e-book

Ferdinand-Steinbeis-Institut

UNLEASHING AUTONOMIZATION: A HOLISTIC ONTOLOGY OF DATA-DRIVEN SOLUTION DESIGN REALIZING BUSINESS OPPORTUNITIES

BASED ON DISTRIBUTED LEDGER TECHNOLOGY AND MACHINE LEARNING

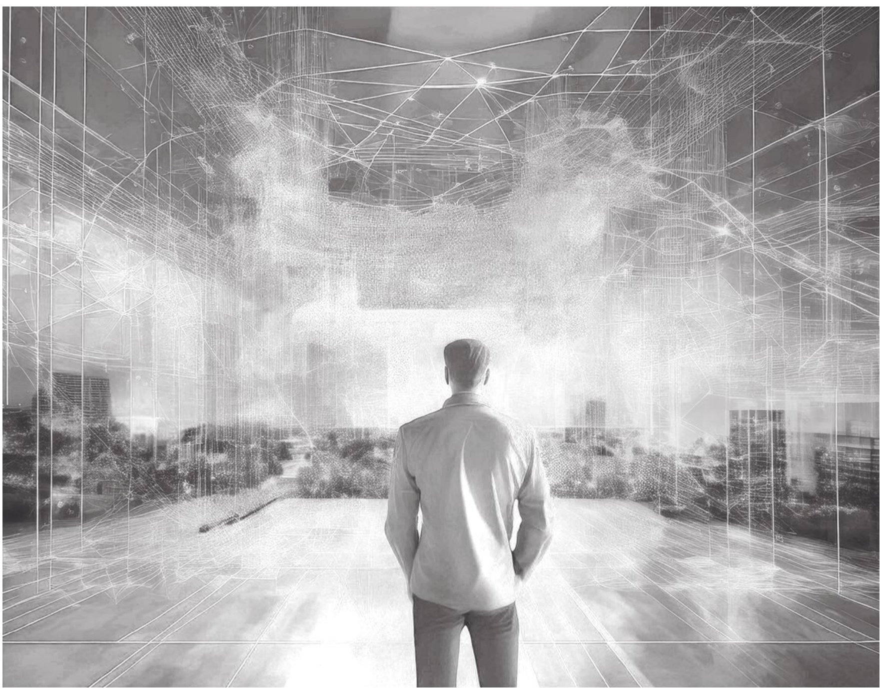
Daniel Burkhardt

Steinbeis-Edition

We can’t solve problems by using the same kind of thinking we used when we created them.

Albert Einstein

Image generated using Adobe Firefly.



Wir können Probleme nicht mit der gleichen Art des Denkens lösen, mit der wir sie geschaffen haben.

Albert Einstein
(übersetzt aus dem Englischen)

Ferdinand-Steinbeis-Institut

UNLEASHING AUTONOMIZATION: A HOLISTIC ONTOLOGY OF DATA-DRIVEN SOLUTION DESIGN REALIZING BUSINESS OPPORTUNITIES

BASED ON DISTRIBUTED LEDGER TECHNOLOGY AND MACHINE LEARNING

Daniel Burkhardt

Steinbeis-Edition

Notice pursuant to the German General Act on Equal Treatment:
In the interest of readability, no distinctions are drawn between different genders. All personal designations apply equally to all genders within the meaning of the German General Act on Equal Treatment.

Manufacturer within the meaning of the General Product Safety Regulation (GPSR):
Steinbeis-Stiftung für Wirtschaftsförderung | Steinbeis-Edition, Adornostraße 8, 70599 Stuttgart, DE | edition@steinbeis.de

Impressum

Published in the series of the Ferdinand Steinbeis Institute on Digital Business Transformation | Volume 2
Heiner Lasi, David Rygl, Dirk Slama, Daniel Werth, Jens Lachenmaier, Marlene Gottwald (Eds.)

Daniel Burkhardt
Unleashing Autonomization: A Holistic Ontology of Data-Driven Solution Design realizing Business Opportunities. Based on Distributed Ledger Technology and Machine Learning.

1st edition, 2025 | Steinbeis-Edition, Stuttgart
ISBN 978-3-95663-315-7 | This book is also available as printed version: ISBN 978-3-95663-314-0
Likewise Steinbeis University, dissertation 2024

Layout: Daniel Burkhardt | Technical Editing: Steinbeis-Edition
Publishing house: Steinbeis-Edition | Steinbeis-Stiftung, Adornostraße 8, 70599 Stuttgart

The platform provided by Steinbeis makes us a reliable partner for company startups and projects. We provide support to people and organizations, not only in science and academia, but also in business. Our aim is to leverage the know-how derived from research, development, consulting, and training projects and to transfer this knowledge into application – with a clear focus on entrepreneurial practice. Over 2,500 business enterprises have already been founded on the back of the Steinbeis platform. The outcome? A network spanning approximately 5,000 experts in over 1,000 business enterprises – working on projects with more than 10,000 clients every year. Our network provides professional support to enterprises and employees in acquiring competence, thus securing success in the face of competition. Steinbeis-Edition publishes selected works mirroring the scope of the Steinbeis Network expertise.

231659-2025-05 | www.steinbeis-edition.de | edition@steinbeis.de

Preface

Choosing appropriate technologies for testing and implementing business scenarios, particularly in designing autonomous systems, presents various challenges for companies. Technical solutions must thoroughly account for design aspects at the business level, system level, data level, and technology level, while considering interdependencies.

With the increasing importance of data-driven business models based on autonomous systems for the future viability of numerous companies, the swift and targeted implementation of business model innovations has become essential for economic success. Consequently, this issue is gaining substantial relevance in both academic and practical domains. Mr. Burkhardt's dissertation addresses this gap by bridging the divide between business model innovation and the targeted design of technical aspects within the realm of autonomous systems.

The goal of the doctoral project is to create an artifact that bridges data-driven business opportunities with technological capabilities. This dissertation addresses a crucial area for both science and practice. By answering the main research question and developing the artifact, it contributes significantly to business informatics and offers practical solutions to relevant problems.

I congratulate Daniel Burkhardt on his successful work that serves scientific and economic benefit and emphasizes the principle of the Ferdinand Steinbeis Institute: "Science must sustainably benefit the economy and society."

Heilbronn, March 2025
Prof. Dr. Heiner Lasi

Preface by Editors

The Ferdinand Steinbeis Institute (FSTI) is a non-profit and interdisciplinary non-university affiliated research institute with a focus on transformation in business and society. Since its foundation in 2015, FSTI employees have been committed to ensuring that all research projects meet the requirement of creating sustainable benefits for the economy and society through relevant knowledge artifacts against the background of social responsibility.

Daniel Burkhardt's doctorate is also based on this claim. It addresses the technical design of “autonomous systems” and, in this context, addresses the problem that currently only rudimentary empirical knowledge is available among a small number of experts. This results in a critical situation from a business and social perspective, which on the one hand hinders progress and on the other hand can lead to untrustworthy systems. The approach of using expert systems to help design autonomous systems is therefore highly relevant both from a scientific point of view and in terms of practical application. In accordance with the FSTI's claim, the author not only develops a scientific artifact in the form of an ontology for the design of data-driven solutions in this dissertation but also implements it in a software pilot. With the help of this tool, practitioners can gain initial experience in this new subject area and derive requirements for their own tool support.

This dissertation clearly demonstrates that scientific progress and real-world impact are not contradictory but synergetic goals of a demanding and challenging dissertation.

The Editors
Heiner Lasi, David Rygl, Helmut Schneider, Eva Deuchert, Jens Lachenmaier,
Dirk Slama, Daniel Werth

Acknowledgments

I had the pleasure of creating this thesis while being at the Ferdinand Steinbeis Institute. Therefore, I would like to give thanks for the provided environment that challenged my ideas and presented new pathways for creative solution finding.

Special thanks go to Prof. Dr. Heiner Lasi, who trusted me in my proceedings and gave me scientific freedom for exploration. His constructive feedback showed me new ways of thinking and broadened my academic vision. Besides that, he gave me access to an international business network that supported me in understanding the current complex field of data-driven solution design. Participating in the Industry IoT Consortium helped me to find new information and methods for improving the emerging artifact. Furthermore, I want to express gratitude to Prof. Dr. Dirk Slama. Through interactions with him and during our collaboration that started the AIoTLab, I received an environment of technological developments.

The interdisciplinarity at the Ferdinand Steinbeis Institute made it possible to work with colleagues from different fields of study. Thus, I got challenged by new impulses to problems that I was facing, for which I am thankful. Dr. Patrick Weber provided a condensed business perspective that let me align on this level. Alexander Neff tweaked my methodological understanding. Furthermore, the development and testing of new technologies were only possible with a great team of students that worked together with me or conducted their thesis in my field of research. For collaboration with all my other colleagues, I am the most thankful.

To my colleagues from the chair of ABWL and Information Systems I of the Institute of Business Administration at the University of Stuttgart, I am incredibly grateful for their support while finalizing my thesis. With the help of Dr. Dominik Morar, the thesis received its necessary focus and Dr. Jens Lachenmaier challenged the Technology level. Both guaranteed a well-founded academic placement of this thesis.

This research result was only achievable with the patience, support, and interest of my family and friends. The relentless conversations with my parents and brothers about my progress motivated me to show new results and insights. Discussions with my wife, Ines

Burkhardt, about creativity and her artwork inspired me to try new ways of reflection and ideation. Many thanks to all for this.

Heilbronn, 10.07.2024

Daniel Burkhardt

Abstract

Dall-E, ChatGPT, and Cruise self-driving vehicles are good examples of solutions that seamlessly incorporate massive amounts of data, advanced hardware resources, and complex AI models. These data-driven solutions are designed to generate corporate value by strategically leveraging changing knowledge. Emerging technologies play an important role in unlocking this knowledge, which then influences existing processes, supports corporate decisions, and facilitates interactions between different organizational layers. However, achieving these objectives is difficult due to the high dynamics involved, quick technological improvements, increasing data complexity, and complicated implications across organizational levels. Despite such challenges, increasing competitive pressure forces organizations to construct information systems that not only assure flexibility and stability, but also rely on emerging knowledge to generate business value. Constant changes in the environment necessitate rapid alterations to system requirements, emphasizing the importance of building data-driven solutions for autonomization. For example, the development of large language models in applications such as ChatGPT has provided novel alternatives to typical human question-and-answer systems, showing tremendous untapped business possibilities.

The uniqueness of collecting such knowledge distinguishes enterprises. Information technology facilitates the implementation of processes while also allowing for the management of complexity, which generates knowledge. Technologies such as the Internet of Things, Distributed Ledger Technology, and Machine Learning were developed to help with complexity management and transparency. However, building data-driven solutions that integrate diverse technologies with an extensive variety and intricate data-based relationships is complex, making complete and value-added implementation difficult.

The conceptualization of technologies improves transparency across their diversity and allows for the selection of technological variants that bring value to business opportunities. However, there is a lack of guidance to assist organizations in managing their diverse requirements for data-driven solution design in the context of digital transformation and innovation. The goal of this thesis is to create a model that can help with technology selection for data-driven solutions. To this purpose, Distributed Ledger Technology and

Machine Learning are described using several conceptual models based on requirements and design principles observed in three case studies from the healthcare, mobility, and risk management industries. The study follows a Design Science Research approach, basing its deductive reasoning on literature studies to find technological concepts and relationships within the existing knowledge base. Furthermore, expert interviews within these case studies aid inductive reasoning by highlighting domain-specific concepts, patterns, and relationships. A logistics case study that focused on assessing new business opportunities in relation to autonomization requirements aided in the development of specific requirement dimensions and characteristics. All of the created artifacts, including requirements dimensions, an architectural patterns matrix, and conceptual models, were incorporated into an ontology to ensure a thorough design of data-driven solutions focused at autonomization. The qualitative research resulted in a prototype implementation of a knowledge graph-based recommendation system, which was further reviewed through expert interviews.

Ultimately, the ontology translates potential business opportunities into the domain of data-driven solution development with the goal of achieving autonomization. During the solution design process, the ontology incorporates evolving knowledge across business and technology layers. This helps in managing complexity and allows for more focused value generation. Essentially, this approach enables the evaluation and explanation of data-driven solutions based on ontological concepts, encouraging the exploration of different elements and innovative solution paths, thus providing assurance and guidance. Furthermore, the created artifact promotes scientific discourse on the abstraction of technologies and enhances the associated knowledge repository by introducing design principles, patterns, conceptual models, and the ontology itself. Subsequent investigations should enhance the ontology by integrating more technologies, implementing it in various domains, and developing prototypes based on it. In the future, the ontology has the potential to improve creative problem-solving by connecting existing solutions with various business opportunities based on conceptual similarities. This approach entails the identification, abstraction, and application of basic ideas or principles, like to those associated with razor blades or natural mechanisms, across other fields.

Abstract (deutsch)

Dall-E, ChatGPT und selbstfahrende Fahrzeuge von Cruise sind Beispiele für Lösungen, die große Datenmengen, fortschrittliche Hardware-Ressourcen und KI-Modelle effektiv integrieren. Diese datengesteuerten Lösungen werden mit dem Ziel entwickelt, Nutzen durch den Einsatz sich entwickelnden Wissens zu generieren. Neue Technologien spielen eine entscheidende Rolle bei der Erschließung dieses Wissens, das wiederum bestehende Prozesse beeinflusst, Geschäftsentscheidungen unterstützt und die Interaktion zwischen verschiedenen organisatorischen Ebenen erleichtert. Die Realisierung dieser Aspekte steht jedoch Herausforderungen gegenüber, die durch die hohe Dynamik, schnelle technologische Entwicklungen und zunehmende Datenkomplexität sowie komplexe Verflechtungen über Organisationsgrenzen hinweg bedingt sind. Trotz dieser Herausforderungen zwingt der wachsende Wettbewerbsdruck Organisationen dazu, Informationssysteme zu entwickeln, die nicht nur Flexibilität und Stabilität gewährleisten, sondern auch das sich entwickelnde Wissen nutzen, um Nutzen zu schaffen. Ständige Veränderungen in der Umgebung erfordern schnelle Anpassungen der Systemanforderungen, was die Bedeutung des Designs datengesteuerter Lösungen zur Autonomisierung unterstreicht. Zum Beispiel hat die Entwicklung großer Sprachmodelle in Applikationen wie ChatGPT innovative Alternativen zu traditionellen Systemen der menschlichen Fragebeantwortung eingeführt, die bedeutendes, noch unerschlossenes Geschäftspotenzial offenbaren.

Die Individualität des Erwerbs solchen Wissens differenziert Organisationen. Informationstechnologie unterstützt die Implementierung von Prozessen und ermöglicht auch die Handhabung von Komplexität, aus der Wissen generiert wird. Technologien wie das Internet der Dinge, Distributed Ledger Technology und maschinelles Lernen wurden entwickelt, um die Komplexitätshandhabung zu unterstützen und Transparenz zu ermöglichen. Allerdings ist das Design datengesteuerter Lösungen, die verschiedene Technologien mit großer Vielfalt und komplexen datenbasierten Beziehungen integrieren, komplex, was eine umfassende und wertschöpfende Implementierung erschwert.

Die Konzeptualisierung von Technologien bietet Transparenz ihrer Vielfalt und ermöglicht die Auswahl einer geeigneten Technologievariante für die zu realisierende geschäftliche Opportunität in nutzenstiftende Lösungen. Es fehlt jedoch eine Anleitung, die

Organisationen unterstützt, eine Vielfalt von Anforderungen für das Design innovativer Lösungen zu verwalten, um die digitale Transformation und Innovation zu ermöglichen. Ziel dieser Arbeit ist die Entwicklung eines Modells, das die Auswahl von Technologien in datengesteuerten Lösungen unterstützt. Daher wurden Distributed Ledger Technology und maschinelles Lernen durch die Erstellung verschiedener konzeptioneller Modelle konzeptualisiert, basierend auf Anforderungen und Gestaltungsprinzipien aus drei Fallstudien in den Bereichen Gesundheitswesen, Mobilität und Risikomanagement. Einem umfassenden Design Science Research-Ansatz folgend, bildeten Literaturrecherchen die Basis für deduktives Schlussfolgern zur Identifikation von technologischen Konzepten und Beziehungen aus der Wissensbasis. Zusätzlich garantierten Experteninterviews als Teil der Fallstudien induktives Schlussfolgern für die Identifikation von domänenspezifischen Konzepten, Mustern und Beziehungen. Anforderungsdimensionen und Charakteristiken für die Autonomisierung wurden in einer Fallstudie im Bereich der Logistik entwickelt, um neue Geschäftsopportunitäten auf Anforderungen zur Autonomisierung zu definieren. Alle entwickelten Artefakte in Form von Anforderungsdimensionen, einer Architektur-Mustermatrix und konzeptionellen Modellen wurden in einer Ontologie kombiniert, um ein umfassendes Design datengesteuerter Lösungen für Opportunitäten der Autonomisierung zu garantieren. Die qualitative Exploration wurde mit der prototypischen Implementierung eines auf einem Wissensgraphen basierenden Empfehlungssystems abgeschlossen, das in weiteren Experteninterviews diskutiert wurde.

Die Ontologie ermöglicht eine Übersetzung von Geschäftsopportunitäten in den Bereich des Designs datengetriebener Lösungen zur Autonomisierung. Im Verlauf des Lösungsdesignprozesses integriert die Ontologie entwickelndes Wissen über die Ebenen von Geschäft und Technologie hinweg, was die Handhabung von Komplexität erleichtert und eine gezieltere Wertschöpfung ermöglicht. Praktisch gesehen erlaubt dieser Ansatz die Erläuterung datengesteuerter Lösungen, fördert die Entdeckung alternativer Komponenten und innovativer Lösungswege und bietet somit Sicherheit und Orientierung. Darüber hinaus regt das entwickelte Artefakt den wissenschaftlichen Diskurs über die Abstraktion von Technologien an und bereichert die entsprechende Wissensbasis durch die Einführung von Designprinzipien, Mustern, konzeptuellen Modellen und der Ontologie selbst. Zukünftige Arbeiten sollten die Ontologie weiter verfeinern, indem sie zusätzliche

Technologien integrieren, sie in verschiedenen Domänen anwenden und Prototypen darauf basierend erstellen. In zukünftigen Arbeiten könnte die Ontologie kreatives Problemlösen fördern, indem sie bestehende Lösungen durch konzeptuelle Ähnlichkeiten mit verschiedenen Geschäftsopportunitäten verbindet. Dieser Ansatz beinhaltet das Identifizieren, Abstrahieren und Anwenden grundlegender Ideen oder Prinzipien, ähnlich jenen, die hinter Rasierklingen oder natürlichen Mechanismen stehen, in diversen Domänen.

List of Publications

The content of this thesis builds on the following **peer-reviewed publications**:

1. Burkhardt, D. 2015. “Analyse und Bewertung von Architekturvarianten mithilfe von Enterprise Architecture Management-Methoden dargestellt am Beispiel der Service Brokering Plattform im Projekt Stuttgart Services“, Master Thesis at Institute of Business Administration, Chair of Information System Department VII.
2. Burkhardt, D., Werling, M., and Lasi, H. 2018. “Distributed Ledger Definition & Demarcation”, 2018 IEEE International Conference on Engineering, Technology and Innovation, ICE/ITMC 2018 – Proceedings. https://doi.org/10.1109/ICE.2018.8436299.
3. Burkhardt, D., Frey, P., and Lasi, H. 2019. “The Symbiosis of Distributed Ledger and Machine Learning as a Relevance for Autonomy in the Internet of Things”, Proceedings of the 52nd Hawaii International Conference on System Sciences HICSS. https://doi.org/10.24251/hicss.2019.559.
4. Burkhardt, D., and Lasi, H. 2020. “A Conceptual Model of Data-Driven Solutions”, in Americas Conference on Information Systems.
5. Burkhardt, D., Agyei-Kena, N., Frey, P., Kurrle, S., and Lasi, H. 2020. “Design Patterns Based on Deep Learning Analyzing Distributed Data”, in Wirtschaftsinformatik.
6. Burkhardt, D., Enders, N., and Lasi, H. 2021. “A Model Design to Evaluate Processes for Autonomization at the Example of a Transportation Case”, in 25th Asia Conference on Information Systems PACIS.

Non-refereed scientific publications:

7. Burkhardt, D., Frey, P., Hiller, S., and Neff, A. 2019. “Distributed Ledger Enabled IoT Platforms Symbiosis Evaluation”, in Business Transformation through Blockchain Volume II, Springer International Publishing.

Other publications and accompanying material of the author for transfer and presentation of intermediate results:

Articles:

1. Burkhardt, D. 2017. “Blockchain – Facts behind the Hype”, Steinbeis Transfer, Edition 4.

2. Burkhardt, D. 2020. “Autonomisierung: Realität Oder (Noch) Fiktion”, Steinbeis Transfer, Edition 1.
3. Burkhardt, D. 2020. “Kalliope: Digitale Unterstützung Der Pflege Im Ländlichen Raum”, Steinbeis Transfer, Edition 2.
4. Burkhardt, D. 2023. “Kooperation am AIoTLab Lecksuche mit KI und IoT”, IT&Produktion Online, https://www.it-production.com/allgemein/kooperation-am-aiotlab/

Supervised theses and common publications:

1. Schwende, I. 2019. “Die Analyse der Fähigkeiten von Blockchain in landwirtschaftlichen Supply Chains”, Bachelorthesis.
2. Busch, R. 2019. “Konzeptionelles Prozessmanagement entlang der Additive Manufacturing-Wertschöpfungskette basierend auf Blockchain Technologie“, Bachelorthesis.
3. Wittkämper, L. 2020. “Entwicklung eines Leitfadens zur Prozessautomatisierung unter Einsatz Künstlicher Intelligenz”, Steinbeis School of Management and Innovation, Steinbeis-Hochschule, Masterthesis.
4. Scherl, T. 2020. “Einsatz Künstlicher Neuronaler Netze zur automatischen Nummernschilderkennung”, T-Systems, Steinbeis-Hochschule, Masterthesis.
5. Lebioda, A., Lachenmaier, J. F., and Burkhardt, D. 2019. “Control of Cyber-Physical Production Systems: A Concept to Increase the Trustworthiness within Multi-Agent Systems with Distributed Ledger Technology”, in Proceedings of Pacific Asia Conference on Information Systems PACIS, common publication.
6. Frey, P. 2019. “Design and Prototypical Evaluation of a Blockchain-Based System for the Storage of Electronic Medical Records,” Karlsruhe Institute of Technology (KIT), Masterthesis.
7. Enders, N. 2020. “Wie lässt sich die Bewertung von Autonomen Geschäftsprozessen am Beispiel der DB Netz AG gestalten?”, DB Netz AG, Steinbeis-Hochschule, Masterthesis.
8. Deshpande, P. 2021. “Design and Implementation of Deep Learning Models for Analyzing German Mobility Data,” Hochschule Kaiserslautern, Master thesis.
9. Renken, S. 2022. “Maßnahmen, um Vertrauen in kooperativen Datenräumen herzustellen“, Steinbeis Hochschule, Bachelorthesis.

Lectures and workshops:

1. Burkhardt, D. 2019–2021. "Data-driven Solution Design based on Machine Learning and Distributed Ledger Technology for Opportunities of Autonomization", Steinbeis-Hochschule, Berlin, Hohenheim, 12 hours each, Master studies, 20 to 35 participants.
2. Burkhardt, D. 2019–2021. "Distributed Ledger Blockchain and IoT", Steinbeis-Hochschule, Berlin, Hohenheim, 6 hours each, Master and Bachelor studies, 20 to 35 participants.
3. Burkhardt, D. 2019–2021. "Autonomization and Digitalization – a Technology Mix for Solution Design", Steinbeis-Hochschule, Berlin, Hohenheim, 5 hours each, Master studies, 20 to 35 participants.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

Table-Abbreviations (it is recommended to put this list next to the document while reading. The chapter numbers indicate where an abbreviation occurs)

Chapter	Abbreviation	Term
1	ADM	Architecture Development Method
	AI	Artificial Intelligence
	AIS	Analytical Information System
	AM	Architecture Modeling
	ArchiMate	is an open enterprise architecture modeling language
	ASUM-DM	Analytics Solutions Unified Method for Data Mining / Predictive Analytic
	CRISP-DM	Cross-Industry Standard Process for Data Mining
	DAMA DMBoK	Data Management Association Data Management Body of Knowledge
	DL	Deep Learning
	DLT	Distributed Ledger Technology
	DoDAF	Department of Defense Architecture Framework
	DT	Digital Transformation
	EA	Enterprise Architecture
	EAM	Enterprise Architecture Management
	EM	Enterprise Modeling
	GAIA-X	based on the European values of transparency, openness, data protection, and security. It relates to the Greek goddess Gaia.
	IDS-RAM	International Data Space Reference Architect Model
	IIC	Industry IoT Consortium
	IIRA	Industrial Internet Reference Architecture
	IoT	Internet of Things
	IS	Information Systems
	IT	Internet Technology
	KDD	Knowledge Discovery in Databases
	KM	Knowledge Management
	ML	Machine Learning
	OWL	Web Ontology Language
	RAMI 4.0	Reference Architectural Model Industrie 4.0
	RDF	Resource Description Framework
	RQ	Research Question
	SEMMA	Sample, Explore, Modify, Model, and Assess
	TOGAF	The Open Group Architecture Framework
2	App.	Application
	AutoML	Automated Machine Learning
	BCT	Blockchain Technology

	CNN	Convolutional Neural Network
	Comp.	Component
	Dev.	Development
	EVM	Ethereum Virtual Machine
	FN	False Negative
	FP	False Positive
	GAI	General Artificial Intelligence
	GLM	Generalized Linear Model
	GPT-3	Generative Pre-trained Transformer
	GPU	Graphics Processing Unit
	HPO	Hyperparameter Optimization
	IPFS	Interplanetary File System
	KG	Knowledge Graph
	KMS	Knowledge Management System
	MLP	Multi Layered Perceptron
	MVC	Model View Control
	NAI	Narrow Artificial Intelligence
	NLP	Natural Language Processing
	NN	Neural Network
	OMG	Object Management Group
	p2p	Peer-to-peer
	PCA	Principal Component Analysis
	PIN	Pattern Instance Notation
	PoS	Proof of Stake
	PoW	Proof of Work
	RBM	Restricted Boltzmann Machine
	RNN	Recursive Neural Network
	RPA	Remote Procedure Call
	RSK	Rootstock Platform
	SARSA	State Action Reward State Action
	SPARQL	SPARQL Protocol and RDF Query Language
	SVM	Support Vector Machine
	TCP/IP	Transmission Control Protocol / Internet Protocol
	TN	True Negative
	TP	True Positive
	TPU	Tensor Processing Unit
	URL	Uniform Resource Locator
	cf.	confer (latin), meaning compare
	CS	Computer Science
3	DP	Design Principle
	DSR	Design Science Research
	fDP	functional Design Principle

	fRQ	functional Requirement
	gfDP	generic functional Design Principle
	gfRQ	generic functional Requirement
	KB	Knowledge Base
	MFCC	Mel-frequency cepstral coefficient
	MoP	Mobility Panel
	nfRQ	non-functional Requirement
	sDP	system Design Principle
	sRQ	system Requirement
	UML	Unified Modeling Language
4	AGI	Artificial General Intelligence
	AP	Architectural Pattern
	CNN	Convolutional Neural Network
	CPU	Central Processig Unit
	DAG	Directed Acyclic Graph
	DAU_SUM	average time spent on mobility by an individual per day, in minutes
	DHT	Distributed Hash Table
	E	Expert
	EAI	Enterprise Application Integration
	EHR	Electronic Health Record
	GDPR	General Data Protection Regulation
	IaaS	Infrastructure as a Service
	iOS	iPhone Operating System
	IPNS	Inter Planetary Name System
	JSON	JavaScript Object Notation
	PoC	Proof of Concept
	KIT-IfV	Karlsruhe Institute for Technology – Institute for Trannsportation
	KL	Kulback-Leibler
	LSTM	Long Short-Term Memory
	MCTS	Monte Carlo Tree Search
	MOP	German Mobility Panel
	MR	Meta-Requirements
	Neo4J	is a graph data platform
	NN	Neural Network
	NutzPKW	indicates use of a car by a specific person on a particular day of the week
	PaaS	Platform as a Service
	RL	Reinforcement Learning
	SaaS	Software as a Service
	SL	Supervised Learning

	SOA	Service Oriented Architecture
	STASA	Steinbeis Angewandte Systemanalyse GmbH
	W3C	World Wide Web Consortium
	XML	Extensible Markup Language
5	API	Application Programming Interface
	AWS	Amazon Web Service
	CQ	Competency Questions
	EVAL	Evaluation
	FAIR	Findability, Accessibility, Interoperability, and Reusability
	GraphQL	query language for an API
	JinaAI	Machine Learning operating platform for multimodal AI
	LeanIX	enterprise architecture management tool
	MAXQDA	tool for conducting many methods of analysis
	n10s	Neosemantics
	NER	Natural Language Understanding
	NeuSpell	Neural Spelling Correction Toolkit
	NLP	Natural Language Processing
	NLU	Natural Language Understanding
	OoV	Out-of-Vocabulary
	OpenCyc	gateway to Cyc, a general knowledge base and commonsense reasoning engine
	ROI	Return of Investments
	RS	Recommender System
	SpaCy	a free open-source library for Natural Language Processing
	STACKIT	cloud computing solution
	tf-idf	term frequency-inverse document frequency

1 Introduction to Data-driven Solution Design in Complex Systems

Organizations face an enhanced complexity[1] of their environment and feel increased pressure to innovate digitally (Kohli and Melville 2018, pp. 1-2; Yoo et al. 2010, pp. 2-3; Zapadka 2020, p. 2). Additionally, stakeholders expect organizations to apply recent technologies to create novel **data-driven solutions**[2] that satisfy needs ideally on an individual level. The evolving environmental pressure forces organizations to collaborate in networks, adapt their business model and organizational structure, as well as change the design and development of solutions accordingly (Henfridsson 2009, pp. 4-5; Hylving and Schultze 2013, pp. 2-3; Hylving and Selander 2012, pp. 2-3; Zapadka 2020, p. 3). Such a digital transformation[3] raises multi-faceted challenges that create complexity (Svahn et al. 2017, pp. 239-240; Yoo et al. 2010, pp. 224-227). Consequently, organizations must translate identified business opportunities into data-driven solutions in increasingly shorter cycles. Businesses require **transparency**[4] about the available solution space and **design knowledge**[5] about applied solutions to guarantee value-add. By this, organizations establish speed, judgment, and flexibility, which are vital in a dynamic and competitive environment (Davenport and Prusak 1998, pp. 4-8; Kohli and Melville 2018, pp. 2-3; Zapadka 2020, p. 4).

---

[1] *Complexity* describes a state that contains non-linear relationships and feedback loops so that the system in its totality cannot be analyzed. Thus, complex systems are not only constituted by their components but also by the relationships between these components (Cilliers 1998, pp. 2–3; Dacrema et al. 2019, p. 8).

[2] Combining 'data-driven' with 'solution' to '*data-driven solutions*' describes a solution that uses data and technological capabilities based on a system to realize business-related opportunities by solving identified challenges.

[3] *Digital transformation* defines the organizational change triggered by digital technology. For example, the emergence of social media led to new digital solutions, which initiated the development of new digital business concepts, as illustrated by social media marketing concepts. Iteratively, the new digital business concept can trigger new digital solutions (e.g., big data analytics), which lead to the creation of complementing digital business concepts (Wiesböck and Hess 2020, p. 77).

[4] *Words are highlighted* in this thesis in order to emphasize their conceptual importance in the respective paragraph.

[5] *Design and (Enterprise or Architecture) Modeling:* As the complexity of systems and technologies increases, the pre-implementation phase of modeling becomes more critical. It is essential for understanding and communicating knowledge as an information asset (Stirna and Persson 2018, p. V). Modeling results in the definition of a specific model abstracting from requirements and the actual situation. See also Section 2.2.2.

Series of the Ferdinand Steinbeis Institute on Digital Business Transformation | Volume 2

In the immersive complexity of digital innovation, data-driven solution design often begins with a specific technology or potential solution already in mind. However, this predisposition narrows the design space and reduces the probability of discovering more suitable or novel alternatives. This thesis introduces a formally developed ontology to support *data-driven solution design*, offering a meta-structure that facilitates abstraction and translation from business opportunities to technological concepts. Developed through extensive literature reviews and four empirical case studies, the ontology was constructed using both deductive and inductive methods. It uniquely integrates the layers of *Business*, *System*, *Data*, and *Technology*, enabling structured reasoning and navigation across abstraction levels. The hierarchical ontology serves as the foundation for a *multi-agent large language model system (MALLMs)* that assists users in identifying appropriate solution concepts aligned with provided business opportunities. Central to this system is a *Knowledge Graph (KG)* acting as a meta-agent, providing structural guidance and validation. Additional agents generate solution-relevant information through *retrieval-augmented generation (RAG)* based on a curated corpus of over 14,000 research papers, academic articles, and code fragments related to data-driven solution design. The findings suggest that combining symbolic representations with neural models enhances output effectiveness by leveraging both user and domain-specific knowledge.

Dr. Daniel Burkhardt is a researcher in the field of neuro-symbolic AI, specializing in the integration of knowledge graphs and language models. His work focuses on adaptive, explainable systems for generation of effective reasoning chains. He co-authored DIN SPEC 91526 and leads research on enhancing reasoning and explainability of large language models using symbolic structures.

ISBN 978-3-95663-315-7

Steinbeis-Edition