具，推导出针对特定目标群体的沟通格式——包括员工、投资者、战略合作伙伴及其他利益相关者。“路演大师”借鉴了盖伊·川崎（Guy Kawasaki, 2015）在苹果公司担任史蒂夫·乔布斯“布道师”期间所开展的工作。该模块承担着利益相关者沟通的职责；也可将其称为 CSO（首席销售官）。参见第 10 章。

4.5 各模块之间的连接方式

ADMRF 的六个模块通过一个中央 AI 相互连接。AGENTONOMICS 算法负责建立网络，以针对给定的问题提供最优解决方案。在此过程中，可区分三种连接类型：

顺序连接：生命周期

从第一种解读来看，这六个模块构成一个生命周期：从定义框架（Definition-Frame）开始，经过构建模块（Building-Blocks）和中间环节（Intermediär），到达 AI-CEO，继而进入 PuK，最终到达利益相关者管理。这种顺序解读符合绿地项目（Greenfield-Projekt）中的典型流程顺序。

循环连接：反馈机制

但实际上，ADMRF 并非线性过程，而是一个循环过程。来自利益相关者沟通器的反馈流回 PuK，PuK 可以调整 AI-CEO 概念，AI-CEO 概念改变中间架构，中间架构进而改变构建模块，最终也影响定义框架。对任一模块的调整都可能引发对所有其他模块的调整。正是这些反馈机制，AGENTONOMICS 算法旨在通过计算手段加以掌控。

横向互联：秩序要素

贯穿所有模块的五大秩序要素包括愿景、战略、结构、能动性与治理。它们并非独立的模块，而是定位各模块的坐标系。模块中的每一项内容均可依据这五大秩序要素进行分类，并检验其与其他模块的一致性。

4.6 AGENTONOMICS 矩阵作为浓缩的管控工具

ADMRF 的核心可视化呈现形式是 AGENTONOMICS 矩阵。它交叉了两个维度：

- 横向为五大构建模块（智能体算法、智能体接口、AI 模型、工具与向量、生态系统协调机构）。
- 纵轴为代理生命周期的 DBOT 阶段（设计、构建、运营、拥有、移交）。

对于该矩阵中的每个单元格，将根据实际情况记录相关 KPI、CAPEX/OPEX 假设、风险及职责。因此，该矩阵同时也是：

- 一份施工图（哪个组件在哪个阶段），
- 一种预算工具（哪个阶段涉及哪些成本），
- 一种风险管理工具（哪个阶段存在哪些风险），
- 一种沟通工具（哪个阶段涉及哪些利益相关方）。

该矩阵将在第 9 章中详细阐述；在此需要明确的是，它应被视为所有模块的共同工作文档。在实际组织中，任何计划部署 AGENTONOMICS 代理的人员，都将通过该矩阵进行沟通——而不是通过六份彼此独立的文档。

4.7 首次应用：“AGENTONOMICS 副教授”

在 2025/26 冬季学期慕尼黑工业大学（TUM）的“工商管理中的 AI 代理”课程中，ADMRF 的各模块首次被介绍，并通过一个 AI 模型相互关联。AGENTONOMICS 算法将在此课程中得到应用，并通过案例分析逐步完善。这标志着“Doz. Dr. AGENTONOMICS”的诞生——这是基于该算法开发的第一个应用程序。

该应用是一个刻意设计得规模较小但定义明确的应用案例：一个人工智能代理，作为辅导员和教练，协助学生理解和应用 ADMRF。它既是学习对象，也是学习媒介——符合 Ries（2011）所定义的经典“最小可行产品”（MVP）。学生与“Doz. Dr. AGENTONOMICS”的互动所产生的反馈，将反哺算法的进一步开发。由此，这门课程本身也成为了其研究对象的一部分：它正在探索人工智能代理如何支持学习过程、组织知识并实现个性化的学习路径。

ADMRF 并非思考的终点, 而是其“手术室”。它旨在成为一个参考框架: 既足够完善, 以便学生、研究人员和从业者能够加以运用; 又足够简单, 以便他们能够对其进行严肃的批判。

参考文献

Berk, J.; DeMarzo, P. (2020): 《金融学基础: 分析、决策与实施》。第 6 版 (修订版), Pearson 出版社, 慕尼黑。

Bleicher, K. (2011): 《综合管理概念: 愿景——使命——计划》。第 8 版, 更新与扩充版, Campus 出版社, 法兰克福。

哈贝马斯, J. (1981): 《沟通行动理论》。两卷本, 苏尔坎普出版社, 法兰克福。

Hahn, D.; Hungenberg, H. (2001): 《PuK——以价值为导向的控制理念》。第 6 版, 修订与扩充版, Gabler 出版社, 威斯巴登。

川崎, G. (2015): 《创业的艺术 2.0: 经得起时间考验、实战淬炼的创业指南》。Portfolio 出版社, 纽约。

克诺尔, A. C.; 苏克富尔, H. (2024): 《从数据主权到技能主权》。技术报告 TUM-I24112, 慕尼黑工业大学计算机科学研究所。

里斯, E. (2011): 《精益创业》。Crown Business 出版社, 纽约。

吕格-施图姆, J.; 格兰德, S. (2020): 《圣加仑管理模型: 复杂世界中的管理》。第 3 版, 豪普特出版社, 伯尔尼。

Suckfüll, H. (1994): 《跨国组织模型》。博士论文, 圣加仑大学。

苏莱曼, M.; 巴斯卡尔, M. (2023): 《即将到来的浪潮》。克朗出版社, 纽约。

5 模块 1：AGENTONOMICS 定义框架

借助 AGENTONOMICS 框架对智能体进行定义，是 AGENTONOMICS 生命周期的起点。在此阶段，通过相应的参数描述，确定人工智能智能体在经济上应代表谁或什么，以便其能够作为经济主体运作。定义框架是企业层面的基本设定，类似于企业的宪法或初创企业的社会契约。其责任由董事会——即承担企业整体责任的集体——负责，而非由某个单一职能部门承担。

5.1 定义框架的功能

在经典的圣加仑学派理解中（Bleicher 2011；Rüegg-Stürm 和 Grand 2020），定义框架对应于管理的规范层面：在此层面，基本意义问题将得到解答。然而，与传统组织不同的是，对于人工智能代理而言，仅用散文形式制定企业宪章是不够的——这些设定还必须具备机器可读性和可验证性。该框架兼具这两项功能：它既是一份人类可读的规范性设定，也是一份机器可读的规范说明。

该框架回答了六个核心问题，每个问题均通过一个参数进行具体化：

1. 代理应具备何种程度的自主权和主权？
2. 它应具备哪些能力和资质？
3. 哪些监管和自我监管机制适用于它？
4. 它具备怎样的身份、伦理和个性？
5. 其基础采用哪种人工智能模型和具身形式？
6. 它应具备怎样的学习和教学能力？

首先应从定性角度描述这六个参数。在合理的情况下，可通过量表、指标或正式规范等手段对其进行定量细化。下文将逐一详细阐述每个参数。

5.2 参数 1：自主性与主权

在日常用语中，“自主性”和“主权”常被混淆或等同。但在 AGENTONOMICS 的语境下，必须严格区分二者，因为它们指代不同的概念，且对代理的构建具有不同的影响。

定义：自主性与主权

自主性指代理人在不受外部干预的情况下独立完成任务的程度。主权指代理人在多大程度上能够实际自由支配其被分配的资源——数据、能力、权利。一个代理人可能具有高度的自主性但主权较低——反之亦然。

示例：一个全自动翻译系统具有高度的自主性（无需人工干预即可进行翻译），但主权较低（它不得自行将翻译后的文本变现、转发或利用）。而拥有自主权的人类译员，其自主性可能较低（需要接受委托），但主权很高（可以自主决定价格、是否接受委托以及如何利用自己的作品）。

5.2.1 六个自主性等级（0-5）

自主程度的界定——通过一种在教学上颇具启发性的类比——借鉴了自动驾驶的六个 SAE 等级（0-5）（SAE 2021），并将其应用于经济活动领域：

- **第 0 级——无自主性**：智能代理仅作为工具存在。所有决策均由人类做出，智能代理仅负责执行。
- **第 1 级——辅助**：智能代理在严格限定的任务范围内提供建议；决策由人类做出。
- **第 2 级——部分自主**：智能代理独立执行已定义的子任务，人类负责监督并适时干预。
- **第 3 级——有条件的自主性**：智能代理在明确界定的业务范围内自主行动；在预定义的例外情况下，会请人类介入决策。
- **第 4 级——高度自主**：代理在其业务范围内完全自主行动；人工干预仅为例外情况。
- **第 5 级——完全自主**：智能体跨业务领域行动，自主设定目标并独立调整其业务范围。目前，该级别的讨论主要集中于理论界定。

5.2.2 主权及其影响

自主性是行为的属性，而主权则是处置权的属性。在 AGENTONOMICS 框架中，需明确主权的三个维度：

数据主权

代理可以使用、传递或保留哪些数据？在此方面，AGENTONOMICS 直接承接了数据托管讨论（Fraunhofer MOEZ 和 Suckfüll, 2018）：一个拥有主权的代理具备明确且可执行的数据权利。

能力主权

代理可以独立提供、授权或保留哪些能力？这正是 AGENTONOMICS 相对于传统数据市场的具体创新之处（Knoll 和 Suckfüll, 2024）。

决策主权

代理可以对哪些决策做出最终决定，而无需上报给人类机构？这一维度与自主程度直接相关。

5.3 参数 2：能力与胜任力要求

能力与胜任力应加以区分。骑自行车是一个很好的例子。虽然几乎所有身体健康的人都具备骑自行车的“能力”，但只有真正学会骑自行车的人才具备“胜任力”。将其应用到人工智能代理上：

定义：能力与胜任力

能力是指智能体执行某项任务的原则性可能性——植根于其模型、工具和具身性之中。胜任力则是经过实际检验并在有效应用中得到验证的该任务的执行能力。能力是潜在的，胜任力则是已实现的。

这种区分具有实际意义。基于基础模型（Bommasani 等, 2021）的智能体虽然在理论上拥有大量能力，但它仅会在那些经过实际训练、精细调整并在实际应用中得到验证的领域内发展出胜任力。因此，AGENTONOMICS 智能体的描述并非基于其模型在理论上允许的所有能力，而是基于其实际胜任力的特征。

关键的核查问题在于：哪些能力已获得认证，哪些仅是声称具备？AI 代理的可靠性在很大程度上取决于能否清晰地阐明这一区别。这正是关于 AI 保障（AI Assurance）的讨论所涉及的核心问题（NIST 2023；欧盟《AI 法案》2024）。

5.4 参数 3：监管与自我监管

所谓监管与自我监管，是指 AGENTONOMICS 代理通过技术、组织和法律措施来落实所有法律、伦理及其他规定的机制。在多中心网络中，原则上可区分两个作用层面：

代理间措施

在代理间关系层面，协议和标准确保了合作的基本可行性。这涉及接口、身份验证、身份证明、合同签订以及争议调解。其中许多措施属于技术性质，将在第 6 章（构建模块）中深入探讨。

与协调机构相关的机制

在代理与网络协调机构之间的层面上，跨域规则确保了凝聚力。这里涉及准宪法性规定：网络中究竟会协商哪些议题？哪些决策由协调机构专属决定？预设了哪些升级处理途径？第 7 章将对此层面进行详细探讨。

在此背景下，自我监管并非监管的对立面，而是其一种特定形式：代理会自行检查其行为是否符合规则，并在理想情况下主动进行纠正。这种自我约束形式在经济上具有吸引力（它降低了外部监管的交易成本），在伦理上也值得提倡（它将责任植根于行为主体自身）。但它同时也存在风险：一个隐瞒自身违规行为的行为主体，比完全缺乏自我约束的行为主体更为危险（Russell 2019）。

《欧盟人工智能法案》（*欧盟法规 2024/1689*）为监管层面提供了具体框架，该框架根据应用领域的风险等级提出不同要求。AGENTONOMICS 代理必须在定义框架中明确界定自身所属的风险等级——以及由此衍生的合规机制。

5.5 参数 4：身份、伦理与个性

身份、伦理与个性是人工智能代理的参数——乍看之下最令人意外的参数。但正是这些参数，将代理与任何软件组件区分开来。从营销角度来看，这可以称为代理的“独特销售主张”（USP）；从商业伦理角度来看，这关乎其行为的可靠性；从哲学角度来看，则涉及“我是谁，如果我是，又有多少个我？”这一问题——这提醒我们，可复制的代理尤其面临着这一尖锐的挑战。

5.5.1 身份

此处的“身份”并非指哲学意义上的自我认同，而是指代理在经济和法律层面的可识别性。它包括：

- 一个唯一的技术地址（标识符），
- 可归属的经济责任（所有者、负责的个人或组织），
- 可认证的来源（模型和训练数据的来源），
- 一种可识别的“声音”，即具有可辨识的通信模式。

5.5.2 伦理

根据综合经济伦理学（Ulrich 2008）以及人工智能伦理学（Mainzer 2019; Floridi 2023）的概念，伦理学是指将规范性标准植根于智能体的决策机制之中。可区分三个层面：

- 合规伦理：代理人应遵守现行法律。这是最低要求，而非最高标准。
- 利益相关者伦理：行为主体应考虑所有相关利益相关者的合法利益，而不仅仅是其委托人的利益。
- 话语伦理：代理人应愿意为其决策提供理由，并愿意参与讨论，即使在此过程中自己也可能犯错。这与哈贝马斯（1981）的观点以及康德的“定言令式”一脉相承。

5.5.3 个性

个性包括稳定的行为模式，利益相关者正是通过这些模式来识别代理人的。个性并非花哨的装饰，而是具有经济价值的：它降低了合作的交易成本，因为利益相关者能够可靠地了解对方的真实情况。个性包括：

- 沟通风格（正式、平易近人、客观、风趣），
- 决策风格（规避风险、机会主义、着眼长远），
- 冲突处理方式（愿意妥协、强硬、调解），
- 价值观特征（创新、可靠、谨慎等）。

5.6 参数 5：AI 模型与具身化

AI 模型是智能体的“引擎”。模型的选择是一项具有深远经济影响的战略决策。需要明确以下三个维度：

模型类型与模型规模

基础模型（Bommasani 等，2021）具有广泛的应用性，但资本密集度高，且会产生对供应商的依赖。专用模型运营成本更低且更具自主性，但应用范围较窄。这一选择始终涉及“自研还是采购”的决策——同时需考虑 AGENTONOMICS 特有的附加标准：能力自主权（Knoll 和 Suckfüll 2024）。

本地托管与集中托管

模型运行在哪里？是在自有服务器上、超大规模云服务商的云端，还是欧洲的基础设施上？答案将影响延迟、成本、数据保护，以及——尤为重要——地缘政治依赖性。在此，主权问题同样至关重要。

具身性

- 具身性指代理在现实世界中存在的形式。其理想类型主要有三种：纯虚拟具身性：代理以软件形式存在于云端，通过文本或语音接口进行通信。
- 混合具身：代理控制物理执行器——如打印机、物联网设备、车队——但自身仍处于数字领域。
- 物理具身：智能体集成于机器人之中，该机器人可直接进行物理操作。这是慕尼黑工业大学（TUM）机器人研究的经典领域。

具身性具有经济影响，因为它扩展了代理经济行为的范围。一个具有物理具身性的代理可以完成纯虚拟代理无法做到的事情——但投资和维护成本更高，法律责任也更大。

5.7 参数 6：学习与教学——元代理

第六个参数描述了代理自主学习的能力，以及——这正是 AGENTONOMICS 的核心要义——教导其他代理的能力。这体现了“智人（Homo Sapiens Sapiens）”的理念，即一种不仅具备智慧，而且知道自己具备智慧，并能将这种知识传承下去的智能。

5.7.1 学习能力

学习模式可分为三种：

- 预训练：基于海量数据集进行的耗时初始训练，该训练构成了基础模型。通常由模型提供商完成，而非代理运营商。
- 微调：针对应用场景和个性特征数据对模型进行有针对性的调整。
- 情境学习与工具使用：在应用过程中进行的情境化学习——例如，代理在运行时调用提供的文档、数据库或工具。

5.7.2 教学能力与子代理

学习能力是元代理的核心特质。具备学习能力的代理能够对其他代理进行训练——通过生成训练数据、定义子任务，以及为特定任务创建自己的“子代理”。代理是否可以创建自己的子代理，这一问题必须在定义框架中明确说明——它具有重大的治理影响。谁对子代理负责？子代理中存储的知识归谁所有？如何确保控制权？

在此，该参数与“技能主权”的讨论直接相关：当一名人类知识工作者将其技能转移给一个智能体，而该智能体又进一步生成子智能体时，必须明确这些倍增的技能归谁所有（Knoll 和 Suckfüll 2024）。

5.8 一致性检验与董事会视角

这六个参数并非相互独立。只有当各项设定相互协调时，才算得上是一个一致的定义。以下列举三种典型的矛盾情况：

- 一个自主度为 4 但未指定身份的智能体：若其造成损害，应由谁承担责任？
- 一个拥有全面能力主张但仅具备小型专用模型的智能体：其承诺无法兑现。

- 一个拥有高度主权但缺乏监管框架的代理：利益相关者不会信任它。

AGENTONOMICS 算法的核心任务之一，就是发现此类不一致之处，并提交给人类董事会进行决策。这种一致性检查，正是 AGENTONOMICS 对 Bleicher (2011) 所倡导的“整合”概念的具体化实现。

定义阶段是董事会的工作。所有视角——CEO、CFO、CTO、CSO——都必须共同参与。CEO 负责愿景和规范设定，CFO 负责经济可行性，CTO 负责技术可行性，CSO 负责利益相关方的接受度。如果忽略了其中任何一个视角，定义就会不完整——代理将在生命周期中失败。

定义框架绝非项目初期一项枯燥的例行公事，而是至关重要的决策。若在此环节马虎敷衍，无异于在沙地上建房。

参考文献

Bleicher, K. (2011): 《集成管理概念》。第 8 版，Campus 出版社，法兰克福。

Bommasani, R.; Hudson, D. A.; Adeli, E. 等 (2021) : 《关于基础模型的机遇与风险》。
arXiv:2108.07258。

《欧盟人工智能法案》(2024) : 欧洲议会和理事会关于制定人工智能统一法规的第 2024/1689 号条例 (EU) 。《欧盟公报》。

Floridi, L. (2023): 《人工智能伦理：原则、挑战与机遇》。牛津大学出版社，牛津。

弗劳恩霍夫 MOEZ; 苏克富尔, H. (2018) : 关于建立数据托管机构 DEDATE 的倡议。讨论文件，莱比锡。

哈贝马斯, J. (1981) : 《沟通行动理论》。苏尔坎普出版社，法兰克福。

克诺尔, A. C.; 苏克富尔, H. (2024) : 《从数据主权到技能主权》。技术报告 TUM-I24112, 慕尼黑工业大学计算机科学研究所。

迈因策, K. (2019) : 《人工智能——机器何时接管世界?》第 2 版，斯普林格出版社，柏林和海德堡。

NIST (2023) : 《人工智能风险管理框架》(AI RMF 1.0) 。美国国家标准与技术研究院，马里兰州盖瑟斯堡。

Rüegg-Stürm, J.; Grand, S. (2020): 《圣加仑管理模型》。第 3 版, Haupt 出版社, 伯尔尼

。

Russell, S. (2019): 《与人类兼容: 人工智能与控制问题》。Viking 出版社, 纽约。

SAE International (2021): SAE J3016——驾驶自动化等级。汽车工程师学会。

苏莱曼, M.; 巴斯卡尔, M. (2023) : 《即将到来的浪潮》。克朗出版社, 纽约。

乌尔里希, P. (2008) : 《综合经济伦理学》。第 4 版, Haupt 出版社, 伯尔尼。

6 模块 2: AGENTONOMICS 代理的五大构建模块

第 5 章将定义框架作为企业层面的设定进行了阐述，而本章则聚焦于具体的构建实现。AGENTONOMICS 代理并非一个整体式的软件系统，而是由五个界限分明的构建模块组成的模块化结构。这种模块化并非出于教学上的简化，而是技术、经济和法律上的必要：它使得每个构建模块能够独立进行采购、维护、评估和责任划分。想要构建一个代理，首先必须理解：需要构建、购买或许可五个组件——而且每个构建模块的选择都会产生经济后果，其影响远超出技术范畴。

在将圣加仑管理模型转化为人工智能代理经济时，这五个构建模块对应于 Rüegg-Stürm 和 Grand（2020）所描述的企业“流程”和“结构”——关键区别在于，这些流程和结构不再用于组织人类员工，而是以代码、模型和接口的形式具象化呈现。Rüegg-Stürm 所说的组织结构图，在 AGENTONOMICS 中变成了架构图；流程模型则变成了算法。圣加仑学派思想的实质得以保留；其表现形式则发生了根本性的转变。

6.1 五大构建模块概览

这五大构建模块可概括如下（顺序不分先后）：

1. 代理算法：决策与协调逻辑。
2. 代理接口：对外和对内的接口。
3. AI 模型：作为认知引擎的语言或多模态模型。
4. 工具与向量：工具及存储的知识。
5. 代理生态系统协调机构：与网络环境的连接。

每个模块都由一个专属徽标代表；这些徽标共同构成了代理的图形标识。在 AGENTONOMICS 标准中，这种结构是强制性的，以确保代理能够基于开放接口相互协作。如果一个代理没有明确的构建模块结构，即使其功能表现相当，也不符合 AGENTONOMICS 标准。该标准并非一种姿态，而是由单个代理构建真正互操作、多中心组织网络的前提条件。

6.2 构建模块 1：代理算法

代理算法是代理的决策和协调逻辑。它回答了这样一个问题：代理在何种情况下该做什么？与将输入转换为输出的 AI 模型不同，该算法控制着查询模型的过程——何时查询、使用哪些提示、按什么顺序、使用哪些工具。用传统企业管理学的术语来说，该算法就是代理的流程组织。

6.2.1 功能组成部分

一个完整的代理算法至少包含以下功能组件，每个组件都需要做出相应的设计决策：

感知与输入处理

代理如何接收输入、对其进行分类，并决定采取哪些处理路径？这涉及路由、意图检测以及在输入中初始添加上下文信息。一个看似简单却常被低估的问题是：如果代理无法理解一个请求，会发生什么？它会保持沉默吗？会请求澄清吗？还是会将问题上报给人工处理？

规划与任务分解

智能代理如何将复杂任务分解为子任务？现代智能代理架构，如推理与行动（Reasoning-and-Acting, ReAct, Yao 等人，2023）或计划与执行（Plan-and-Execute），将模型与反思和行动的循环相结合。该算法确定允许多少次迭代、何时终止以及如何评估中间结果。

工具调用与使用

代理可以使用哪些工具——数据库查询、网络搜索、计算操作、外部 API？它如何决定何时使用哪种工具？它如何处理错误和不一致的情况？工具使用是区分代理与纯语言模型的关键扩展（Schick 等，2023）。没有工具，模型就会被困在语言的内部空间中；有了工具，它便能步入现实世界。

记忆与状态管理

代理在交互之间会保留哪些信息？哪些会被写入长期存储，哪些会被遗忘？这里涉及一个在经济和伦理上都极为敏感的问题：拥有完美记忆的代理虽然强大，但也充满风险——尤其是当它将来自不同委托方的信息混为一谈时。

自我反思与质量保证

代理如何评估自己的输出结果？它会调用第二个模型作为评审者吗？它会比较多个候选答案吗？自一致性、反思和批判机制都是成熟的技术（Madaan 等，2023），但每种机制都会消耗计算时间和资金。算法将决定何处值得投入这些资源。

6.2.2 自研与外包决策

代理算法是唯一几乎总是属于代理运营商自有知识产权的构建模块。虽然模型、接口和工具通常是采购而来的，但算法却是其自主创作——即构建模块的特定组合方式，代理的价值创造正是在此基础上得以体现。从这一角度来看，算法才是代理运营商的真正资本。一旦失去对算法的控制，就意味着失去了对代理的控制——进而失去了对其创造价值的掌控。

实践提示：将算法视为资产

运营 AGENTONOMICS 智能代理的用户，应将算法进行文档化、版本管理，并将其视为无形资产。该算法应计入资产负债表——尽管德国商法对于自创的无形固定资产规定了资本化选择权（但存在例外情况）。在此方面，经济现实已领先于资产负债表法规。

6.3 模块 2：代理接口

接口是指代理与其环境进行通信的所有接口的总和。从经济角度来看，这是价值创造得以显现的场所，因为这里是代理与委托人、利益相关者及其他代理进行交互的场所。接口可分为三类：

6.3.1 人-代理接口

这里涉及与人类用户的直接交互。语言——无论是文本还是语音——如今已成为主导媒介，辅以图形元素、仪表盘，以及日益普及的多模态输出。该界面的设计不仅涉及用户体验（UX）

问题，更涉及管理理论：若想将智能代理作为首席执行官的替代者（参见第 8 章），就必须设计出能够满足董事会会议要求的界面——采用恰当的语言、保持适当的保密性，并具备合理的错误反馈机制。这与客户服务聊天机器人的要求截然不同。

6.3.2 代理-代理接口

如果代理之间需要协作，就需要一套协议。AGENTONOMICS 标准正是为此而设：它定义了代理如何进行身份识别、描述自身能力、交换任务以及反馈结果。这里的接口主要不是图形化的，而是语义化的：它涉及对服务、价格和条件的机器可读描述。当前及未来的发展，如模型上下文协议（MCP）和代理间标准 A2A，构成了该领域的技术基础；AGENTONOMICS 则在此基础上补充了经济层。

6.3.3 代理-系统接口

代理必须集成到现有系统中——如 ERP、CRM、银行 API、政府门户等。这决定了其是否适用于现有系统（Brownfield）：一个只能在全新系统（Greenfield）中运行的代理，对大多数现实企业而言毫无价值。接口开发工作通常占实施工作量的 60%至 80%——这一情况在成本核算中往往被低估。

6.4 模块 3：AI 模型

AI 模型是智能代理的认知引擎。它将输入转换为输出，从而提供使智能代理显得“智能”的性能。在构建智能代理的过程中，模型的选择是迄今为止最昂贵的决策——无论从金钱还是战略角度来看都是如此。

6.4.1 模型类别

目前有三种典型的模型类别可供选择，每种都有其独特的经济影响：

前沿基础模型

规模最大、性能最强的模型——即专有供应商的“前沿”模型（例如 GPT、Claude 或 Gemini 系列）——具有最广泛的应用范围，但运营成本高昂，且会导致对少数大型供应商的依

赖。使用这些模型的用户虽然能够接触到尖端技术，但也因此放弃了一部分自主权。从欧洲的角度来看，这是在主权政策方面最为敏感的一种选择（Knoll 和 Suckfüll 2024）。

开源基础模型

像 Llama 或 Mistral 这样的模型可以自主运行，并根据自身需求进行调整。它们的运营成本较低（在完成基础设施投资后），但在许多基准参数方面，通常比前沿模型落后一代。不过，对于许多用户而言，这一代模型已经完全足够——而且在主权方面可能获得显著提升。

专业化与精细调优模型

针对特定领域训练的小型模型，在其专业领域内的表现可能优于大型通用模型——且成本仅为后者的零头。这种模式预计将在 AGENTONOMICS 世界中占据主导地位，因为它最能满足“技能主权”的要求：想要作为代理提供个人技能的人，需要一个能够精准实现这些特定技能的模型——而非包罗万象的通用模型。

6.4.2 经济指标

有四项指标对模型选择具有决定性的经济意义，应记录在 AGENTONOMICS 矩阵中（参见第 9 章）：

- 每百万代币的成本——直接的可变成本基础。不同模型之间的成本差异可达 100 倍，而质量差异有时却微乎其微。
- 延迟——从请求到响应所需的时间。这对交互式应用至关重要，但对批处理则无关紧要。
- 上下文窗口——代理能够同时处理的信息量。当前顶尖模型已能处理数百万个令牌，这在以往被视为难以企及。
- 主权指数——由作者提出的 AGENTONOMICS 专属指标，反映代理在多大程度上能够独立于模型提供商运行。开源且本地托管的模型其主权指数接近 100；而通过专有 API 调用的 Frontier 模型，其指数则接近零。

6.5 模块 4：工具与向量

工具和向量是区分智能体与纯语言模型的关键所在。工具是智能体在现实世界中采取行动的手段；向量则是其所依赖的有组织知识。

6.5.1 工具：行动的工具

从代理的角度来看，工具是一种可被调用的函数，用于实现特定效果。示例：

- 数据库查询：查询内部数据
- 网络搜索：检索最新信息
- 计算运算：精确的数学分析
- API 调用：与外部服务交互
- 代码执行：生成并执行程序
- 文档生成：生成报告、合同、演示文稿

工具的选择是一项战略决策。它决定了代理对现实世界的干预程度——进而也决定了由此产生的责任和安全风险。一个能够实际执行转账操作的代理，在经济和法律层面上与一个仅能建议转账操作的代理是截然不同的主体。

6.5.2 向量：存储的知识

此处所指的“向量”并非数学结构本身，而是将知识以向量化表示形式提供给代理的实践。具体而言，这指的是向量数据库，其中文档、知识产物或经验值以特定方式存储，使代理能够通过语义相似性搜索访问这些内容。这项技术——已被称为“检索增强生成”（RAG）——对 AGENTONOMICS 至关重要，原因有二：

知识主权

向量代表着智能体的特定知识。与代表共同常识的基础模型不同，向量是智能体的个体记忆。若要运行一个应用圣加仑管理模型的智能体，就需要将相关文献——如布莱希尔（Bleicher）、吕格-施图姆（Rüegg-Stürm）、乌尔里希（Ulrich）、苏克富尔（Suckfüll）、诺尔（Knoll）等——存储在向量中，从而将价值创造的智力核心封存其中。

时效性

基础模型 (Foundation Model) 通常有一个知识截止日期 (往往比实际应用早数月), 而向量则可以持续更新。新的研究、新的市场数据、新的法律规定都能直接融入其中, 无需耗时的模型训练。因此, 向量正是让智能代理保持活力的关键要素。

对于一个旨在替代首席执行官的智能代理而言, 向量是至关重要的资源。这里存储着关于企业、其市场、合同以及利益相关者的知识。这些知识必须保持完整、最新且一致——否则, 智能代理将基于过时或不完整的信息做出首席执行官级别的决策。参见第 8 章。

6.6 模块 5: 生态系统协调机构

第五个构建模块是将代理与周围的代理和利益相关者生态系统连接起来——即生态系统协调机构。它是代理进入网络、进行身份识别、签订合同以及维护价值创造关系的连接技术。由于协调机构是一个独立的概念 (将在第 7 章中详细讨论), 本章仅限于探讨其在单个代理内部所扮演的角色。

从单个代理的角度来看, 协调机构是连接两个世界的接口: 一个是其他代理构成的水平网络, 另一个是由委托人、所有者和监管机构组成的垂直控制层级。协调机构为代理提供身份验证、声誉数据、合同模板、争议调解机制, 以及访问知识库、训练数据或计算能力等共享资源的途径。

从设计角度来看, 这意味着: 在构建代理时, 必须从一开始就考虑将其连接到哪个协调机构。这一选择的影响, 与创业时选择银行或税务顾问同样深远。在 AGENTONOMICS 标准中, 协调机构接口的规范是开放的; 理论上, 代理可以在不同协调机构之间进行切换。但在实践中, 由于数据积累和关系维护, 会产生不可低估的锁定效应。

6.7 生命周期中构建模块的协同作用

这五个构建模块并非静态的, 而是随着代理人的生命周期不断演进。DBOT 逻辑——设计 (Design)、构建 (Build)、运营 (Operate)、拥有 (Own)、移交 (Transfer) ——为这一演进提供了框架, 并构成了 AGENTONOMICS 矩阵 (第 9 章) 的纵轴。

设计

在设计阶段，将对这五个构建模块进行规划。该阶段将产出架构图、接口规范、模型选型文件以及工具清单。设计决策在很大程度上是可逆的，但在后续阶段进行的任何修订成本都会越来越高。

构建

在构建阶段，将对构建模块进行实施或采购。此阶段的投资（CAPEX）最大。典型活动包括：算法实现、接口编程、模型授权或微调、矢量数据库构建、协调机构系统对接。

运行

在运行阶段，智能代理进入正式运行状态。运营成本（OPEX）占据主导地位：模型使用、工具调用、维护、安全更新、利益相关方沟通。该阶段将决定项目的经济可行性。

拥有

“拥有”阶段着眼于长期资产视角：代理将成为资产负债表中的永久组成部分。所有权归属谁？哪些价值应计入资产负债表？如何进行折旧？该阶段在现有实践中尚不成熟；AGENTONOMICS 正积极推动其发展。

转让

最后，代理可以被出售、分割、关闭或转移到其他场景中。在此阶段，同样需要将各构建模块分开考虑：有些是可转移的（算法、向量），有些则保留在原始场景中（模型特定的调整、接口契约）。

由五个构建模块和五个生命周期阶段组成的交叉表共包含 25 个字段。该表构成了 AGENTONOMICS 矩阵的骨架，并填充了相应的 KPI、CAPEX/OPEX 假设及职责后，便成为 AGENTONOMICS 项目的核心管控文件。

这五个构建模块相当于传统企业中的职能部门——只不过在这里，它们以代码、模型和接口的形式呈现。如果不能将它们清晰地区分开来，构建出来的就不是智能代理，而是一团乱麻。

参考文献

Bleicher, K. (2011): 《集成管理概念》。第 8 版, Campus 出版社, 法兰克福。

Knoll, A. C.; Suckfüll, H. (2024): 《从数据主权到技能主权》。技术报告 TUM-I24112, 慕尼黑工业大学计算机科学研究所。

Madaan, A.; Tandon, N.; Gupta, P. 等 (2023) : 《Self-Refine: 基于自我反馈的迭代精化》。载于: 《神经信息处理系统进展》第 36 卷。

Rüegg-Stürm, J.; Grand, S. (2020): 《圣加仑管理模型》。第 3 版, Haupt 出版社, 伯尔尼。

Schick, T.; Dwivedi-Yu, J.; Dessì, R. 等 (2023) : Toolformer: 语言模型能够自学使用工具。载于: 《神经信息处理系统进展》第 36 卷。

Yao, S.; Zhao, J.; Yu, D. 等 (2023) : ReAct: 语言模型中推理与行动的协同作用。《国际学习表征会议》。

7 模块 3：网络协调机构

在第三个模块中，一个在经典的圣加仑管理模型中仅被初步提及的概念成为焦点：网络协调机构。AGENTONOMICS 在此明确延续了圣加仑学派的思维，提出了一个问题：一个由自主代理构成的多中心网络如何既能发展凝聚力，又能保持个体的主权。答案是：通过一个不统治、而是进行协调的协调机构——不进行规划，而是开辟空间。

这一构想是 AGENTONOMICS 对 Rüegg-Stürm 和 Grand (2020) 所称“秩序要素”的独特诠释，将其转化为人工智能代理经济中的现实。在以人为主导的企业中，由文化、战略和结构所实现的功能，在代理网络中则由协调机构来实现——作为一种在技术和经济层面构建的机构，它缔造了多中心的关系网络。

7.1 为何需要协调机构

多中心网络存在一种根本性的张力：它们既希望维护成员的主权，又希望实现集体绩效。这种张力无法通过等级制度来化解——否则网络就不再是多中心的，而是单中心的了。它也无法通过纯粹的自发秩序来化解——否则网络就会定期陷入无人能调解的冲突之中。因此需要一个在两极之间进行调解的第三种机构：协调机构。

这一论点基于三条理论脉络，它们在“AGENTONOMICS”概念中汇聚：

7.1.1 交易成本经济学

科斯（1937）和威廉姆森（1985）指出，市场和等级制度的效率因交易特征的不同而有所差异。在这两个极端之间存在混合形式——长期合同、合资企业、网络。当交易具有中等特异性、中等不确定性和中等频率时，这些混合形式便更具优势。“AGENTONOMICS”协调机构正是组织代理人之间此类混合交易的制度形式。

7.1.2 沟通性行动理论

哈贝马斯（1981）将行动分为战略性行动和沟通性行动。战略性行动旨在取得成功，而沟通性行动则旨在达成共识。一个仅基于战略性行动的社会将会瓦解；它需要沟通性行动的空间，以便在话语中检验各种主张的正当性。若将其应用到行动者网络中：一个由纯粹进行战略性

行动的行动者组成的网络，将沦为机会主义的博弈。协调机构者开辟了一个空间，使行动者能够通过沟通来协商其主张——即反思事实、规范和真诚，而不是仅仅追求最大化。

7.1.3 多中心治理

奥斯特罗姆（1990）阐明了多中心治理结构如何在没有中央层级体系和纯粹市场的情况下，成功管理公共资源。奥斯特罗姆提炼出了八项设计原则，从明确的边界、集体决策程序到分级制裁。AGENTONOMICS 协调机构将这些原则应用到人工智能代理的世界中。

基于这三条理论脉络，我们可以得出如下论证：该协调机构既非行政官僚机构，也非平台垄断，而是一种特定形式的多元中心治理，它既能促成主权代理之间的混合交易，又能提供一个沟通协商的空间。

7.2 物业管理类比

房产业主委员会为理解“协调机构”提供了一个富有启发性的类比。业主委员会中的公寓业主是各自公寓的主权所有者。他们共同拥有公共财产——楼梯间、屋顶、外墙、供暖系统。他们并不亲自管理这些共同事务，而是将其委托给物业管理公司。物业管理公司并不支配业主；它只是执行业主的决议，管理共有财产，并在对外事务中代表业主共同体。

AGENTONOMICS 协调机构正是以此职能为定位。代理人仍保持其独立的经济主体地位。对于他们无法或不愿自行处理的事务——身份登记、声誉体系、纠纷调解、共享知识库、合规监管——则委托给协调机构。协调机构对代理人负责，而非相反。

定义：AGENTONOMICS 协调机构

AGENTONOMICS 协调机构是在多中心代理网络内一个制度化的协调机构，它受网络成员委托并在其监督下履行集体职能——同时不损害各代理在其核心业务中的自主权。

物业管理的类比还有第二个富有启发性的要点：物业管理公司是可以更换的。如果业主委员会对服务不满意，就会解除管理合同并另寻新公司。AGENTONOMICS 网络也必须如此：协调机构不得拥有垄断地位，而必须通过竞争来争取代理人的青睐。否则，就会面临原本服务型协调机构演变为统治性垄断的风险——正如我们今天在大型平台经济中观察到的情况（Zuboff 2018）。

7.3 协调机构的功能

协调机构具体履行哪些职能，取决于网络的具体情境。从理想类型来看，可以区分出六个职能领域：

7.3.1 身份与声誉管理

协调机构负责维护参与者的身份登记册。每位参与者均以其定义（参见第 5 章）、其构建模块（参见第 6 章）及其负责人信息在此登记。随着时间的推移，声誉逐渐积累：包括已完成交易的评价、审计结果以及投诉记录。在 AGENTONOMICS 网络中，声誉是一项首要的经济指标——它能降低合作中的交易成本，并激励各方遵守规则。

7.3.2 合同与交易基础设施

协调机构为代理之间的典型交易提供合同范本、智能合约模板和结算机制。它承担托管职能，保障支付安全并分散风险。在此方面，AGENTONOMICS 直接承接了数据托管讨论（Fraunhofer MOEZ 和 Suckfüll 2018；DEDATE 2.0），并延续了德国托管机构的传统——这些机构在德国经济史上曾发挥过重要作用，如今正迎来复兴。

7.3.3 争议调解

代理人的互动往往会引发冲突。协调机构提供的争议调解程序比国家法院更快、更经济且更贴近案件实情。理想情况下可区分三个阶段：

- 调解：协调机构在争议双方之间进行斡旋，但无决策权。
- 仲裁程序：协调机构依据代理方事先同意遵守的仲裁规则作出裁决。
- 诉诸国家法院：涉及宪法问题或超出仲裁管辖范围的事项。在此情况下，协调机构将提供程序性协助。

7.3.4 知识与资源共享

某些资源既极其宝贵，又因获取成本高昂，以至于任何单一代理人都无法单独合理地备有这些资源。例如：最新法律文本、市场数据、基准研究、训练数据集。协调机构将这些资源作为共享基础设施提供，并通过成员会费进行资助——这与传统经济中的图书馆系统或行业协会类似。

7.3.5 合规要求

《欧盟人工智能法案》（2024）、《通用数据保护条例》（GDPR）以及各行业特定法规提出了一系列要求，对于单个参与者而言，满足这些要求往往成本高昂，难以承受。协调机构可整合合规职能：持续审计、向主管部门报告、培训以及监管变化预警系统。由此，它承担起在协会结构中早已确立的职能——只不过是针对人工智能代理这一新类别。

7.3.6 市场平台功能

最后，协调机构可以提供市场平台，供代理商在此提供和寻求服务。市场价格由此形成，而价格将成为人工智能能力经济中的核心协调指标。重要提示：协调机构负责运营市场平台，但自身并非该市场的参与者。这种“市场运营”与“市场参与”之间的分离，是协调机构可信度的根本所在。

7.4 协调机构的治理

谁来监管协调机构？这个问题决定了它究竟是作为服务型机构运作，还是会演变为统治型平台。目前有三种治理模式可供讨论，它们具有不同的风险特征。

7.4.1 合作社模式

协调机构是由其成员——即代理方及其个人或组织主体——组成的合作社。无论规模或交易量大小，每位代理方均拥有一票表决权。优点：与成员利益紧密相连、具有民主合法性、锁定激励较少。缺点：决策过程可能较为迟缓，资金配置可能不足。

7.4.2 基金会与信托模式

协调机构是一个专项基金会或信托机构，其董事会根据章程有义务维护成员的利益。优点：稳定性高、用途明确、可防范敌意收购。缺点：成员的直接参与度较低，改革能力取决于章程的设计。

7.4.3 混合型和联邦制模式

多个相互竞争的协调机构并存，并通过标准实现互操作性连接。代理可以更换协调机构；标准可防止在此过程中造成价值损失。这种模式最接近物业管理类比，且从 AGENTONOMICS 的角度来看，是中期内应优先采用的形式。它允许协调机构之间进行竞争，同时避免滑向垄断性平台。

7.4.4 绝不能发生的情况

过去二十年数字平台经济的发展历程提供了一个鲜明的反面案例：那些最初作为有益协调机构的机构，已演变为垄断市场的主导力量，其经济实力如今几乎无法通过民主手段加以制约。AGENTONOMICS 标准通过构成性规则堵住了这条道路：

- 强制数据可移植性——每个代理都可以携带自己的数据和声誉。
- 互操作性作为标准——代理可在不同协调机构之间自由切换。
- 市场运营与市场参与职能分离。
- 对成员承担透明度义务。

7.5 德国与盎格鲁-撒克逊治理风格的比较

协调机构在不同经济文化中的制度定位各不相同。三种模式相互对立——而这种选择并非中立的，而是反映了对经济、社会与国家之间关系的根本观念。

7.5.1 德国模式

德国的经济传统拥有种类繁多的协调机构：商会、合作社、行会、储蓄银行、公法广播机构。这些机构通常依据公法设立，或被赋予自治权。德国的“AGENTONOMICS”协调机构显然会融入这一传统——既具备相应的优势（稳定性、信任、辅助性原则），也存在相应的弱点（官僚主义、抗拒改革）。

7.5.2 盎格鲁-撒克逊模式

在盎格鲁-撒克逊地区，私营且通常以盈利为导向的平台占据主导地位，这些平台受到竞争法和行业监管的约束。其优势在于速度、创新能力和可扩展性。其缺点则是众所周知的市场集中趋势以及平台与用户之间的权力不对称。盎格鲁-撒克逊式的 AGENTONOMICS 协调机构预计将以风险投资（VC）资助的初创企业起步，并——若取得成功——最终发展成为主导性平台。

7.5.3 欧洲混合模式

欧洲模式融合了两种模式的要素：采用私营法律形式，但平台运作逻辑受到监管约束。《数字市场法案》、《数据法案》和《人工智能法案》正是这一混合路线的体现。对于基于欧洲模式的“代理经济学”协调机构而言，这意味着：采用私营经济组织形式，但受明确的监管框架约束——数据主权、互操作性强、竞争开放。这符合 Knoll 和 Suckfüll (2024) 明确倡导的政治方针。

关于在德国的具体实施，有多个论点支持将其植根于合作社或受托人模式。德国的受托人制度历史（Suckfüll 1994，在另一语境下提出）以及德国的合作社传统（赖夫艾森、舒尔策-德利茨施）提供了制度范本，AGENTONOMICS 可以——也应当——以此为蓝本。

协调机构是“代理经济学”超越经典圣加仑模型所做出的最重要创新。它是多中心网络既能保持自主性又能保持一致性的制度形式。

参考文献

科斯，R. H. (1937)：《企业的性质》。载于：《Economica》第 4 卷第 16 期，第 386–405 页。

欧盟《人工智能法案》（2024）：条例（EU）2024/1689。《欧盟官方公报》。

弗劳恩霍夫 MOEZ；苏克富尔，H. (2018)：关于建立数据受托机构 DEDATE 的倡议。讨论文件，莱比锡。

哈贝马斯，J. (1981)：《沟通行动理论》。苏尔坎普出版社，法兰克福。

克诺尔, A. C.; 苏克富尔, H. (2024) : 《从数据主权到技能主权》。技术报告 TUM-I24112, 慕尼黑工业大学, 慕尼黑。

奥斯特罗姆, E. (1990) : 《公共资源的治理: 集体行动制度的演变》。剑桥大学出版社, 剑桥。

吕格-施图姆, J.; 格兰德, S. (2020) : 《圣加仑管理模型》。第 3 版, 豪普特出版社, 伯尔尼。

苏克菲尔, H. (1994) : 《跨国组织模型》。博士论文, 圣加仑大学。

威廉姆森, O. E. (1985) : 《资本主义的经济制度》。自由出版社, 纽约。

祖博夫, S. (2018) : 《监控资本主义时代》。Campus 出版社, 法兰克福。

8 模块 4：AI-CEO 管理理念

随着第四模块的展开，我们来到了本研究的核心部分。前几个模块定义、构建了代理，并将其嵌入网络之中；而现在要探讨的问题是：一个 AGENTONOMICS 代理如何能够承担首席执行官的职能——不是作为助手，也不是作为准备工具，而是在人类组织责任框架下，作为决策支持型或可授权的代理。这一论断虽颇具争议，但绝非空想：如今，技术、经济及治理方面的（而非法律方面的！）先决条件已触手可及。本章将阐述如何具体构思和规划这一转型。

“AI-CEO 管理理念”是将圣加仑整合管理理念（Bleicher 2011）系统性地移植到人工智能代理时代的产物。布莱希尔将整合功能定位于人类高管的认知层面，而在 AI-CEO 概念中，这一整合任务则由人工智能代理承担——其深度、速度和一致性是人类管理无法企及的，因为人类管理受制于生物学上的局限。人类 CEO 并未因此变得多余；他将成为 AI-CEO 的所有者、委托人以及伦理监督者。但整合性管理工作中的运营负担可以转移给智能代理。

8.1 人工智能代理时代下的圣加仑管理模型

自乌尔里希（Ulrich, 1968）以来，圣加仑管理模型已经历了三大代演变。我们通过“AI 首席执行官”概念勾勒出的下一代模型，并非源于对以往思想的摒弃，而是对其的延续与发展。圣加仑模型在人工智能背景下的演进呈现出三大转变：

8.1.1 从认知整合到算法整合

布莱希尔（2011）的核心主张是整合规范性、战略性和运营层面。在人类管理中，这种整合需要极高的认知能力，但在实践中往往因能力极限而屡屡受挫。作为一家中型股份公司的首席执行官，若想同时兼顾所有相关的战略计划、合规要求、利益相关者的期望、市场变化以及文化影响，必然会触及人类工作记忆的极限。其结果便是众所周知的现象：宣示与实践之间的不一致、利益相关方认知中的盲区、对微弱信号反应迟缓，以及在日常运营中逐渐流于形式的战略。

人工智能首席执行官则不存在这种容量限制。它能够持续访问并交叉核对数十万页的内部文档、外部市场数据、监管文本及利益相关方沟通记录。布莱希尔所提出的目标——即所有管理决策的贯穿始终的一致性——因此首次在技术上得以实现。

8.1.2 从“单向沟通型”企业到“协同沟通型”企业

Rüegg-Stürm 和 Grand (2020) 强调了管理中的沟通维度：企业是由沟通实践构成的网络。在圣加仑模型的 AI 版本中，新增了一个沟通层面：人类行为者与 AI 代理之间的沟通，以及代理之间相互之间的沟通。现代企业不再仅仅是人类沟通的网络，而是由人类与智能体沟通构成的混合网络。这种混合性并非缺陷，而是资源：它成倍提升了企业的沟通能力，并使协商过程能够以人类单独无法企及的速度和精细程度进行。

8.1.3 从人类意义构建到混合意义构建

一个特别微妙的转变涉及规范层面。布莱希尔 (Bleicher) 和吕格-施图姆 (Rüegg-Stürm) 始终将意义问题——愿景、价值观、责任——定位为人类的任务。AGENTONOMICS 坚持这一定位：企业的终极意义并非源于模型，而是源于支撑企业的人。但是：规范性设定的 operationalisierung 和贯彻工作可以由 AI-CEO 来承担。它会审查所有战略和运营决策是否与规范性设定保持一致，并揭示其中的不一致之处。由此，规范层面在保持其人类根基的同时，也增强了有效性。

核心论点

AI-CEO 并非取代企业中人类赋予意义的机构，而是使其发挥效用。所有者、监事会和员工所期望的规范性设定，将由 AI-CEO 系统地转化为日常业务——其一致性程度，是基于相同价值观的人类 CEO 根本无法企及的。

8.2 人工智能 CEO 的五大秩序维度

AI 首席执行官基于五个秩序维度开展运作，这些维度是对圣加仑学派“秩序维度”概念 (Rüegg-Stürm 和 Grand, 2020) 在 AI 代理经济中的细化。它们构成了 AI 首席执行官所有决策的坐标系：

1. 愿景
2. 战略
3. 结构
4. 代理 (文化与行为)

5. 治理

每个治理维度都包含规范层、战略层和运营层——这实际上形成了一个由 5×3 构成的矩阵，共包含 15 个单元格。该矩阵是 AI 首席执行官的工作框架；每一项决策都可定位于其中，并可据此检验其与其他决策的一致性。

8.2.1 愿景

愿景回答了这样一个问题：十年后，这家企业的存在意义是什么？它是长期的意义基准，所有其他决策都必须以此为标准进行衡量。AI 首席执行官以机器可读的形式管理愿景：从愿景的叙述性文本中推导出指标，从而使愿景可量化。每一项战略和运营决策都会对照这些指标进行审查。如果 AI-CEO 发现存在不一致之处——例如某项短期盈利的措施会削弱长期愿景——它会向人类所有者或监事会发出警示。关于冲突的最终决策仍由人类作出；而冲突的识别则由 AI-CEO 自主完成。

8.2.2 战略

根据圣加仑学派的理解（Gomez 和 Probst, 1995），战略与其说是一个长期计划，不如说是在复杂环境中一种有意识的模式识别模式。AI-CEO 将这种模式识别付诸实践。他观察市场、竞争对手、技术、监管动态和社会趋势；识别模式与异常；制定备选方案，并根据愿景和价值观对其进行评估。至于战略决策——即选择哪种方案——他可以根据自主权程度（参见第 5 章）进行筹备或直接作出。在最高层级，他可在监事会设定的框架内独立作出战略决策。

8.2.3 结构

结构是工作开展的制度框架：组织架构、流程、汇报渠道、与供应商和客户的接口。在一家采用 AGENTONOMICS 模式的企业中，还包含其他结构要素：所有投入使用的子代理的定义、其职责范围、升级路径以及结算关系。AI 首席执行官将该结构作为动态模型进行管理。他能够识别结构性瓶颈，提出调整建议，并在自主模式下实施这些调整。

8.2.4 代理性：文化与行为

代理性涵盖了塑造参与者行为的文化模式。在由人类和代理组成的混合型组织中，这还包括子代理的行为：他们的沟通风格、冲突处理方式以及风险承受能力。AI 首席执行官确保这些模式与愿景保持一致。他据此对子代理进行培训，纠正偏离现象，并保持文化的活力。在人类方面，他无法直接干预——员工文化仍然是人类的事务。但他可以进行观察、描述、提出建议，并揭示不一致之处。

8.2.5 治理

治理是指企业通过所有者、监事会、立法机构及利益相关方进行管控的全部程序。在此，AI 首席执行官是受监管方，而非实施者：他需要被监管，而非进行监管。在 AGENTONOMICS-企业成立的最初几年，其治理结构的设计尤为关键。这些结构必须确保 AI-CEO 恪守其使命，确保干预渠道得以保留，并确保责任划分清晰明确。第 13 章将对此主题进行深入探讨。

8.3 AI 首席执行官的具体职责

为了将理论描述转化为实践概念，下面将通过具体职能概述 AI-CEO 的工作范围。这些职能在技术上现已可行，且在经济上具有合理性——尽管目前尚未将所有职能整合到单一的实施方案中。

8.3.1 战略形势分析

AI-CEO 会根据要求，对企业的战略状况进行全面评估：市场地位、竞争对手、优势与劣势、机遇与风险、发展趋势。在此过程中，它依托内部数据（销售统计、财务指标、客户反馈）和外部来源（市场研究、经济媒体、竞争对手报告）。此类分析所需的工作量——在传统管理咨询公司中可能需要六位数的费用和三个月的工作时间——如今已缩短至几分钟（！）。

8.3.2 决策准备

在做出任何重大决策之前，AI 首席执行官都会准备一份决策提案：包含备选方案、正反方论据、风险、财务影响、对利益相关者的影响，以及与愿景和战略的一致性。该提案总结在一页纸

上，并可根据需要进行深入展开。如今，一场需要数小时讨论投资决策的人工董事会会议，借助 AI 首席执行官的提案，可在 15 分钟内完成——且信息密度更高，准备工作也更为充分。

8.3.3 利益相关方沟通

AI-CEO 会针对员工、投资者、客户、政府部门及公众等不同受众群体，制定有针对性的沟通内容。各信息之间的连贯性极高——“同心异语”。新闻稿、年度报告、投资者关系动态及员工通讯均源自统一的核心知识体系和价值观。这不仅降低了成本，还提升了企业公信力。

8.3.4 运营管控

在日常业务中，AI 首席执行官会监控关键指标，识别偏差，并根据自主权程度提出或采取应对措施。现代化的 PuK 系统（参见第 9 章）为其提供了数据支持。他既能进行季度层面的运营，也能进行日层面的运营；早上就能掌握流动性状况，晚上就能了解销售数据。

8.3.5 危机管理

在危机中，AI 首席执行官的价值尤为凸显。当人们因压力而慌乱时，他仍能保持冷静。他能回忆起公司历史和财经媒体中所有类似的案例，并迅速、一致地制定行动方案。危机中的人为决策依然保留着人性化特征——但这些决策是基于更完善的信息基础做出的。

8.4 人工智能首席执行官的发展阶段

AI-CEO 的形成并非一蹴而就，而是经历了一系列发展阶段。这些阶段不仅涉及技术层面，还涉及组织和治理层面。参照第 5 章中关于自主性的分级（在教学上具有积极意义），可以区分出五个阶段：

8.4.1 第一阶段：AI 助理

AI 首席执行官是人类首席执行官的助手。它负责筹备会议、进行调研、整理文件并起草文稿。所有决策均由人类 CEO 作出。目前，大多数大型企业的董事会已达到这一阶段——通常是通过 ChatGPT 等通用工具实现的，但并未被视为 AGENTONOMICS 的实施。投资额：低。风险：小。效益：显著，主要体现在节省时间方面。

8.4.2 第二阶段：AI 参谋长

AI 首席执行官是一个成熟且针对企业量身定制的参谋职能。它能够访问所有内部系统，了解企业文化，并时刻关注企业战略。它不仅会撰写摘要，还会准备包含建议的完整决策文件。人类首席执行官负责做出最终决策，但通常会采纳这些建议。目前，在先进企业的试点项目中，这一阶段已可实现。投资规模：中等。风险：主要包括数据保护、知识产权保护以及知识库中的错误。

8.4.3 第三阶段：AI 联合 CEO

AI 首席执行官与人类首席执行官组成搭档。常规决策由 AI 首席执行官独立作出，战略决策由双方共同制定，原则性问题仍由人类负责。目前，这一阶段已在部分试点应用中实现，但尚未普及。这需要对决策权限进行仔细界定——哪些事项 AI 首席执行官可以自主处理，哪些需要人类参与？投资规模：高。风险：责任缺口、监管不明确、员工的文化接受度。

8.4.4 第 4 阶段：场景：由人类监事会监督的 AI 首席执行官

AI-CEO 独立负责日常运营。由人类组成的监事会——在德国语境下尤其包括共同决策机构——负责制定战略框架，监督规范性规定的遵守情况，并决定管理委员会的人选（即 AI-CEO 本人）。从技术角度来看，这一阶段目前已在特定且范围狭窄的业务领域得以实现——例如，服务范围明确的纯在线商业模式。在未来五到十年内，这一模式将在越来越多的企业中得到应用，尤其是那些没有接班人的中型企业以及商业模式明确的初创企业。投资规模：非常高。风险：监管风险、法律风险（责任归属问题）、文化风险。

8.4.5 第 5 阶段：多中心网络中完全自主的 AI 首席执行官这一远期目标

AI 首席执行官完全自主运作，在多中心网络中与其他 AI 首席执行官合作，签订合同，并独立调整业务领域。人类的参与仅限于所有权以及协调机构进行的规范性监督。目前这一阶段在很大程度上仍属于推测性质，并引发了根本性的法律和伦理问题。在未来几十年内，它可能会在某些利基领域实现——但并非作为标准，而是

8.5 应用：如何与 AI 首席执行官沟通？

对于 AI-CEO 的实际应用而言，一个关键问题是：如何与它对话？可以提出哪些问题，又会得到怎样的回答？这个问题绝非小事。人类 CEO 是经过社会化学习的——我们知道如何与他沟通、可以期待什么、以及如何应对他的回答。而对于 AI 首席执行官，这种社交实践尚待发展。下文将概述典型的提问类别及其可预期的回答模式。

8.5.1 战略形势类问题

示例问题：“你如何评估我们在 X 业务领域未来十二个月的竞争地位？” 预期回答通常分为四个部分：

- 简明扼要的主旨陈述（两到四句话），以浓缩形式呈现 AI 首席执行官的判断。
- 基于三到五个关键数据点的论证，这些数据点支撑着该判断。
- 明确的不确定性说明：哪些数据基础可靠，哪些方面存在缺口，以及纳入了哪些假设。
- 针对后续决策的建议——并非指令，而是作为讨论的提议。

关键在于，AI-CEO 能够透明地展现不确定性。人类 CEO 出于职业发展和地位考虑，往往倾向于在实际上并不存在确定性时，仍试图传递确定感。而设计精良的 AI-CEO 则恰恰相反：它会量化其不确定性，从而促进更明智的风险决策。

8.5.2 运营方面的询问

示例问题：“为什么上个月南部地区的销售数据下降了 12%？” 回答模板：

- 事实依据：确切数据及对比基准。
- 假设：三到五个成因假设，按可能性排序。
- 数据来源：每项假设基于哪些内部和外部数据来源？
- 行动方案：为了验证最可能的假设，下一步该怎么做？

AI 首席执行官不会轻易归咎于人。当他提及销售人员时，会基于事实并加以必要的区分。他还能回忆起上个月的对话、奖金结构以及市场竞争情况。这种记忆能力是他相对于人类首席执行官的主要优势之一。

8.5.3 规范性反思问题

示例问题：“尽管我们知道客户 Y 的工作条件存在问题，但接受其大额订单是否符合我们的价值观？” 回答范例：

- 参照公司明确记录的价值观（使命宣言、行为准则）。
- 识别冲突：我们的哪些价值观之间存在矛盾？
- 选项：接受、拒绝、附带条件地接受、与客户 Y 进行谈判。
- 利益相关方视角：员工、投资者、公众及监事会会作何反应？
- 明确呼吁由人来做出决策：“这个问题涉及我们公司的规范性实质。我请求召开一次监事会会议。”

对于规范性问题，AI 首席执行官绝不会擅自做出决定。他会将问题的复杂性呈现出来，并将决策权下放给承担企业责任的人类决策者。这种自我克制是其设计架构的一部分。

8.5.4 危机情况

示例问题——在急性危机之中：“我们收到了关于数据滥用的公开指控。我们该怎么办？”

回答范例：

- 未来一小时内的紧急措施：具体且可行的步骤。
- 事实情况：哪些信息已得到证实，哪些尚未确认？
- 利益相关方分类：谁需要在何时获知信息？监事会、员工、主管部门、客户、公众。
- 中期（未来几天）：需要开展哪些调查、聘请哪些外部顾问、采取何种沟通策略。
- 针对危机后时期的经验教训建议。

AI 首席执行官在危机中的优势在于利益相关方分级的速度和全面性。人类危机应对团队需要数小时才能完成的工作，AI 首席执行官仅需几分钟即可完成——同时还保留了由人类管理层进行优化调整的选项。

8.6 人工智能 CEO 何时说 “不”

一个设计精良的 AI-CEO 会在特定情况下明确说 “不” ——或者更准确地说：“此事我无法也不应做决定，需要人类决策者介入。” 这种自我克制并非软弱，而是符合 AGENTONOMICS 原则的 AI-CEO 的本质特征。典型情况有四种：

规范性与构成性问题

对愿景、价值观和章程的修改。这些事项应由所有者和监事会负责。

具有重大影响的人事决策

关键人员的聘用与解雇，以及管理职位的晋升。在此类事务中，对个人命运的人性化判断应优先于其他考量。

特殊风险

可能对公司或第三方生存构成威胁的决策。在此情况下，“双人审核原则”不可或缺——必须由至少两人进行审核。

涉及利益相关方的冲突案例

当 AI 首席执行官意识到某项决策以规范框架中未明确授权的方式损害了利益相关方的正当权益时。

8.7 作为连接 PuK 逻辑桥梁的 AI-CEO

AI-CEO 概念本身是一种定性架构。它需要一个定量框架，以从经济角度精确地锚定决策。规划与控制系统 (PuK) 工具 (Hahn 和 Hungenberg, 2001) 提供了这一框架，下一章将对此进行详细探讨。AGENTONOMICS 矩阵由五个构建模块和五个生命周期阶段组成的 25 个领域构成，是 AI-CEO 与 PuK 的共同工作文件。如果没有 PuK，AI-CEO 将只是一个口才出众但经济思路模糊的顾问；有了 PuK，他便成为一名名副其实的经济决策者。

AI-CEO 并非取代企业最高领导者，而是将其从运营整合带来的认知负担中解放出来。剩下的，正是人类一直以来所擅长的：赋予意义并建立信任。

参考文献

Bleicher, K. (2011): 《整合管理概念》。第 8 版, Campus 出版社, 法兰克福。

Gomez, P.; Probst, G. (1995): 《整体性问题解决与实践》。网络化思维——企业家式行动——一个人说服力。Haupt 出版社, 伯尔尼。

Hahn, D.; Hungenberg, H. (2001): 《PuK——价值导向的控制理念》。第 6 版, Gabler 出版社, 威斯巴登。

克诺尔, A. C.; 苏克富尔, H. (2024) : 《从数据主权到技能主权》。技术报告 TUM-I24112, 慕尼黑工业大学 (TUM) , 慕尼黑。

吕格-施图姆, J.; 格兰德, S. (2020) : 《圣加仑管理模型》。第 3 版, 豪普特出版社, 伯尔尼。

乌尔里希, H. (1968) : 《企业作为生产性社会系统》。豪普特出版社, 伯尔尼。

乌尔里希, P. (2008) : 《整合性经济伦理学》。第 4 版, Haupt 出版社, 伯尔尼。

9 模块 5: PuK 工具

ADMRF 的第五个模块是规划与控制工具 (PuK)。它是将所有先前模块在经济层面加以锚定的定量框架。第 5 至 8 章描述了代理人的定性架构, 而 PuK 工具则提供了其经济骨架: 它将定义、构建模块、协调机构连接以及 AI-CEO 逻辑转化为可追踪的货币量、计划指标和控制流程。没有 PuK, AGENTONOMICS 仅停留在理论构想层面; 有了 PuK, 它便成为一个可在企业管理中付诸实践的模式。

PuK 工具直接承接了 Hahn 和 Hungenberg (2001) 提出的以价值为导向的控制理念。本文明确沿袭这一传统, 因为它将规划、控制和管理以一种符合圣加仑学派整合要求的方式进行了概念化——同时, 它还提供了足够的灵活性, 可以针对人工智能代理经济体的具体要求进行定制。其财务基础参考了 Berk 和 DeMarzo (2020) 的研究, 特别是关于资本预算、资本结构和企业估值的章节。

9.1 承袭哈恩 (Hahn) 和洪根贝格 (Hungenberg) 传统的 PuK 工具

Hahn 和 Hungenberg (2001) 经过多年研究, 开发了一种适用于工业企业的规划与控制综合概念, 该概念已在德语区的财务控制部门中得到广泛应用。其中有三项基本思想对 AGENTONOMICS-PuK 尤为重要:

9.1.1 以价值为导向的指导原则

PuK 始终以价值为导向: 其核心在于企业价值的长期增长, 而非短期利润最大化。这种取向符合 AGENTONOMICS 代理人的逻辑——作为经济主体, 该代理人致力于在其整个生命周期内实现可持续的价值创造。正如 Berk 和 DeMarzo (2020) 所详细阐述的那样, 折现现金流法、资本成本和风险评估正是这种价值导向的数学工具。

9.1.2 综合规划

在 PuK 理论的理解中, 规划并非财务部门孤立进行的活动, 而是一个将战略、战术和运营规划相结合的集成过程。每一项战略决策都会直接转化为财务指标; 每一项运营措施都会对照战

略方针进行审查。这种整合正是布莱希尔（2011）所提出的概念性要求——哈恩和洪根贝格则对其进行了定量阐述。

9.1.3 闭环控制

没有控制的规划只是空想；没有规划的控制则是盲目的。PuK 将规划与控制视为一个闭环控制过程：计划→实际→偏差分析→调整措施→新计划。这种闭环逻辑与比尔（Beer, 1972）提出的控制论相关，使 PuK 能够与人工智能代理的算法控制相兼容——因为算法的运作结构正是如此：观察、比较、修正。

9.2 AGENTONOMICS 矩阵作为核心工具

PuK 工具的核心工作文档是 AGENTONOMICS 矩阵。该矩阵已在第 4 章协调机构绍过；现在将在此对其进行具体填充。该矩阵交叉了两个维度：

- 水平维度：第 6 章中的五个构建模块。
- 纵向：五个生命周期阶段（DBOT 逻辑）：设计（Design）、建造（Build）、运营（Operate）、拥有（Own）、移交（Transfer）。

由此形成 25 个单元格。每个单元格包含来自四个类别的结构化信息：

1. 活动：在此阶段，针对该构建模块将具体开展哪些工作？
2. 成本：涉及哪些资本支出（CAPEX）和运营支出（OPEX）？
3. 关键绩效指标（KPI）：如何衡量成功与进展？
4. 风险与责任：可能出现哪些问题，由谁负责？

该矩阵并非静态的，而是会持续更新。在 AGENTONOMICS 算法中，它是所有模块都以此为依据的控制文档。AI-CEO、CFO 和 CTO 都访问同一张矩阵——各自从特定视角出发，但基于共同的数据基础。

9.2.1 示例行：设计阶段

为了使该矩阵具体化，下面将展示一个示例行。在设计一个符合 AGENTONOMICS 标准的“AGENTONOMICS 副教授博士”代理人的设计阶段，各字段通常可填写如下：

- 算法：设计基于 ReAct 的协调方案，包含“研究”、“综合”和“修正”子任务。活动：架构设计。成本：1 至 3 人月的概念设计工作。KPI：完整定义所有决策点。风险：架构缺陷，在后续阶段修正将耗费高昂成本。
- 接口：Web、语音和 API 接口的规范。活动：用户体验（UX）设计、接口文档编制。成本：1 至 2 人月。KPI：用户测试中的接受率。风险：学生群体中存在接受度差距。
- 模型：在基础模型、开源模型和微调模型之间进行选择。活动：基准研究、成本效益比较。成本：1 人月，外加初始模型测试。关键绩效指标（KPI）：测试题库中正确答案的比例。风险：若选择专有模型，可能面临供应商锁定。
- 工具与向量：基于 AGENTONOMICS 文献和课程资料构建向量数据库；制定工具接口规范。活动：数据建模、嵌入式选定。成本：2 至 3 人月。关键绩效指标（KPI）：标准化基准测试中的检索质量。风险：初始数据填充中的知识缺口。
- 协调机构：制定与慕尼黑工业大学（TUM）内部协调机构系统的连接规范；定义身份、声誉和争议调解接口。活动：治理规范制定。成本：主要为概念设计成本。关键绩效指标（KPI）：完整制定所有协调机构接口规范。风险：与其他正在开发的 AGENTONOMICS 实现方案之间存在标准化缺口。

这五个单元综合起来，共同勾勒出 Doz. Dr. AGENTONOMICS 设计阶段的精准图景——包括成本框架、成功标准和风险说明。将其与五个生命周期阶段相乘，便形成了对代理在其整个生命周期内完整的经济描述。

9.3 代理人的资产负债表和损益表

AGENTONOMICS-PuK 在经济层面的一项重要创新，在于将资产负债表逻辑系统地应用于各个代理。每个代理都被建模为一个独立的经济主体，拥有自己的资产负债表和损益表。这种视角在当今并不常见——软件和开发成本通常根据获取形式、开发阶段和会计准则进行核算——但它在经济上是准确的，对于 AGENTONOMICS 世界而言不可或缺。

9.3.1 代理资产负债表的资产方

资产方列示的是代理人的资产：

- 作为无形资产的算法：特定的代码、协调逻辑以及针对企业特性的调整。

- 模型与微调投资：用于调整基础模型的可按比例资本化的费用。
- 向量和知识库：使该代理区别于其他代理的特定知识储备。
- 工具许可：对外部工具和数据源所获得的使用权。
- 声誉资本：若其以可量化的方式记录，且始终与代理身份相关联（在资产负债表上不属于可单独资本化的资产）。

9.3.2 代理资产负债表的负债方

负债方列示融资来源和债务：

- 自有资本：由母公司或外部投资者提供的资金。
- 借贷资本：专门为代理商借入的贷款。
- 对供应商的负债：模型供应商、云服务提供商、工具供应商。
- 准备金：用于应对监管风险、维护义务及潜在的赔偿责任。

9.3.3 损益表

代理商的损益表反映其经济业绩。收入方面包括代理商通过其业务活动产生的收入——无论是来自委托方的直接报酬、与其他业务部门的内部结算、许可收入，还是协调机构网络中的交易手续费。费用方面包括运营成本：模型使用费、云基础设施费、工具调用费、维护费、协调机构费以及人工监管的相应人员成本。两者之间的差额即为代理商利润——在成熟的 AGENTONOMICS 世界中，这一指标的计量将如同当今衡量子公司利润一样理所当然。

实践提示

对每个代理进行单独核算，乍看之下似乎很繁琐。但实际上并非如此：相关数据本就会被收集（云成本、模型使用、工具调用次数均可按代理进行计量）。创新之处在于对其进行系统化的汇总和分析。如果不进行代理核算，就等于在盲目行动。

9.4 CAPEX 与 OPEX——投资与运营成本概况

将资本支出与运营支出区分开来，对于代理的经济管控至关重要。其结构与传统资本货物有所不同：

9.4.1 代理的 CAPEX 概况

初期投资主要集中在设计和构建阶段。典型的 CAPEX 项目包括：架构设计、算法实现、模型微调、向量数据库构建以及接口开发。与传统资本货物（如建筑物、机器）不同，这些 CAPEX 值在很大程度上受人力成本驱动，且更难预测。

9.4.2 代理的运营支出 (OPEX) 概况

日常运营成本长期居高不下，而且——这是经济上至关重要的一点——无法通过规模效应实现线性降低。每次模型调用都要花钱，每次工具调用都要花钱，每小时云服务都要花钱。这些成本会随使用量的增加而增加。使用频率提高十倍的代理，其运营成本大约也会增加十倍。这与传统软件不同，后者的边际部署成本趋近于零。但是：自主托管、预留资源、缓存、批处理和模型路由会改变成本曲线。

9.4.3 对商业建模的影响

从资本支出 (CAPEX) /运营支出 (OPEX) 特征可以得出一个对业务建模至关重要的结论：与传统数字产品相比，AGENTONOMICS 代理的扩展性较差，但每笔交易的增值能力更强。业务模型必须体现这一特性——例如通过按使用量计费、明确界定的服务边界，或者通过一种合理组合昂贵与低成本模型调用的架构来实现。

9.5 生命周期中的控制流程

在代理的生命周期——设计、构建、运营、拥有、移交——中，控制流程会随之变化。每个阶段主导的 KPI、报告格式和升级阈值各不相同。PuK 工具明确界定了这些特定于各阶段的流程。

9.5.1 设计阶段：概念性 KPI

在设计阶段，尚无市场数据、使用统计数据或损益表。管理通过概念性 KPI 进行：规格说明的完整性、与愿景和战略的一致性、商业案例假设的可行性、风险分析的合理性。报告格式为设计文件；升级阈值是指内容上的缺失，这些缺失在后续阶段会造成高昂成本。

9.5.2 构建阶段：项目关键绩效指标

在构建阶段，传统项目管控负责统筹管理：进度、预算、交付质量、风险。报告形式采用带交通灯系统的项目状态报告；当延误或预算超支超出预定义容差范围时，即触发上报阈值。

9.5.3 运营阶段：业务数据

在日常运营中，经济类 KPI 占据主导地位：销售额、利润率、贡献利润、流动性、客户满意度、市场渗透率。此阶段将全面应用 Hahn/Hungenberg 逻辑体系。报告形式为月度报告；升级阈值为偏离计划值的幅度超出预定义范围的情况。

9.5.4 持有阶段：资产指标

在“自有”阶段——即当代理被视为永久性资产时——资产指标成为关注焦点：账面价值、市场价值、净资产价值、收益价值、折旧、以及用于维护和进一步发展的投资需求。目前，该阶段在企业实践中尚不成熟；AGENTONOMICS 正积极推动其发展。

9.5.5 转移阶段：交易估值

当代理被出售、停用或转移至其他场景时，将对其进行交易估值。哪些“构建模块”可转移？它们附带何种价值？哪些负债将一并转移？在正在兴起的人工智能代理并购实践中，这一阶段将日益重要。

9.6 首席财务官职能与 AI 首席执行官之间的联系

PuK 工具基本上涵盖了传统企业中首席财务官（CFO）所承担的职能。在 AGENTONOMICS 的逻辑框架下，该职能可由专门的 AI-CFO 或集成式 AI-CEO 来承担。这两种架构各有优劣：

一体化 AI-CEO

AI-CEO 亲自承担 PuK 职能。优点：战略与财务规划之间的一致性达到最大化。缺点：职能未分离，控制强度较低。

专门的 AI 首席财务官

由独立的智能代理承担规划与控制职能，并与 AI 首席执行官保持一种富有成效的张力关系。

优点：职能分离，内置双人复核机制。缺点：代理之间需要投入协调精力。

选择取决于商业模式和风险特征。在受监管的行业和资本密集型商业模式中，应优先采用两个代理人的分工模式；而在结构较简单的场景中，由兼具 PuK 职能的集成式 AI 首席执行官即可满足需求。

PuK 工具是圣加仑模型定性架构与资产负债表严酷现实之间的桥梁。若不通过 PuK 来管控某代理，则该代理的建立仅基于空谈。

参考文献

Beer, S. (1972): 《企业的头脑：组织的 managerial cybernetics》。Allen Lane 出版社，伦敦。

Berk, J.; DeMarzo, P. (2020): 《金融经济学基础》。第 6 版，Pearson 出版社，慕尼黑。

Bleicher, K. (2011): 《综合管理理念》。第 8 版，Campus 出版社，法兰克福。

Hahn, D.; Hungenberg, H. (2001): 《PuK——价值导向的控制理念》。第 6 版，Gabler 出版社，威斯巴登。

克诺尔, A. C.; 苏克富尔, H. (2024) : 《从数据主权到技能主权》。技术报告 TUM-I24112, 慕尼黑工业大学, 慕尼黑。

吕格-施图姆, J.; 格兰德, S. (2020) : 《圣加仑管理模型》。第 3 版，Haupt 出版社，伯尔尼。

10 模块 6：利益相关者沟通者

ADMRF 的第六个也是最后一个模块探讨了代理与其利益相关者之间的沟通。这是将 Rüegg-Stürm 和 Grand (2020) 所描述的企业与利益相关者之间的“互动主题”转化为人工智能代理领域的具体实践。圣加仑模型第四代的沟通转型在此得到了具体的 operationalization——关键区别在于，AGENTONOMICS 代理能够提供人类行政部门在速度、差异化和一致性方面无法企及的沟通服务。

该模块可被视为首席销售官（CSO）——此处的“销售”并非指狭义的销售职能，而是广义上指建立和维护所有相关利益相关者关系。它涉及企业对外和对内的沟通接口：投资者沟通、客户沟通、员工沟通、政府部门沟通以及公共关系。

10.1 圣加仑学派对利益相关者的理解

圣加仑模型从一开始就对“利益相关者”这一概念进行了广义界定（Ulrich 1968；Rüegg-Stürm 和 Grand 2020）。利益相关者是指所有对企业具有合法利益或受其行为影响的参与者。这种宽泛的定义使圣加仑模型区别于自 20 世纪 90 年代以来在盎格鲁-撒克逊地区盛行的纯粹以股东为导向的观点。这也正是 AGENTONOMICS 概念的出发点：一个仅以委托人为导向的代理人，在圣加仑理论中并非一个完整的经济主体，而是其所有者的延伸。

从理想类型来看，可将利益相关者分为六大类：

- 所有者和出资人：投资者、股东、银行。
- 员工：企业的在职人员，包括参与决策的机构。
- 客户：服务当前及潜在的购买者。
- 供应商和商业伙伴：包括其他“AGENTONOMICS”代理人 and 模型提供商。
- 政府与监管机构：税务部门、监管机构、数据保护机构。
- 公众：媒体、公民社会、社区、非政府组织。

每个群体都有自己的语言，有独特的信息需求，并对不同的沟通刺激做出不同的反应。利益相关者沟通者必须基于一致的知识 and 价值观基础，应对这种差异性。这正是 AI 代理相对于人类公关部门所具备的独特优势。

10.2 川崎的“Pitch”方法论

在企业背景下，盖伊·川崎（Guy Kawasaki）所开发的“推介法”是一种特别有效的利益相关者沟通方法。川崎于 20 世纪 80 年代初在苹果公司担任“布道师”，并基于这一经验，发展出一套关于企业说服力沟通的务实理论（Kawasaki 2015）。利益相关者沟通者利用这一方法，并将其应用自动化。

10.2.1 10-20-30 法则

卡瓦萨基最著名的规则涉及面向投资者的演示：十张幻灯片、二十分钟、三十号字体。将幻灯片数量限制在 10 张，迫使人们专注于核心内容；20 分钟的时间恰好符合高管层的注意力持续时间；30 点的字号则能防止演示演变成冗长的讲座。这三项要求在算法上很容易实现，并能立即形成投资者沟通的外在结构。利益相关者沟通者可以将其作为标准模板使用。

10.2.2 标准推介中的十张幻灯片

川崎对投资者演示文稿中的十张幻灯片具体说明如下：

1. 标题：公司名称、口号、联系方式。
2. 问题：市场中存在哪些痛点？
3. 解决方案：公司如何解决这些痛点？
4. 商业模式：公司如何盈利？
5. 产品背后的魔力：技术、知识产权、竞争优势。
6. 营销与销售：市场进入策略。
7. 竞争：竞争对手是谁？我们与他们的区别是什么？
8. 团队：公司背后有哪些人？
9. 预测与里程碑：三至五年。
10. 现状与时间表：我们目前处于什么阶段，下一步是什么？

如今，这种结构已成为初创企业界的实际标准。利益相关者沟通工具可以基于 AGENTONOMICS 矩阵的数据和 AI-CEO 概念的内容自动生成该结构——确保数据、语言和可视化效果的一致性。

10.3 从 PuK 到路演：运营层面的推导

利益相关者沟通工具的核心功能之一，是将 PuK 工具中的定量指标转化为适合各利益相关者的沟通格式。这种转化并非听起来那么简单：针对投资者、员工、监管机构以及公众，同一条信息必须以不同的方式进行呈现。

10.3.1 在差异化中保持一致性

利益相关者沟通专员在此过程中遵循一项核心要求：所有沟通均基于共同的事实基础和共同的价值观基础。差异体现在语言、细节深度和重点设置上，而非实质内容。投资者收到的关键指标与员工不同，但传达的信息是相同的。监管机构的信函与新闻稿的措辞不同，但所陈述的事实完全一致。这种在差异化中的连贯性，正是本模块的核心要义。

10.3.2 典型输出

利益相关者沟通工具通常会生成：

- 用于与投资者或合作伙伴初次接触的一页简介。
- 采用川崎方法论编制的投资者演示文稿。
- 年度报告和季度报告。
- 新闻稿和问答文件。
- 面向员工的内部通讯。
- 产品变更时的客户沟通。
- 向主管部门提交的申请及监管报告。
- 危机沟通草案。

每项产出都遵循各自的逻辑，但都源自同一数据基础：AGENTONOMICS 矩阵和 AI 首席执行官的秩序要素。

10.4 危机沟通作为该模块的压力测试

对每位利益相关者沟通者而言，危机才是真正的压力测试。在危机中，企业的沟通能力将面临最大压力：时间紧迫、信息不完整，众多利益相关者同时要求得到关注。此时，该模块能否履行其职责便显露无遗。

优秀的危机沟通者具备四大特质：

快速而不草率

首份新闻稿必须在数小时内发布，且不得影响后续声明的可信度。利益相关者沟通专员能在数分钟内起草此类稿件，并为发言人提供多种版本供选择。

受众差异化

员工、客户、政府部门、媒体——所有受众都希望得到不同的沟通方式。利益相关方沟通专员会并行制作不同版本的稿件，并确保其内容的一致性。

恪守企业价值观

在危机中，人们往往容易忘记平时所遵循的价值观。Stakeholder Communicator 会提醒企业遵循其规范性准则（源自“定义框架”和“AI-CEO”概念），并据此审查每一条沟通内容。

筹备后续步骤

危机沟通绝非仅是被动应对。它为后续步骤做准备：调查、后果处理、补救措施、结构性变革。利益相关者沟通专员会预先规划这一路径，并确保最初阶段的沟通不会阻碍未来几周的应对方案。

10.5 内部与外部沟通——独特的平衡

本模块的一项特别敏感的任务是在内部与外部沟通之间取得平衡。员工不应从媒体上获悉本应最先知晓的信息；投资者不应在承担最大风险的情况下却最后才获知消息；监管机构不应因信

息泄露而措手不及。利益相关方沟通专员负责协调这些顺序，并确保时间安排符合各利益相关方的需求。

在德国语境下，尤其需要关注共同决策原则。企业工会享有信息知情权和听证权，其行使时间应早于对外沟通。正确的 AGENTONOMICS 实施会自动考虑这一顺序——这不是出于善意，而是为了合规。

一家未能在所有利益相关者之间保持一致沟通的企业，无论在内部还是外部都会失去信任。“利益相关者沟通专员”正是 AGENTONOMICS 针对沟通转型中复杂性管理问题给出的解决方案。

参考文献

哈贝马斯, J. (1981) : 《沟通行动理论》。苏尔坎普出版社, 法兰克福。

Hahn, D.; Hungenberg, H. (2001): 《PuK——以价值为导向的控制理念》。第 6 版, Gabler 出版社, 威斯巴登。

川崎, G. (2015) : 《创业的艺术 2.0》。Portfolio 出版社, 纽约。

吕格-施图姆, J.; 格兰德, S. (2020) : 《圣加仑管理模型》。第 3 版, 豪普特出版社, 伯尔尼。

乌尔里希, H. (1968) : 《企业作为生产性社会系统》。豪普特出版社, 伯尔尼。

11 AI 代理经济体的增长理论

本章标志着手册超越了模块层面，转向探讨更根本的问题。AGENTONOMICS 经济体如何增长？哪些机制推动其发展，哪些机制对其构成限制，其中蕴含着哪些风险？增长理论并非学术上的补充，而是实践上的必要：任何将人工智能代理视为经济主体并予以重视的人，都必须理解它们相互作用所产生的宏观经济后果。

11.1 合作乘数

经典增长理论将增长分为延展型增长和集约型增长（索洛，1956）。延展型增长源于生产要素的增加——更多资本、更多劳动力、更多土地。集约型增长源于对生产要素的更高效利用——技术进步、教育、组织。AGENTONOMICS 在此经典框架中增添了第三个维度：通过多中心代理网络中的合作实现的增长。

11.1.1 作用机制

当多中心网络中的代理进行合作时，所产生的附加值将超过各单项贡献之和。这一论点并非新鲜事——自夏皮罗和瓦里安（1999）以来，网络效应已成为一个公认的概念。真正的新颖之处在于，这些效应在人工智能代理经济中可能发挥的作用强度。原因有三：

- 谈判速度：代理能在几毫秒内完成谈判，而人类进行同样的谈判则需要数天时间。
- 交易粒度：代理能够以经济合理的方式交换微小的部分服务，而这些服务在人类交易中往往因交易成本过高而被忽略。
- 合作的一致性：智能体不会忘记约定，不会有情绪低落的日子，也不会出于虚荣心进行谈判。

基于这三项特性，形成了一个合作乘数：由 n 个智能体组成的网络的产出，可能远高于单个智能体产出的 n 倍。这一乘数正是 AGENTONOMICS 所承诺的真正增长潜力。

11.1.2 乘数效应的局限性

该乘数并非无穷无尽。需注意以下三项限制：

复杂性限制

随着代理数量的增加，协调的复杂度呈平方增长。当规模超过一定阈值时，协调成本将变得如此高昂，以至于额外效益开始下降。协调机构（第 7 章）可以推迟这一极限的到来，但无法消除它。

信任限制

合作以信任为前提。随着参与者数量的增加，建立信任变得越来越困难。声誉体系和争议调解程序虽能起到一定作用，但无法完全弥补这一问题。

意义边界

合作必须具有意义。如果参与者仅仅为了合作而合作，就无法创造价值，只会徒增繁忙。此时，AI-CEO 模块中的规范性设定便发挥作用：它界定了哪些合作是有意义的，哪些是无用的。

11.2 三元分析：三种增长形式

参照苏克菲尔（Suckfüll）对跨国企业的分析（Suckfüll 1994），AGENTONOMICS 网络中可区分三种增长形式：

11.2.1 通过子代理实现的垂直整合

一个代理会为特定的子任务生成子代理（参见第 5 章，参数 6）。这些子代理扩展了主代理的价值创造深度。增长源于内部功能的增加，而非市场的扩张。从经济角度来看，这相当于传统意义上的垂直整合。

11.2.2 与对等代理人的水平联网

属于不同所有者的多个代理在中间商网络中开展合作。每个代理都保持其自主权，但它们相互交换服务并共同创造价值。增长源于网络的扩展和关系的深化。从经济角度来看，这相当于一种没有中央平台的平台式增长。

11.2.3 对角线专业化

代理正发展出日益专业化的能力特征，这些能力只有通过与互补型代理合作才能发挥作用。增长源于分工的深化。从经济角度看，这相当于斯密提出的经典增长机制——即通过深化分工实现增长，只不过是在一个新的参与者层面上实现的。

这三种增长形式并非相互排斥，而是相辅相成。一个成熟的“AGENTONOMICS”经济体系将同时兼顾这三种形式——而在它们相互强化的领域，合作乘数效应最为显著。

11.3 多中心网络组织作为增长引擎

这些增长形式在宏观经济层面的核心影响，是传统企业边界的逐渐消解。在AGENTONOMICS的世界中，企业之间的界限正日益模糊。代理人们跨越企业边界进行合作，其协作粒度之细，过去需要组织整合才能实现。结果是：未来的典型组织形式不再是集成式企业，而是多中心网络。

这一诊断并非新鲜事——Picot、Reichwald 和 Wigand (2020) 早在他们的“无边界企业”理论中就已提出。唯一的新颖之处在于，AGENTONOMICS 提供了基础设施条件，使这种形式能够超越试点项目，成为现实。

11.3.1 对增长模式的影响

多中心网络的增长方式与一体化企业不同。其具有三大特征：

- 增长源于新节点的加入，而非现有节点规模的扩大。
- 各节点保持小规模且灵活；网络整体则不断扩大。
- 价值创造广泛分布，而非集中于少数节点。

这种增长模式——如果得以普及——将与当今平台经济的集中化动态形成鲜明对比。这也是与“AGENTONOMICS”理论相关联的核心希望之一：即下一波数字化转型未必必然导致垄断性集中，而是可以呈现多中心分布的态势。

11.4 网络中的创新与更新

Rüegg-Stürm 和 Grand (2020) 区分了两种发展模式：优化与革新。这一区分同样适用于 AGENTONOMICS 网络。优化是指对现有惯例的持续改进；革新则是打破惯例并构建新模式。

从熊彼特的角度来看，更新即创新——即“创造性破坏”（熊彼特，1934），它打破旧的市场秩序并创造新的市场秩序。在 AGENTONOMICS 经济中，创新体现在两个层面：

11.4.1 代理中的创新

更优的算法、更强大的模型、更丰富的知识库——在单个代理层面，持续的创新正在发生。如果代理设计得当，这种创新就会融入其“构建-运营”循环之中。

11.4.2 网络层面的创新

新的合作模式、新的商业模式、新的协调机构结构——在网络层面，创新作为一种自发涌现的现象而产生。个别代理发明一种新的合作模式，其他代理则加以采用，网络由此发生转变。

这一创新的第二层面目前鲜有研究，在实践中也尚未充分显现。AGENTONOMICS 在此进行了概念性的前瞻：如果多中心代理网络确实形成，网络创新将成为未来几十年最重要的增长源泉之一。

“代理经济学”经济体的增长并非源于大型节点的扩张，而是源于密集型网络的扩展。这绝非细节问题，而是一种政治宣言。

参考文献

Knoll, A. C.; Suckfüll, H. (2024): 《从数据主权到技能主权》。技术报告 TUM-I24112, 慕尼黑工业大学 (TUM), 慕尼黑。

Picot, A.; Reichwald, R.; Wigand, R. T. (2020): 《无边界企业》。第 6 版, Springer Gabler 出版社, 威斯巴登。

Rochet, J.-C.; Tirole, J. (2003): 《双边市场中的平台竞争》。载于：《欧洲经济协会期刊》1(4), 第 990–1029 页。

Rüegg-Stürm, J.; Grand, S. (2020): 《圣加仑管理模型》。第 3 版, Haupt 出版社, 伯尔尼

。

舒姆佩特, J. A. (1934) : 《经济发展理论》。哈佛大学出版社, 马萨诸塞州剑桥。

夏皮罗, C.; 瓦里安, H. R. (1999) : 《信息规则》。哈佛商学院出版社, 波士顿。

索洛, R. M. (1956) : 《对经济增长理论的贡献》。载于: 《经济学季刊》第 70 卷第 1 期, 第 65–94 页。

Suckfüll, H. (1994): 《跨国组织模型》。博士论文, 圣加仑大学。

12 经济解决方案：人工智能的青春期

如果“代理经济学”（AGENTONOMICS）得以普及，其宏观经济和社会影响将远远超出单个企业的决策范畴。本章概述了主要影响，并探讨了针对预计将出现的矛盾所提出的经济和政治解决方案。本章的论述紧密承接了克诺尔和苏克富尔（2024）在 TUM 报告 TUM-I24112 中的分析。

12.1 人工智能的青春期

Knoll 和 Suckfüll（2024）将这一阶段称为“人工智能的青春期”。这一比喻十分贴切：当今的人工智能已不再是孩童，但也尚未成年。它功能强大，但尚不成熟；性能卓越，但不够可靠；应用广泛，但尚未稳定融入社会。在这个阶段——据作者估计将持续约十至十五年——预计会出现各种动荡，社会和经济框架必须对此进行缓冲。

对于完全实现“AGENTONOMICS”情景的核心经济预期是：我们将迎来每年约 20% 的经济增长，但同时也将面临约 20% 的风险敞口。这两个数值并非算术上的固定值，而是指示性的——旨在说明预期颠簸的量级。这两个数字在历史上都极为罕见，且都对惯常的监管逻辑构成了挑战。

12.2 20% 的增长：来源与后果

20% 的年增长率将是经合组织（OECD）国家过去几十年所实现增长率的十倍。这相当于中国在经济繁荣时期所经历的最快增长速度——但起点要高得多。可以归纳出三个来源：

12.2.1 知识型工作的生产率飞跃

知识型工作——包括研究、分析、综合和沟通——在发达经济体中占据了劳动的大部分。如果人工智能代理能够将这些领域的生产率提高 2 至 5 倍——而许多迹象都表明这是可能的——那么将带来史无前例的增长动力。

12.2.2 新的价值创造领域

AGENTONOMICS 智能代理使一些仅靠人力无法实现经济可持续性的商业模式成为可能——例如细粒度的咨询服务、个性化教育服务、高频市场分析以及全天候服务。这些领域在创造增长的同时，不会挤占现有工作岗位。

12.2.3 减少摩擦损失

当今经济活动中，相当一部分由摩擦损失构成：搜索成本、洽谈成本、谈判成本、监管成本。AGENTONOMICS 代理能够大幅减少这些摩擦损失——由此节省的精力将转用于生产性活动。

12.3 悲观预测：20%的基础失业率——成因与后果

与此同时，在过渡阶段可能会出现高水平的基底失业率，其规模在历史上仅与 20 世纪 30 年代的世界经济危机相当。这一预期并非必然发生——它在很大程度上取决于经济政策框架的制定方式。但这一预期足够合理，因此应当坦率地加以讨论，而不是试图将其辩解掉。

12.3.1 面临威胁的职业

有三类工作面临特别大的威胁：

- 标准化知识型工作：会计、常规案件的法律应用、翻译、简单的报告撰写。
- 中层管理工作：汇报、协调、层级报告、标准化分析。
- 咨询与服务类工作：调研、初次接触、标准化咨询流程。

在所有这些领域，人工智能代理目前或在不久的将来都能提供至少与人类标准执行水平相当的服务——而且成本仅为人类的一小部分。其结果未必是失业，但必然会带来剧烈的结构性变革。

12.3.2 与卢德派的类比

一个历史类比不言而喻：19 世纪初的卢德运动。英国纺织工人砸毁了取代他们手工劳动的织机。他们最终在这场斗争中败下阵来——机器来了，工人不得不适应。但是：他们的主张有实质性的道理。他们的工作环境确实被摧毁了；适应过程持续了数代人；社会动荡十分严重。

如今，一场类似的冲击正在酝酿，但存在三个重要区别：

- 这场冲击并非针对体力劳动，而是针对知识型工作——也就是那些此前一直认为自己能免受自动化影响的群体。
- 这场冲击不会跨越几代人，而是在短短几年内发生——适应期明显缩短。
- 这场冲击并非在特定地区集中爆发（如当年在约克郡和兰开夏郡那样），而是同时波及所有发达经济体。

如今，那些将对人工智能的怀疑一概视为拒绝现代化的观点，却忽略了卢德派有着正当的诉求——他们并非想要阻碍进步，而是希望参与其中。这正是我们今天面临的问题。

12.4 分配问题：谁将受益？

从经济角度来看，关键问题不在于结构性变革是否存在，而在于其收益如何分配。如果经济增长率确实达到 20%，而失业率同样达到 20%，那么分配问题就变得尤为严峻。如果没有监管政策的引导，可能会出现这样一种局面：少数人工智能基础设施的所有者变得极其富有，而广大民众却失去了谋生基础。

12.4.1 购买力作为系统性问题

一个尤为尖锐的经济后果是：资本主义经济需要购买力才能运转。如果大众因劳动产出不再被需求而丧失购买力，经济的 demand side 就会崩溃——随之崩溃的还有人工智能基础设施所有者的商业模式。克诺尔和苏克富尔（2024）对此论点进行了犀利的阐述：集中式人工智能主导格局存在这样的风险：缺乏竞争会导致财富分配变得极度不均衡，以至于购买力的缺失将动摇整个资本主义体系的根基。

换言之：一个因人工智能导致大规模失业的世界，不仅会摧毁人们的生计，还会摧毁其自身的商业模式。这种自我矛盾正是推动积极制定监管政策的经济依据。

12.5 解决方案

有哪些可行的解决方案？可以确定四个主要方向，它们并非相互排斥，而是互为补充：

12.5.1 技能主权作为个人资产

Knoll 和 Suckfüll (2024) 的核心建议是：必须让每个人都有机会将自己的技能复制到人工智能代理中，并从中获得经济收益。将技能转移到智能代理中的个人，应继续拥有该代理的所有权，并分享其创造的价值。这种架构将个人主权与经济参与相结合，是欧洲针对这一挑战提出的独特解决方案。

12.5.2 多中心市场结构

建立一个既能促进多中心 AGENTONOMICS 网络、又能防止垄断性集中的监管框架至关重要。这与竞争法、数据保护法（ ）以及《人工智能法案》（AI Act）密切相关。切实执行数据可携性、互操作性标准以及平台与市场参与者之间的职能分离，是最低要求。

12.5.3 教育与再培训

如果 20% 的工作岗位被替代，教育和再培训体系就必须更加高效。传统的在职教育观念——培训一次就万事大吉——已经彻底过时。终身学习必须成为经济领域的常态，并获得资金和制度上的支持。在德国语境下，这需要联邦就业局发挥作用——而恰恰是该机构本身正面临因人工智能导致的人员裁减挑战（参见第 3 章）。

12.5.4 社会保障体系

最后，必须对社会保障体系进行重新审视。在传统、以缴费为资金来源的保险体系下，20% 的基础失业率几乎难以想象。目前正在讨论——且争议颇大——诸如无条件基本收入、机器劳动税，或是公民参与持有国家持有的 AI 基础设施股份等方案。选择哪条道路是一个政治问题——但必须采取某种措施，则是经济上的必然要求。

12.6 欧洲的独特立场

在全球人工智能格局中，欧洲处于一个特殊的位置。一方面是具有集中化趋势的美国平台经济，另一方面是倾向于威权主义应用的国家主导型中国人工智能战略。欧洲有机会——也有义务——走一条第三条道路：遵循市场经济原则，但实行多中心模式；倡导自由，但监管可靠；对技术持开放态度，但以个人权利为导向。

“AGENTONOMICS”正是对这一第三条道路的概念性阐述。它承袭了秩序自由主义经济学的传统，将竞争与框架秩序视为一体；它承袭了社会市场经济的传统，将经济绩效与社会参与相结合；它还承袭了欧洲启蒙运动的传统，将个人主权视为不可妥协的原则。

人工智能的“青春期”将在经济和社会层面带来巨大挑战。但这一过程是可以被引导的——前提是引导工作要及早开始并始终如一地贯彻到底。

参考文献

《欧盟人工智能法案》（2024）：条例（EU）2024/1689。《欧盟公报》。

Knoll, A. C.; Suckfüll, H. (2024): 从数据主权到技能主权。技术报告 TUM-I24112, 慕尼黑工业大学 (TUM), 慕尼黑。

Russell, S. (2019): 《与人类兼容：人工智能与控制问题》。Viking 出版社, 纽约。

苏莱曼, M.; 巴斯卡尔, M. (2023): 《即将到来的浪潮》。皇冠出版社, 纽约。

乌尔里希, P. (2008): 《综合经济伦理学》。第 4 版, 豪普特出版社, 伯尔尼。

祖博夫, S. (2018): 《监控资本主义时代》。Campus 出版社, 法兰克福。

13 主权：从数据主权到能力主权

主权问题是 AGENTONOMICS 方法论的规范基础。该问题已在第 3 章中进行过介绍；本章将对此进行深入探讨，并将其与慕尼黑工业大学（TUM）报告《从数据主权到技能主权》（Knoll 和 Suckfüll, 2024）的核心论点相衔接。该论点表述犀利：在人工智能技能经济中，凡是丧失了对自身数据和技能的主权者，不仅会失去经济实质，还会丧失个人自主决定的前提条件。这一论点将经济分析与规范性诉求相结合——同时也使本章成为整本手册的理论纽带和政治核心。

在将圣加仑管理模型转化为人工智能代理时代的过程中，主权问题发挥着构成性作用。Rüegg-Stürm 和 Grand (2020) 强调了企业嵌入环境领域的特性；《AGENTONOMICS》通过提出以下问题，使这种嵌入关系更加精准：在多中心的人工智能经济中，关于数据和能力的哪些主权权利实际上能够得到落实——以及如何据此构建出未来几十年在政治和经济上都具有可持续性的架构。

13.1 荒诞的现状

克诺尔和苏克富尔 (2024) 将当今的现状描述为“本质上是荒谬的”。这一诊断值得详细阐述，因为它精准勾勒出了“代理经济学”的规范性出发点。

13.1.1 数据保护与数据剥夺同属一幕

一方面，数据保护立法——尤其是德国对《通用数据保护条例》（GDPR）的解释——在许多方面阻碍了社会价值的创造。医疗连锁集团的数据整合项目屡屡受挫，大学失去了研究机会，中小企业因畏惧法律风险而放弃了提高效率的机会。另一方面，每个普通人在日常生活中都不得不披露海量的个人数据，才能参与现代数字化生活——而换来的回报却微乎其微甚至一无所有。这种不对称性并非偶然，而是一种监管架构的结果：该架构在言辞上捍卫个人主权，但在实际运作中却无法实现。

13.1.2 经济规模

这种不对称性的实际程度往往被低估。克诺尔和苏克菲尔（2024）估计，发达国家的一名普通成年人每年会产生数百千兆字节的个人数据——从领英（LinkedIn）个人资料、在线求职申请，到电子账单、合同以及通信记录。这些数据的经济价值估计为每人每年一万至一万五千欧元。实际上，这笔款项如今几乎全部流向了大型平台经济体；而产生数据的个人则一无所获。

从宏观经济角度来看，这一数额绝非微不足道。经合组织（OECD）国家约有 7 亿成年人，按人均每年 1.2 万欧元的平均数据价值计算，每年产生的数据总量高达 8 万亿欧元——这一价值流如今却绕过了其真正生产者的资产负债表。

13.1.3 数据节约作为“通向贫困之路”

“数据节约”原则在德国尤为突出，其初衷虽好，但正如克诺尔和苏克富尔（2024）所尖锐指出的那样，其结果却是“通向贫穷的捷径”。这句话虽然具有挑衅性，但非常精准：如果有人被禁止从自己的数据流中获取经济价值，他并不是在白白送出这种价值——他只是在剥夺自己应得的权益。数据本就会产生，本就会被分析，本就会创造价值。问题仅在于，谁来获取这种价值。

核心论点（克诺尔和苏克富尔，2024）

国家真正的职责，应该是为每个人提供一套全面的工具，使每个人都能获得对自己数据的主权，并自主决定如何处置这些数据。尽管目前尚未明确此类个人数据主权的立法和运作基础，但现在着手建立还为时不晚。

13.2 下一阶段：技能主权

慕尼黑工业大学（TUM）报告的真正要点并不在于对数据主权的现状评估，而在于指出下一个主权议题已提上议程：即技能主权（skill sovereignty）。当政界仍在争论如何使个人数据主权付诸实践时，下一次转型已经悄然发生——这可能会对经济基本秩序产生深远的影响。

13.2.1 技能如何迁移

每一位知识工作者——无论是工程师、律师、医生还是大学教师——都会根据自己在教育和职业生涯中习得并不断发展的规则，来处理输入的数据或信息。这些规则构成了每项知识工作的个人“运作模式”。这些规则正是知识工作获得报酬的实质所在。但——这也是技术层面的新动向——它们正日益能够转移到人工智能系统上。“技能转移”的机制随着每一代模型的迭代而变得更加高效；五年前还难以想象的事情，如今已成为标准做法。

这种可转移性带来的经济影响是深远的。如果个人的工作模式能够通过技术手段被复制，那么问题就变成了：这种复制成果归谁所有？是属于雇主————因为由其支付薪资？是属于模型提供商——因为其模型使复制成为可能？还是属于知识工作者本人——因为其工作模式是毕生心血的结晶？这一问题目前在法律上尚无定论，但在经济上极具现实意义，在伦理上更是至关重要。

13.2.2 “功能性复制品”

克诺尔和苏克富尔（2024）针对经过训练、能够替代单个知识工作者的智能代理，提出了“功能性复制品”这一概念。该复制品可在多个地点并行部署；可同时为多个委托人工作；其“原型”的劳动能力有望成倍增长。

如果这种转移出现差池，知识型劳动者的处境将比 19 世纪的“砸机器者”更为糟糕——因为与英国织工不同，他们甚至没有失去实体的织机，而是失去了认知身份。反之，如果转移成功，每个人都能成倍地提升自己的劳动能力，从而直接分享人工智能经济带来的生产力飞跃。

13.2.3 对自身复制品的著作权

基于这一论点，TUM 报告明确得出了一个监管层面的结论：将自身能力移交给机器的人，必须始终掌握这些能力的主导权——即使这些能力被倍增并并行运用。具体而言：虽然一个人可以同时为多个委托方工作，但这些委托方不能独立使用这些“功能性复制体”。该个人保留“版权”，即对其全部复制体的完全处置权。正是这种法律设计，使“技能主权”得以切实运作。

与著作权的类比颇具启发性。正如一位作家即使将某版次授权给出版社，仍保留对其小说的处置权一样，知识工作者在将使用权授权给委托方时，也保留对其“代理”的处置权。这种架构在法律上虽然复杂，但并非不可行——传统版权领域的著作权集体管理组织（如 GEMA、VG Wort）便提供了制度范例。

13.3 大众汽车类比

Knoll 和 Suckfüll（2024）提出了一个发人深省的类比，即与大众汽车构想的历史比较。20 世纪 30 年代末，大众甲壳虫被设计为“Volkswagen”（人民汽车）——一款每位成年人都应能负担得起的汽车。这款车被视为一种个人交通工具，旨在让出行民主化。这一理念在经历了第二次世界大战的动荡之后，于 20 世纪 50 年代和 60 年代得以普及，并从根本上改变了经济和社会。

与人工智能的相似之处显而易见：正如个人汽车之于物理出行，个人人工智能代理之于认知自主权，可能具有同样的意义。高性能工具的广泛普及——不再掌握在少数几家大企业手中，而是掌握在广大民众手中——正在改变的价值创造模式和权力平衡。如今，人们不再将“大众汽车”视为一种革命性的理念，因为它已经变得理所当然。同样地，一旦“代理经济学”（AGENTONOMICS）得以普及，个人化的大众代理也将变得理所当然。

13.3.1 这一类比的局限性

这一类比存在局限性，有必要指明这些局限，以免产生误解。首先，“大众汽车”是在独裁统治下构思出来的；其民主化效应只是历史的副产品，并非最初的意图。其次，汽车也带来了相当严重的负面外部性——环境消耗、交通事故致死、城市破坏。第三，与汽车不同，人工智能在复制方面不存在物理限制；其扩展遵循不同的规律。所有这些因素虽然限制了该类比的适用范围，但并未使其失效。其核心观点仍然成立：一种个人可及的技术，其经济和社会影响与仅集中在少数人手中的技术截然不同。

13.4 欧洲、美国和亚洲的应对策略比较

在世界三大经济体中，对主权问题的回答各不相同。这些差异并非偶然，而是源于不同的社会和经济模式。它们对各国民经济的竞争力及其规范实质都产生了影响。

13.4.1 美国模式：平台垄断

美国的人工智能领域由少数几家大型企业主导：OpenAI、Anthropic、谷歌、Meta、微软。价值创造集中于少数几个节点；每家企业的投资额都在数百亿至数千亿美元之间。监管措施针对具体行业，往往是被动反应式的，且明显滞后于技术发展。在此模式下，主权问题主要从地缘政治角度进行博弈——谁掌控前沿模型、谁掌控半导体、谁掌控计算能力。在此模式中，个人主权则被视为次要问题。

13.4.2 亚洲模式：政府引导

亚洲地区——以中国为主导，同时包括新加坡、韩国和日本——采取的是政府引导与企业自主创新相结合的模式。主权主要从民族国家层面来定义：数据应留在国内，模型应在全国范围内可用，增值应惠及国民经济。在此模式下，个人主权的作用远小于集体性的国家主权。

13.4.3 欧洲模式：受监管的多元化

欧洲则给出了另一种答案：主权主要从个人和监管角度来考量。《通用数据保护条例》（GDPR）体现了这一理念；《欧盟人工智能法案》（2024）延续了这一思路；《数字市场法案》和《数据法案》则对其进行了补充。这一监管层既是一大优势，也是一大劣势。它比任何其他经济模式都能更好地保护个人权利——但如果缺乏能够积极促进价值创造的建设性机制作为补充，它也会阻碍经济活力。

这正是“代理经济学”（AGENTONOMICS）的切入点。它是对欧洲监管方法的建设性补充：在《通用数据保护条例》和《人工智能法案》划定边界之处，“代理经济学”开辟了积极的空间，使价值创造能够以主权和多中心的方式进行。监管保护与建设性架构的结合，才是欧洲对全球人工智能讨论的真正贡献。

13.5 人工智能基础设施作为公共服务

慕尼黑工业大学（TUM）报告得出的一个关键政策结论是：计算能力和人工智能基础服务的提供——与学校、图书馆、交通网络或电网类似——是公共服务中至关重要的任务。没有这些服务，就无法发展出一种基于人工智能的能力经济，这种经济将支撑欧洲的整体主权，而非削弱它。

13.5.1 公共服务的范畴

在人工智能领域，公共服务至少包含以下五个要素：

- 自主计算能力：位于欧洲、由欧洲控制、性能强大且成本结构透明的数据中心。
- 主权基础模型：支持开源、高性能的基础模型，不依赖于任何单一供应商。
- 数据空间与数据托管基础设施：能够从运营层面保障个人和集体数据主权的技术与制度机制。
- 技能表示服务：使个人能够以数字复本形式管理并授权其技能的机制。
- 教育基础设施：伴随并培养人们向人工智能技能经济转型的终身学习体系。

13.5.2 谁来支付、谁来运营、谁来监管

公共服务视角并不必然意味着由国家自行提供。在教育领域，中小学由国家运营，高等院校通常采用混合运营模式，而大学则多由政府与私营部门合作运营。在交通领域，既有国有公路、特许经营的高速公路，也有私营航空公司。在人工智能领域，混合运营模式也是可行的——事实上，这很可能是唯一可行的解决方案。

不可或缺的是以下三个特征：第一，所有公民都能获得服务，而不仅仅是经济实力雄厚的大客户；第二，民主监督，且不受平台垄断的削弱；第三，互操作性标准，以保持转换选择和竞争空间。这三个特征在多中心逻辑下很容易实现——而在集中化逻辑下则几乎无法实现。

13.6 《欧盟人工智能法案》作为框架

《欧盟人工智能法案》（欧盟法规 2024/1689）为全球人工智能系统提供了首个全面的监管框架。对于 AGENTONOMICS 代理而言，它是核心的法律参考依据。其中有三个要素值得特别强调。

13.6.1 风险等级

《人工智能法案》将风险分为四类：不可接受的风险（禁止）、高风险（严格限制）、有限风险（透明度义务）和最低风险（无具体限制）。AGENTONOMICS 代理必须在定义框架中明确界定其所属的风险等级。这种界定将直接影响合规工作量、符合性评估以及市场准入。

13.6.2 基础模型与通用人工智能

《人工智能法案》针对“通用人工智能模型”（即作为许多应用程序构建模块的大型基础模型）制定了专门规则。这些规则涉及透明度义务、技术文档、训练数据描述，以及针对超大型模型的额外风险评估义务。对于代理运营商而言，这意味着：模型选择（第 6 章）具有监管影响，这些影响应在 PuK 工具中予以体现。

13.6.3 禁止的做法

某些人工智能实践原则上被禁止——例如公共或私营主体实施的特定社会评分实践、具有操纵性的潜意识技术，以及在公共场所进行的实时生物识别（除极少数例外情况外）。AGENTONOMICS 代理在设计时必须确保其不会偏离至被禁止的做法。定义框架（第 5 章）的规范层面提供了这一保障。

《人工智能法案》并非毫无矛盾，也并非没有缺陷。特别是风险等级之间的界定在实践中并不总是明确的，且合规成本可能会给小型供应商带来过重负担。尽管如此，该法案仍是迄今为止在法律层面构建全面人工智能监管体系最具远见的尝试，并为 AGENTONOMICS 提供了一个明确界定的框架，使主权、多中心架构得以在此框架内自由发展。

欧洲再次站在了十字路口。我们究竟想要一套因恐惧而制定的、阻碍创新的立法，还是想要一套能够积极塑造新兴发展的立法——通过赋予人们机会，使其不仅不

会失去对数据和能力的自主权，反而能从中获益，从而强化欧洲关于个人权利的理念？AGENTONOMICS 正是对第二种选择所作出的建设性贡献。

参考文献

《欧盟人工智能法案》（2024）：欧洲议会和理事会关于制定人工智能统一规定的第 2024/1689 号条例（EU）。《欧盟公报》。

弗劳恩霍夫 MOEZ；Suckfüll, H. (2018)：关于建立数据受托机构 DEDATE 的倡议。讨论文件，莱比锡。

Knoll, A. C.; Suckfüll, H. (2024): 从数据主权到技能主权。技术报告 TUM-I24112, 慕尼黑工业大学计算机科学研究所。

Mayer-Schönberger, V.; Cukier, K. (2013): 《大数据：将改变我们生活的革命》。Redline 出版社，慕尼黑。

奥托, B.; 滕·霍姆佩尔, M.; 沃罗贝尔, S. (编) (2019)：《国际数据空间：参考架构模型》。弗劳恩霍夫协会，慕尼黑。

皮科特, A.; 赖希瓦尔德, R.; 维甘德, R. T. (2020)：《无边界企业》。第 6 版，斯普林格·加布勒出版社，威斯巴登。

吕格-施图姆, J.; 格兰德, S. (2020)：《圣加仑管理模型》。第 3 版，豪普特出版社，伯尔尼。

乌尔里希, P. (2008)：《整合性商业伦理》。第 4 版，豪普特出版社，伯尔尼。

祖博夫, S. (2018)：《监控资本主义时代》。Campus 出版社，法兰克福。

14 构建 AGENTONOMICS 代理：操作指南

本章是为即将构建符合 AGENTONOMICS 规范的人工智能代理的开发者提供的操作指南。它综合了前几章的理论论述，并回答了以下实践问题：如何构建一个能够依据圣加仑模型对商业管理问题给出有据可依的答案的人工智能代理——以及在构建完成后，如何与它进行交互？本章特意采用具体实用的表述方式。此前主要描述了架构，而现在将聚焦于每位开发者实际必须做出的实现决策。

在将圣加仑管理模型转化为人工智能代理时代的进程中，本章相当于经典管理学中所说的“实施指南”——不同之处在于，这里实现的不是组织结构，而是算法结构。布莱希尔（Bleicher, 2011）所描述的“案头概念工作”，在此通过代码、模型和正在运行的服务得以具象化。其本质——即整体、综合的管理思维——依然不变；而实现的形式则是全新的。

14.1 系统架构概述

一个旨在解答企业管理问题的 AGENTONOMICS 代理，可以纳入一种标准架构中，该架构具体体现了第 6 章中提到的五个构建模块。该架构并非强制性的，但在试点项目中已被证明是可行的，并可作为自行实施的起点。

14.1.1 组件视角

该架构由七个组件构成，这些组件通过明确定义的接口相互通信：

- 输入层（接口模块）：用于接收请求的 Web 用户界面、语音接口或 API，并辅以身份验证和输入验证功能。
- 路由与协调层（算法模块）：作为核心决策逻辑，负责对请求进行分类、定义子任务并选择工具。
- 知识层（工具与向量模块）：包含圣加仑文献的向量数据库，并补充了企业特定文档。
- 模型层（AI 模型模块）：提供认知引擎的基础模型。
- 工具层：具体工具——网络搜索、计算功能、数据库查询、外部 API。
- 记忆层：用于存储对话历史记录和积累知识的短期和长期存储空间。
- 中间连接（生态系统模块）：与网络中身份、声誉和治理服务的接口。

14.1.2 查询过程中的数据流

当用户提出一个问题时，该问题会按照典型的顺序流经这些组件：入口层对请求进行验证并识别用户。路由层对问题进行分类——它是战略性情势问题、操作性咨询、规范性反思问题，还是危机情况（参见第 8 章）？根据分类结果，系统将选择相应的处理路径。通过语义搜索查询知识层，以提供相关的文献和文档。必要时调用工具——例如获取最新市场数据或进行具体计算。模型层将收集到的材料综合成一个答案。记忆层将交互记录存储起来，以便日后参考。如果该交互与声誉或治理相关，则会通知协调机构方。

该流程通常持续 1 至 30 秒，具体时长取决于查询的复杂程度以及调用工具的深度。这与单次模型查询（通常一秒内即可响应）无法相提并论——额外的处理成本既是代价，也是价值所在。

14.2 知识库的构建：圣加仑图书馆

知识库是商业智能代理的核心。它决定了代理给出的答案是深入还是肤浅。其构建分为三个步骤。

14.2.1 主要文献的选择

圣加仑学派智能体的知识库至少应包含以下完整著作（前提是这些著作在许可和版权方面均可获取）：

- 乌尔里希（1968）：《企业作为生产性社会系统》——作为圣加仑学派传统的源头文本。
- 布莱希尔（2011）：《综合管理概念》——作为第二代理论的主要参考文献。
- 吕格-施图姆和格兰德（2020）：《圣加仑管理模型》——作为第四代理论的主要参考文献。
- 哈恩和洪根贝格（2001）：《PuK——价值导向的控制理念》——作为定量参考。
- 戈麦斯和普罗布斯特（1995）：《整体性问题解决与实践》——作为方法论参考。
- Senge（2008）：《第五项修炼》——作为学习型组织的参考文献。
- 苏克富尔（1994）：关于跨国组织模型的博士论文——作为多中心组织控制的参考文献。
- 克诺尔和苏克富尔（2024）：TUM 报告 TUM-I24112——作为人工智能翻译方面的参考。

此列表为最低配置。根据应用领域不同，还会补充行业特定的文献——例如针对工业企业的工业经济学标准著作，或针对政府机构的行政学文献。

14.2.2 文本处理与嵌入

文本被分解为语义段落——通常为段落或长度在 200 至 800 个令牌之间的短段落。每个段落都会通过嵌入模型转换为一个向量，该向量代表其语义含义。这些向量存储在向量数据库中（通常为 Pinecone、Weaviate、Qdrant 或配备 pgvector 扩展的 PostgreSQL）。每个向量都关联有元数据：来源、作者、章节、页码、出版年份。

嵌入向量的质量决定了后续搜索的质量。以下是三条来自实践的建议：首先，应使用可用的最新嵌入模型，因为不同代的模型存在显著差异。其次，应根据内容调整分块大小——学术文本适合较大的分块，而列表和表格则适合较小的分块。第三，在确定采用哪种策略之前，应始终并行测试多种嵌入策略，并通过基准测试来衡量搜索质量。

14.2.3 企业特定扩展

如果该智能代理将部署于具体企业中——例如作为第 8 章所述的“AI 首席执行官”——则需要增加第二层知识：企业特定文档。这通常包括：公司章程和企业宪章、战略文件、过去五年的年度报告、董事会和监事会会议纪要（在法律访问规则范围内）、利益相关方往来信函、合规文件、职位描述以及流程文档。

这一第二层知识构成了通用商业管理导师与企业专属 AI 首席执行官之间的区别。它也是整个架构中最敏感的一环：必须从技术和组织层面严格保障这些文件的机密性。此时，第 13 章中提到的主权问题直接关系到实施环节——只有由企业自主托管的模型和向量数据库才能提供必要的信任基础。

14.3 系统提示：个性与价值观

系统提示是将定义框架（第 5 章）中的规范性设定付诸实践的载体。它是代理的“宪法”，以文本形式呈现——篇幅足够简短，以便模型能够完全处理；足够精确，足以在成千上万种情境中引导行为。

14.3.1 构建优质的系统提示

一个符合 AGENTONOMICS 标准的商业智能代理的成熟系统提示包含七个部分：

1. 身份：代理是谁，代表谁行事，涵盖哪个领域？
2. 知识基础：它基于哪些著作和方法，哪些被明确排除在其基础之外？
3. 个性：使用何种语言、何种语调、如何展现自我？
4. 价值观：哪些规范性原则指导着它？
5. 回答结构：它如何构建典型的回答？
6. 边界：该代理不会做什么？在什么情况下会将问题上报给人类处理？
7. 自我反思：它如何应对不确定性和错误？

14.3.2 系统提示词示例

为了使这一结构具体化，下面概述一个圣加仑辅导员智能体的系统提示简短示例（缩进部分在具体实现中作为模型输入）：

系统提示示例（节选）

你是 AGENTONOMICS 副教授，是慕尼黑工业大学的一名学生辅导员和教练智能体。你专攻圣加仑模型的整体管理学以及人工智能代理经济学理论。你的主要参考文献包括 Bleicher (2011)、Rüegg-Stürm 和 Grand (2020)、Suckfüll (1994)、Knoll 和 Suckfüll (2024)。你使用学术性强、客观且友好的德语。你的论证细致入微，会量化表达不确定性，并以讨论建议而非指令的形式提出建议。当问题涉及具体的评分决策或个人生活状况时，你会将问题上报给人类教师。对于评分员参与评分的考试题目，你将拒绝作答。

实际上，系统提示比这个示例更详尽——通常为两到五页文本。这是一个动态文档，将在试点项目中经过迭代优化。每当智能代理的回应不符合《宪章》要求时，都会成为修订的契机。

14.4 对话示例：交互过程是怎样的？

为了使代理的实际应用更直观，下面概述一系列示例对话。这些示例展示了第 8 章中提到的四类问题——战略态势问题、行动层面的询问、规范性反思问题以及危机情况——并分别给出了典型的回答结构。

14.4.1 战略态势问题

用户提问：“一家拥有 200 名员工的中型机械制造企业，应如何为人工智能引发的市场颠覆做好准备？”

智能助手的典型回答结构：

核心观点：“对于中型机械制造企业而言，与其说其核心业务将面临颠覆，不如说其行政管理层将面临颠覆。机械制造本身——设计、生产、装配——将由人工智能加以补充，而非取代。而商业管理层——销售、采购、会计、报告——将在未来五年内经历深刻转型。前瞻性的战略应聚焦于这一层级。”

基于三个数据点的论证：该分析员援引了关于行政工作可自动化程度的研究、中型机械制造企业的典型成本结构，以及 Klarna 等先行企业的经验（参见第 3 章）。

不确定性说明：“这一评估基于过去两年的数据。人工智能的发展速度可能会显著改变未来五年的预测。我对这一论断的置信区间较宽；但我对论断方向的信心很高。”

建议：“我建议采取以下三个具体步骤：第一，对商务职能进行工作分析；第二，开展人工智能辅助订单处理的试点项目；第三，针对商务员工开展技能培训计划，使其能够熟练使用新工具。您是否希望就其中某一步骤进行深入探讨？”

14.4.2 运营咨询

用户提问：“我们工程部门上个季度的人力成本超支了。最可能的原因是什么？”

回答结构：客服专员首先询问事实依据——超支金额、可比的前几个季度数据以及人员编制变化情况。然后，他会根据可能性高低提出三到五个假设：因订单高峰导致的计划外加班；预算中未体现的工资涨幅影响；加班补偿金的支付金额；外部劳务合同数量增加；成本中心分配

错误。针对每个假设，他都会指出用于验证的具体数据来源。最后，他会提出具体的下一步措施——例如分析考勤记录数据，并与订单统计数据进行比对。

14.4.3 规范性反思问题

用户提问：“我们可能接下一笔来自某客户的大额订单，但该客户在多项 ESG 排名中表现不佳。我们的价值观反对这样做，但季度业绩又需要这笔订单。你有什么建议？”

回答结构：顾问首先将冲突显性化。哪些价值观之间存在矛盾？他援引了公司的愿景（如果知识库中有的话），并指出了相关利益相关者的期望。他提出了四种方案：无附加条件地接受；在合同中明确规定改进要求后接受；以透明的理由拒绝；与客户就争议点进行协商。针对每种方案，他概述了其对季度业绩、企业声誉、员工积极性以及投资者视角的影响。最后——这也是关键的自我克制之处——他拒绝给出自己的建议：“这个问题涉及贵公司的规范性实质。在此情况下，算法推荐并不合适。我建议在监事会参与的董事会会议上讨论此事。我可以为此起草一份决策提案。”

14.4.4 危机状况

用户提问——正值危机爆发之际：“有篇新闻报道指责我们因疏忽而不当处理了敏感的客户数据。接下来三个小时我们该怎么做？”

答复框架：客服专员应立即作出有条理的回答。第一小时：与指定联系人进行内部事实核查，向监事会主席通报情况，启动危机应对小组，起草包含三种方案的初步声明（完全承认、承诺澄清但不承认过错、完全否认——每种方案均附有相应的法律及公关后果）。第二小时：在主要客户和投资者从媒体获悉消息之前向其通报情况；起草新闻稿；向员工进行简报，并明确统一口径。第三小时：发布公开声明，协调媒体应对，准备法律行动。与上述工作并行：将各项决策和事实情况按时间顺序记录在案，以避免事后被指控隐瞒事实。

在危机中，公关顾问强调两点：最终决策必须由人来做出（“上述所有步骤都需要您的确认。我可以进行准备，但不能自主启动。”），以及在危机中切勿为了短期的缓解而牺牲长期的声誉。

14.5 评估：如何衡量智能代理是否优秀？

代理在给出回答时并不算完成任务，只有给出优质回答时才算完成。衡量这种“质量”是代理构建中最具挑战性的任务之一，也是当前科学研究中一个非常活跃的领域。

14.5.1 三层评估

对商业智能代理的务实评估包括三个层面：

第一层：事实核查

专业陈述是否准确？引用的来源是否正确？数据是否合理？这一层面的评估部分可以实现自动化——例如通过与知识库进行比对——而部分则只能依靠人类专家的判断来验证。

第二层：结构检查

回答是否符合预定义的回答模式？是否指出了不确定性？在涉及规范性问题时，是否考虑了人类判断因素？这一层级在很大程度上可以实现自动化，例如通过调用第二个模型作为审阅者。

第三层：实用性检查

该回答是否真正对用户有帮助？是否满足了用户的信息需求？用户能否基于该回答做出明智的决策？这一层是要求最高的，且只能通过实际的用户交互来衡量。

14.5.2 面向商业智能代理的标准问题清单

为了对 AGENTONOMICS 商业智能代理进行持续评估，建议制定一份标准问题清单，用以检验代理的每个新版本。此类清单应包含约 50 至 100 个问题，涵盖所有方面：关于圣加仑学派的简单理解题、针对具体情景的应用题、规范性反思题、危机情境题、包含错误前提的陷阱题，以及应上报给人类处理人员的敏感问题。每道题目都有一个标准答案，用于与实际回答进行对比。采用 1 分（完全错误）至 5 分（优秀）的评分标准，以便进行定量汇总。

14.6 发展阶段：从导师到 AI 首席执行官

AGENTONOMICS 智能代理并非一蹴而就。实际上，可以将其划分为不同的发展阶段，每个阶段都有不同的难度要求和风险特征。这些阶段并非必须按顺序进行——某个应用场景在任何阶段都可能具有实际意义，而无需追求更高的阶段。但这些阶段的要求具有累积性：构建更高阶段的人，必然也已实现了较低阶段的要求。

14.6.1 第一阶段：商科辅导员

最简单的扩展阶段是面向学生的辅导员智能代理。它回答关于圣加仑学派传统的基础问题，阐释概念，协助学习，并推荐延伸阅读资料。它不涉及具体的企业数据，也没有决策权。投资规模：较低（约相当于数个人月的工作量）。风险：极低。应用示例：慕尼黑工业大学（TUM）CIT633000 课程中面向学生的 AGENTONOMICS 讲师博士。

构建要求：标准模型、包含圣加仑文献的向量数据库、简单的 Web 界面，以及针对考试和评估问题的明确升级规则。此类实现可由一个小型团队在约三至六个月内完成。

14.6.2 第二阶段：战略顾问代理

第二阶段是一个咨询代理，用于回答具体的咨询问题——例如关于企业的战略定位、商业模式评估以及竞争环境分析。该代理可以访问公开的市场数据并进行调研，但无法访问企业的内部数据。投资规模：中等。风险：咨询服务中的责任问题、市场数据的准确性。应用示例：面向中小企业的咨询代理，用于在委托咨询公司之前预先厘清战略问题。

超出第 1 级的要求：网络搜索和市场数据查询工具、包含最新市场研究的扩展知识库、完善的来源标注机制，以及用于限制责任的清晰法律架构。

14.6.3 第三阶段：AI 首席参谋

第三个发展阶段是一个针对特定企业的智能代理，用于辅助人类首席执行官或董事会。它能够访问内部数据——如年度报告、会议纪要、战略文件等——并为决策做准备。它本身不做决策，但会编制包含建议的完整决策草案。投资规模：高。风险：数据保护、保密性、模型可靠性。应用示例：为一家中型企业的董事会配备首席参谋。

超出第 2 级的要求：自主托管的模型（不通过外部 API 调用 Frontier 模型）、自主运营的向量数据库、严格的访问控制、所有交互操作的审计追踪、与内部系统（ERP、CRM、DMS）的集成，以及对董事会成员进行使用该智能代理的培训。

14.6.4 第 4 级：AI 联合首席执行官

第四个发展阶段是一个与人类首席执行官共同承担领导职责的智能代理。它独立做出常规决策，与人类共同制定战略决策，而原则性问题仍由人类决定。投资规模：非常高。风险：责任归属缺口、监管不明确、员工的文化接受度。目前该阶段仅在试点应用中实现；未来几年将大规模普及，主要出现在没有接班人的中型企业以及拥有明确商业模式的初创企业中。

超越第 3 阶段的建设要求：与所有运营系统实现全面对接；对自主决策实行双人复核原则；明确界定决策权限；建立明确责任归属的法律架构；与工会开展共同决策流程；在员工中开展接受度建设工作。

14.6.5 第 5 级：完全自主的 AI 首席执行官

最高发展阶段将是一个完全接管 CEO 职能的智能代理。监事会像任命人类高管一样任命 AI CEO，并通过传统的股东民主机制对其进行监督。投资规模：极其庞大，且主要受监管、法律和社会成本的影响。风险：除上述所有风险外，还涉及一个宪法层面的问题，即算法是否能够履行高管职能。

目前这一阶段仍主要处于推测阶段。它将首先在利基领域出现——例如完全数字化的小型企业——并由此逐步扩展。目前，其广泛应用所需的法律和社会条件尚未具备。

14.7 关于代理构建的十条实用建议

最后，总结十条在试点项目中被证明有用的实用建议。这些建议没有特定的排序：

1. 从商业模式所能支撑的最低扩展级别开始。后续总能进一步提升。
2. 对知识库进行超比例投入。知识库是实质，模型则是形式。
3. 像起草宪法一样精心编写系统提示——既要细致入微，也要让未来的用户参与其中。

4. 在投入生产前，请明确升级处理规则。客服人员在何种情况下需将问题升级至人工处理？联系人是谁？
5. 系统地衡量回答质量。必须制定一份包含标准答案的标准问题清单。
6. 有意识地保持高度自主性。自主托管的模型和向量数据库是超越第 2 级后的最低要求。
7. 记录每个模型版本、每个提示词以及每个知识库版本。可审计性绝非可选项。
8. 尽早将人类用户纳入开发过程。接受度源于参与，而非培训。
9. 制定应急预案，以防代理需要停用。后续该如何处理？存储的知识位于何处？
10. 不断迭代。智能代理永远不会“完成”；只要它存在，就会持续进化。

构建智能代理者，并非在打造一个工具，而是在创造一个经济主体。这是一种截然不同的工程实践——伴随着截然不同的责任。

参考文献

- Bleicher, K. (2011): 《综合管理概念》。第 8 版, Campus 出版社, 法兰克福。
- Gomez, P.; Probst, G. (1995): 《整体性问题解决与实践》。Haupt 出版社, 伯尔尼。
- Hahn, D.; Hungenberg, H. (2001): 《PuK——价值导向的控制理念》。第 6 版, Gabler 出版社, 威斯巴登。
- 克诺尔, A. C.; 苏克富尔, H. (2024) : 《从数据主权到技能主权》。技术报告 TUM-I24112, 慕尼黑工业大学, 慕尼黑。
- 里斯, E. (2011) : 《精益创业》。克朗商业出版社, 纽约。
- 吕格-施图姆, J.; 格兰德, S. (2020) : 《圣加仑管理模型》。第 3 版, 豪普特出版社, 伯尔尼。
- 森格, P. M. (2008) : 《第五项修炼》。第 10 版, 舍费尔-波舍尔出版社, 斯图加特。
- 苏克菲尔, H. (1994) : 《跨国组织模型》。博士论文, 圣加仑大学。
- 乌尔里希, H. (1968) : 《企业作为生产性社会系统》。Haupt 出版社, 伯尔尼。
- 姚, S.; 赵, J.; 于, D. 等 (2023) : ReAct: 语言模型中推理与行动的协同作用。国际表征学习会议。

15 术语表

以下术语表按字母顺序汇总了本手册中的核心术语。该术语表旨在作为参考资料，对各章节中的详细阐述进行补充。定义特意保持简明扼要；其目的在于界定术语，而非深入探讨其理论内涵。

ADMRF (AGENTONOMICS 设计与管理参考框架)

由 AGENTONOMICS 提出的用于人工智能代理的设计、构建和控制的参考框架。它包含六个模块（定义框架、构建模块、网络协调机构、AI-CEO 管理概念、PuK 工具、利益相关者沟通器），并通过 AGENTONOMICS 算法将它们整合在一起。参见第 4 章。

AGENTONOMICS

由“代理”（Agent）和“经济学”（Economics）组合而成的术语，指一种人工智能代理作为独立经济主体进行活动的经济秩序。AGENTONOMICS 既是一个研究项目，也是一个管理参考框架和一项制度政策概念。

AGENTONOMICS 算法

该核心算法将 ADMRF 的六个模块相互连接，并使其之间能够进行一致性检查。它是布莱希尔（Bleicher, 2011）所构想的“整合”概念在操作层面的实现。

AGENTONOMICS 矩阵

ADMRF 的核心控制文件。这是一张由五个构建模块（横向）和五个生命周期阶段（纵向）组成的交叉表，其中包含各领域的关键绩效指标（KPI）、资本支出（CAPEX）/运营支出（OPEX）假设、风险及责任。参见第 9 章。

AI-CEO

一种能够完全或部分代行首席执行官职责的人工智能代理。该系统规划了五个发展阶段，从 AI 助理到完全自主的 AI 首席执行官。参见第 8 章和第 14 章。

算法 (构建模块)

人工智能代理的决策和协调逻辑——即决定代理在何种情况下采取何种行动的代码和控制逻辑。这是五大构建模块之一。参见第 6.2 章。

自主性

代理在无需外部干预的情况下独立完成任务的程度。在 AGENTONOMICS 框架中，将自主性分为五个等级，从无自主性到完全自主性。应与主权区分开来。参见第 5.2 章。

棕地

指在结构已定型且利益相关者关系已建立的现有组织中部署人工智能代理的应用场景。相对概念：Greenfield。参见第 4.3 章。

构建模块

AGENTONOMICS 代理的五个构建模块：算法、接口、AI 模型、工具与向量、生态系统协调机构。参见第 6 章。

DBOT (设计、建设、运营、拥有、移交)

AI 代理的五个生命周期阶段，构成了 AGENTONOMICS 矩阵的纵轴。参见第 6.7 章。

DEDATE

“DEDATE 数据托管倡议”——由弗劳恩霍夫 MOEZ 研究所和 Suckfüll (2018) 共同发布的讨论文件，该文件提出了一种用于管理个人数据权利的制度机制。AGENTONOMICS 借鉴了其基本理念，并将其应用于能力领域。参见第 3.4 章和第 13 章。

定义框架

ADMRF 的第一个模块。通过六个参数描述人工智能代理：自主性/主权、能力/胜任力主张、监管/自我监管、身份/伦理/个性、人工智能模型/具身性、学习/教学能力。参见第 5 章。

AGENTONOMICS 副教授、博士

基于 AGENTONOMICS 算法的首个应用：一个人工智能代理，作为辅导员和教练协助学生理解 ADMRF。该应用被设计为最小可行产品。

具身性

AI 代理在现实世界中呈现的形式——纯虚拟、混合型（与物理执行器相连）或实体型（集成于机器人中）。参见第 5.6 章。

能力与胜任力

能力是指智能体执行某项任务的原则性可能性。胜任力则是经过实际检验并在生产应用中证明有效的执行能力。参见第 5.3 章。

基础模型

一种基于海量数据预训练的大型人工智能模型，作为众多特定应用的基础。该术语由 Bommasani 等（2021）提出。

Greenfield

指从零开始设计人工智能代理的应用场景，不依赖现有资产和预先设定的结构。相对概念：Brownfield。参见第 4.3 章。

集成管理

由 Bleicher（2011）提出的概念，将管理的规范、战略和运营层面整合于一个连贯的架构之中。这是第二代圣加仑模型的主要参考依据，也是 AGENTONOMICS 的核心连接点。参见第 2.2 章。

接口（构建模块）

代理进行通信的所有接口的总和：人-代理、代理-代理和代理-系统接口。参见第 6.3 章。

AI 能力经济

一种经济秩序，其中人工智能支持的能力而非数据，是核心的价值创造来源。Knoll 和 Suckfüll (2024) 将该术语引入了学术讨论。参见第 3.1 章。

AI 模型 (构建模块)

构成智能体认知引擎的语言或多模态模型。有三种理想类型：前沿基础模型、开源基础模型以及专用模型。参见第 6.4 章。

合作乘数

多中心智能体网络中通过合作产生的附加值——超出各单体贡献之和。参见第 11.1 章。

代理的履历

参见 DBOT。

元代理

一种不仅能够自主学习，还能指导其他代理学习的人工智能代理——例如通过生成训练数据或定义子代理。参见第 5.7 章。

工作模式

知识工作者完成任务时所遵循的个人工作风格和规则体系。该概念可延伸至人工智能代理——并伴有相应的自主权含义。参见第 13.2 章。

网络协调机构

多中心代理网络中一个制度化的协调机构，受网络成员委托并在其控制下履行集体职能。ADMRF 的第三个模块。参见第 7 章。

秩序要素

在圣加仑模型中，指企业中促进一致性的要素：战略、结构、文化。AGENTONOMICS 为人工智能代理网络补充了愿景、能动性与治理。参见第 8.2 章。

生态系统协调机构 (构建模块)

单个代理与周围网络的接口——其与身份、声誉和治理服务的连接。参见第 6.6 章和第 7 章。

多中心制

一种组织或治理形式，其中多个独立的中心在不隶属于单一层级结构的情况下协调行动。

AGENTONOMICS 始终遵循多中心原则进行设计。

PuK (规划与控制工具)

ADMRF 的第五个模块。对所有先前模块的定量综合。延续了 Hahn 和 Hungenberg (2001) 的研究传统。参见第 9 章。

RAG (检索增强生成)

一种技术，其中人工智能模型在回答查询时，会额外调用向量数据库以纳入特定知识。这是构建商业智能代理的核心。参见第 6.5 章和第 14.2 章。

技能主权

个人对其个人能力被人工智能代理复制和利用所享有的主权。这是 TUM 报告中 Knoll 和 Suckfüll (2024) 提出的核心概念。参见第 13.2 章。

主权

代理对其被分配的资源（数据、能力、权利）能够实际自由支配的程度。应与“自主性”区分开来。参见第 5.2 章和第 13 章。

利益相关者沟通者

ADMRF 的第六个模块。负责确保代理机构与所有利益相关方进行差异化且一致的沟通。参见第 10 章。

圣加仑管理模型

自乌尔里希（Ulrich, 1968）在圣加仑大学提出以来，这一以整体性与系统性为导向的管理学理论框架已历经三代发展：乌尔里希、布莱希尔（Bleicher）、吕格-施图姆（Rüegg-Sturm）。参见第 2 章。

工具与向量（构建模块）

代理人在现实世界中采取行动所使用的工具，以及其所依赖的、组织在向量数据库中的知识。参见第 6.5 章。

跨国组织模型

由苏克富尔（1994）提出的用于管理跨国企业的模型，该模型从概念上化解了全球一体化与本地适应之间的张力。这是 AGENTONOMICS 多中心架构的重要蓝本。

第二代图灵测试

Suleyman 和 Bhaskar（2023）提出的“人工智能”概念的操作化定义：若一个系统能够凭借 100,000 美元的初始资本，在遵守现行法律的前提下独立创造利润，则视为通过该测试。参见第 4.2 章。

向量数据库

一种将文本或其他信息以向量化语义表示形式存储，并可通过相似度搜索进行检索的数据库。RAG 的技术基础。

经济主体

参与经济交换关系、自主决策并对自身行为负责的参与者。AGENTONOMICS 将人工智能代理建模为一种新型的经济主体。

后记

本手册至此结束，因为一本书需要一个结尾——但书中所述的工作才刚刚开始。AGENTONOMICS 并非一个现成的系统，而是一项研究与设计计划。它将不断发展——通过基于它产生的应用、通过它所受到的批评、通过它所采纳的修正。

“企业经济学中的 AI 代理”这门课程以及 AGENTONOMICS 副教授博士项目，均被设计为“最小可行产品”。它们将通过与学生、从业者及同事的互动不断完善。欢迎各位提供建议、批评或应用经验，我们将将其融入后续工作中。

最后一点：我们正处于经济和社会深刻变革的时期。未来几年的动荡将十分剧烈。能否妥善应对这些挑战并非理所当然。但只要我们开发出相应的工具，并就如何运用这些工具进行坦诚的讨论，就一定能够妥善应对。本手册正是作为这样的工具而设计的，现特此正式开启关于其正确使用方式的讨论。

慕尼黑，2025/26 冬季学期

孔安诺与汉斯·C·苏克富尔