Weinhold, Maximilian

The Governance of Decentralized Innovation Ecosystems

Dissertation
for obtaining the degree of Doctor of Business and Economics
(Doctor rerum politicarum - Dr. rer. pol.)

at WHU – Otto Beisheim School of Management

24.06.2024

First Advisor: *Prof. Dr. Dries Faems*

Second Advisor: *Prof. Dr. Stefan Spinler*

I am grateful to Prof. Dr. Dries Faems for his guidance, support, and expertise, which were fundamental in conducting and completing this research. I also thank Prof. Dr. Stefan Spinler for serving as my second adviser and for the support he has provided in this endeavor.

My sincere thanks go to my beloved family for their endless support and encouragement.

TABLE OF CONTENTS

CHAPTER ONE: GENERAL INTRODUCTION

1.1 INTRODUCTION

Innovation ecosystems as a form of business ecosystems have become an increasingly important research topic in the discussion of business strategies across organizations, providing a perspective that complements the established schools of thought on competition and cooperation (Adner, 2016; Autio, 2022; Thomas & Autio, 2019). The transformation from the Schumpeterian view of innovation centered around individual firms to a more holistic view where innovation is embedded in ecosystems of complementary actors is a significant shift in the field of innovation studies (Baldwin et al., 2024)

While business ecosystems are defined as multilateral constellations of partners that jointly aim to develop a value proposition (Adner, 2012), innovation ecosystems are characterized by "collaborative arrangements through which firms combine their individual offerings into a coherent, customer-facing solution" (Adner, 2006, p. 2). Ecosystem governance has been identified as a core driver of ecosystem success (Daymond et al., 2022; Jacobides et al., 2018; Pidun et al., 2020; Stonig & Müller-Stewens, 2019). This governance involves the rules and policies (Adner, 2016; Jacobides et al., 2018), but also the dynamic interplay between diverse and autonomous actors (Baldwin et al., 2024).

While there is broad agreement on the significance of centralized leadership for fostering ecosystem success and stability (Adner & Kapoor, 2010; Daymond et al., 2022), empirical research of traditional governance models being challenged by ecosystem participants remains underexplored in management research (Chen, Richter, et al., 2020; Cozzolino & Geiger, 2024; Pujadas et al., 2024). My research interest in decentralized governance strategies of innovation ecosystems corresponds to this gap in the existing literature. When I engaged in case studies of innovation ecosystems, I observed how innovation ecosystem initiators planned to embrace the governance principles of centralized orchestration for an industry-wide sustainability initiative. Yet, I also observed how these governance principles were fundamentally rejected by potential ecosystem partners. Exposing the limitations of centralized governance models, this resistance highlights room for a shift toward more alternate governance models and a paradigm that embraces decentralization. Such a paradigm requires governance mechanisms that are suitable and adaptable to the complex interplay between diverse stakeholders within innovation ecosystems, a need that is increasingly recognized in the literature (i.e. Baldwin et al., 2024)

In this context, blockchain technology has been posited in management research as one of the core mechanisms enabling a decentralized governance (i.e. Chen, Pereira, et al., 2020; Du et al., 2023). The technology has been discussed as a potential alternative to established interfirm governance structures (Goldsby & Hanisch, 2023; Smits & Hulstijn, 2020), "distinct from traditional contractual and relational governance" (Lumineau, Wang, et al., 2020, p. 3). However, as research also acknowledges limited understanding of how blockchain integrates with governance frameworks in innovation ecosystems in practicality (i.e. Gan et al., 2023; Lei & Ngai, 2023), it opens the door for further exploration. My research on case studies contributes to this ongoing discourse by positioning blockchain not merely as an alternative but rather as a complementary approach within these governance structures.

The objective of this doctoral thesis, therefore, is to explore governance within decentralized innovation ecosystems from different angles. It will study how governance approaches and mechanisms, including blockchain technology, contribute to the organizing and management of these ecosystems, aiming to identify best practices for fostering successful collaborative innovation.

My consecutive thesis starts with an in-depth NLP-based literature review on the topic of innovation ecosystems alongside the discussion of the feasibility of this enhanced approach for wider knowledge management in management research. This is followed by a study on the challenges and characteristics of ecosystem governance in balancing centralization and decentralization in an industry-wide sustainability initiative. Finally, I explore the managerial challenges and intricacies surrounding the scalability of blockchain applications in the setting of intrafirm innovation ecosystems.

1.2. APPROACH AND MOTIVATION OF THE RESEARCH PROJECTS

The first paper of this cumulative study initially strived to provide a literature review on the topic "Innovation Ecosystems". However, traditional reviews are limited in their effectiveness and efficiency when dealing with the vast bodies of literature (Bawden & Robinson, 2009; Edmunds & Morris, 2000), as seen with our topic, producing several hundred publications per year. Recent holistic literature reviews covering the broader theme of innovation ecosystem remained rare. Notably, Granstrand and Holgersson (2020) did present a conceptual overview of "innovation ecosystems". Their manual method, however, allowed them to analyze only 100 publications from the period 2016 to 2020. This limitation prompts a need for alternative methods.

Topic modeling, especially the latent Dirichlet allocation (LDA) introduced by Blei, Ng, and Jordan (2003), has emerged as an acknowledged text-mining tool for machine-based uncovering of hidden semantic structures in texts. Although this technique has allowed researchers to identify topic patterns in a wide variety of research contexts, it also faces important limitations in contextual interpretation. To address the limitations of LDA for topic modeling, Natural Language Processing (NLP) scholars have recently started developing alternative approaches such as BERTopic (Grootendorst, 2022a) and CorEx (R. Gallagher, Reing, Kale, & Ver Steeg, 2016). Moreover, the launch of large language models such as GPT provides novel options for identifying, categorizing, and labeling topics. To our knowledge, the application of these methods has remained largely absent in management research.

My paper aims to develop an integrative topic modeling approach that enhances knowledge accumulation in entrepreneurship and innovation research. It seeks to address the shortcomings of traditional methods like LDA by incorporating advanced transformer-based approaches, thus offering a more comprehensive and contextually aware analysis.

Using 2567 journal paper abstracts in the research domain of innovation ecosystems, the paper ultimately identifies and analyzes 27 distinct topics, enabling a comprehensive bibliographic review. This approach showcases the potential of a combination of various topic modeling techniques for scalable, contextual text analysis in management research.

Segueing into the exploration of sustainability-centered innovation ecosystems, **the second paper** of this thesis presents insights into various governance approaches. Existing ecosystem governance literature has emphasized the advantages of centralized orchestration and a minimal viable ecosystem configuration ecosystem (Adner & Kapoor, 2010; Daymond et al., 2022). As we engaged in a longitudinal case study of the development of an innovation ecosystem to facilitate data exchange on carbon emissions, we observed that this kind of "traditional" governance design led to initial ecosystem rejection. Triggered by this surprising observation, we engaged in an inductive analysis of governance dynamics in sustainability-centered ecosystems. This research is particularly relevant at a time when environmental concerns are paramount, and the effective governance of sustainable ecosystems is crucial for long-term ecological health (Aarikka-Stenroos et al., 2021; Papalexopoulos & Klein, 2022; Peltola et al., 2016).

Our findings show how both fully centralized and fully decentralized governance approaches triggered important challenges that hampered the growth of the ecosystem. At the same time, we identify two important ecosystem governance principles – i.e., democratic, multi-body governance, and give-first mentality – which can help to address these challenges and navigate the tension between centralization and decentralization in sustainability-centered ecosystems.

Lastly, **in the third paper**, the thesis brings to light the managerial challenges and intricacies surrounding the scalability of blockchain applications in the setting of intrafirm innovation ecosystems. Research on innovation ecosystems has pointed at blockchain technology as one viable mechanism to govern innovation ecosystems. For instance, Du et al. (2023, p. 29) mention in a recent article in the California Management Review that "In general, blockchain helps address pain points at the ecosystem level". Lumineau, Wang, et al. (2020, p. 3) note that "blockchains may be considered a critical turning point in organizing collaborations and can be viewed as a type of governance mechanism that is distinct from traditional contractual and relational governance".

The motivation for this paper is driven by the need to understand and overcome the organizational challenges that impede the scalability and practical implementation of blockchain. The research aims to demystify the complexities of adopting blockchain in large-scale industrial settings, addressing the gap in knowledge regarding how such technology can be integrated into traditional organizational frameworks (i.e. Gan et al., 2023; Lei & Ngai, 2023). Based on an in-depth analysis of two projects in a large industrial corporation, we inductively derive three organizational challenges, explaining why blockchain applications struggle to scale. Our findings contribute to the growing management literature on blockchain in different ways. First, whereas current management literature has positioned blockchain as a substitute for existing interfirm governance structures and as a third-party replacement, our study suggests that blockchain is rather a complementing governance structure. In particular, our findings highlight that traditional collaborative governance frameworks are required before blockchain-based governance can fully realize its promise. Second, as opposed to extant assumptions that blockchain can be seen as a trust-enabling silver bullet, we find that the successful implementation of blockchain actually relies on preexisting trust, provided by sound governance structures.

The following table gives an overview of each of the three papers on the addressed research gap, the research object, the employed methodology, the core finding as well as the core contribution.

Figure 1.1. Research Overview

Paper	**Research Gap**	**Research Object**	**Methodology**	**Core Finding**	**Core Contribution**
Integrated Topic Modeling	Limitations of traditional topic modeling methods in handling large-scale, contextual analysis in management research	2567 journal paper abstracts in the field of innovation ecosystems	Novel approach to a Bibliometric Analysis: Combination of BERTopic, CorEx, and GPT models for topic modeling	Demonstration of the potential of combining various NLP techniques for effective and scalable text analysis in management research	Integrated approach enhancing scalability and contextual relevance of topic analysis in mgmt. research & conducted Bibliographic Literature Review
Governance in Sustainability Ecosystems	The challenges of (rejected) ecosystem governance in balancing centralized and decentralized governance	Exploration of the development of an industrial interfirm sustainability ecosystem for carbon emissions data exchange	Longitudinal case study exploring the balance between centralization and decentralization in ecosystem governance	Democratic multi-body governance and a give-first mentality help address challenges in sustainability ecosystems	Provides insights into managing the tension between centralization and decentralization in environmental sustainability initiatives
Blockchain in Industrial Organizations	Understanding how blockchain technology can be scaled and integrated within existing organizational structures	Two intrafirm governance blockchain projects (NewProtect and NewCert) from a large industrial corporation	Case studies of two blockchain projects exploring organizational challenges	Blockchain relies on preexisting trust and governance structures, not a substitute for existing structures	Contributes to understanding role of blockchain as a complementary governance structure in organizational contexts

CHAPTER TWO:

DEVELOPING AN INTEGRATED TOPIC MODELING APPROACH FOR KNOWLEDGE ACCUMULATION IN ENTREPRENEURSHIP AND INNOVATION

ABSTRACT

This paper introduces an integrative topic modeling approach, leveraging recent advancements in natural language processing (NLP) for knowledge accumulation in entrepreneurship and innovation research. It addresses the limitations of traditional topic modeling methods (i.e., Latent Dirichlet Allocation), by employing novel algorithms such as BERTopic, CorEx, and leveraging large language models like GPT. Using 2567 journal paper abstracts in the research domain of innovation ecosystems, the paper identifies and analyzes 27 distinct topics, enabling a comprehensive bibliographic review. This innovative approach showcases the potential of a combination of various topic modeling techniques for scalable, contextual text analysis in management research.

An adjusted version of this chapter is accepted for presentation at the *Academy of Management Conference 2024* Weinhold, M., & Faems, D. (2024). Developing an Integrated Topic Modeling Approach for Knowledge Accumulation. *Academy of Management Conference 2024*. Developing an Integrated Topic Modeling Approach for Knowledge Accumulation

2.1. INTRO

Natural language processing (NLP) has been recognized as a promising machine learning domain to leverage textual data for knowledge accumulation in management research.

One particular approach that has received substantial attention is topic modeling (for an extensive review, see Hannigan et al. (2019). Topic modeling is a text-mining tool for the discovery of hidden semantic structures in a text body. The conventional method for topic modeling is latent Dirichlet allocation (LDA), which was introduced by Blei et al. (2003). Although this technique has allowed researchers to identify and use topics in a wide variety of research contexts, it also has important limitations.

Those machine-driven, unsupervised models, while impressive in their computational abilities, often falter when it comes to contextual interpretation. Lacoste-Julien et al. (2009) pointed out the potential for such models to neglect the nuanced context of documents. Such an oversight can accidentally push aside vital thematic details, leading to an incomplete or even skewed representation of the literature. Further compounding this issue is the observation that traditionally machine-generated topics sometimes fail to align with human comprehension and understanding of the context. As Chang et al. (2009) noted, a topic that is mathematically coherent doesn't necessarily ensure its semantic or thematic relevance to human researchers. In other words, it is often difficult to adequately interpret the topics that emerge from the algorithm.

To address the limitations of LDA for topic modeling, NLP scholars have recently started developing alternative approaches such as BERTopic (Grootendorst, 2022b) and CorEx (Gallagher et al., 2016). Moreover, the launch of large language models such as GPT provides novel options for identifying, categorizing, and labeling topics.

To our knowledge, the application of these methods has remained largely absent in management research. Closing this gap, however, is important as these alternative topic modeling algorithms have the potential to assist scholars in broadening their portfolio of tools to engage in topic modeling in specific literature domains such as entrepreneurship and innovation.

The core objective of this paper is to develop and demonstrate an integrative topic-modeling approach that leverages the advantages of different novel topic-modeling techniques for knowledge accumulation in entrepreneurship and innovation. To do so, we first provide a brief description of the underlying logic of these different topic modeling approaches. Subsequently, we provide a step-based explanation and illustration of our approach. We rely on 2567 journal paper abstracts in the field of innovation ecosystems. Thereby, we demonstrate how our novel topic modeling approach allows for the identification and analysis of a coherent set of 27 topics, capturing a wide variety of research streams within the field of innovation ecosystems. We also show how these 27 topics can be subsequently used to conduct a bibliographic review of the innovation ecosystem literature. Our core contribution is that we demonstrate the opportunity to leverage and combine novel topic model algorithms to develop a scalable approach toward text-based knowledge accumulation.

2.2. LITERATURE REVIEW

2.2.1. Traditional Topic Modeling: LDA

The foundational framework for topic modeling was established by Blei et al. (2003) with the introduction of Latent Dirichlet Allocation (LDA). By evaluating the distribution of word occurrences across a given corpus of text, this statistical model identified latent structures within large datasets, discerning discrete topics. LDA is classified as an unsupervised topic modeling technique. Hence, it identifies previously unknown patterns in a dataset without pre-existing labels and with minimal human oversight (Hastie et al., 2009). Within its collection of documents, also referred to as a corpus, this method aims to uncover abstract topics. Thus, LDA operates under the premise that words convey significant semantic information and that documents on similar topics will use similar word groups. Essentially, the model assumes that each document contains a distribution of topics, and each topic is represented by a distribution of words (Derstappen, 2021).

LDA represents text by noting the occurrences of words within a document while ignoring the order and structure of those words. Therefore, it is also referred to as a "bag-of-words" model. It does not account for relationships between words. LDA extends beyond a simple word count by incorporating term frequency-inverse document frequency (TF-IDF), which normalizes word counts and assesses the importance of terms by dividing the total number of documents by the number of documents that contain each term (Derstappen, 2021; Liu et al., 2015).

Over time, numerous LDA extensions have been introduced, such as the Author-Topic models (Steyvers et al., 2004), Dynamic Topic Models (Blei & Lafferty, 2006), and supervised LDA topic models (Lacoste-Julien et al., 2009). Despite these extensions, several issues remain present when applying LDA approaches. For instance, the topics which emerge out of LDA-based algorithms often are difficult to interpret from a human perspective (Chang et al. (2009). Moreover, these topic models often struggle with domain-specific language, acronyms, and jargon, which are common in specialized fields like entrepreneurship and innovation (Greene et al., 2014). Further compounding this issue is the observation that while machine-generated topics might indeed emerge from robust mathematical foundations, they sometimes fail to align with human comprehension and understanding of the context. As Chang et al. (2009) noted, a topic that is mathematically coherent doesn't necessarily ensure its semantic or thematic relevance to human researchers. In other words, it is often difficult to adequately interpret the topics which emerge from the algorithm. Moreover, LDA triggers a substantial risk of undercooking and overcooking, which limits the usability of the topics for subsequent analyses. As LDA relies on a bag-of-words approach, the order or structure of the words is neglected, implying that LDA does not consider the context in which particular words are used. Finally, LDA topics often get contaminated by words that are not really representative for the underlying topic, jeopardizing the reliability of topic-based classifications (van Kessel, 2019).

Application of Topic Modeling. Topic Modeling has seen wide-ranging applications across numerous disciplines. For instance, Ni Ki et al. (2022) tapped into Topic Modeling to refine precision medicine in personalized diabetes management by analyzing diabetes literature and genomic data. Similarly, Rahman et al. (2020) sought insights into hydrology literature trends. The adaptability of this technique is further showcased by Ricketts et al. (2023), who extracted safety challenges from occurrence reports, Yasin Musa (2023), who forecasted Zambian election results, and Leonnardo Benjamin and Derwin (2022), who unearthed hoax news in Indonesia. Even discussions as contemporary as COVID-19 vaccination opposition found clarity through this technique, as illustrated by Andre Mediote de and Karin (2021) on Twitter. Another compelling, recent application was presented by Mallick et al. (2023) who analyzed the impacts of climate change on critical infrastructure.

2.2.2. Novel Topic Modeling Approaches

Introduction to Transformer-based techniques. The emergence of transformer-based techniques in machine learning has revolutionized the field of NLP. The foundational paper by Vaswani et al. (2017), introduced the transformer architecture and demonstrated that “self-attention mechanisms” could significantly outperform previous recurrent neural network models in translation tasks. The transformer's ability to process sequences of data in parallel, rather than sequentially, allows for more efficient training and improved performance on complex tasks (Vaswani et al., 2017).

Subsequent studies have built upon this architecture, leading to the development of models like BERT (Devlin et al., 2019), which uses bidirectional training to achieve a deeper understanding of language context, and GPT (Radford & Narasimhan, 2018). The advancements have led to significant improvements in analyzing and understanding textual data. Thus, through leveraging substantial computational power, these models identify patterns and themes in large-scale text data.

BERTopic is an unsupervised topic modeling approach that leverages the BERT transformer model (Grootendorst, 2022b) By doing so, this approach can dynamically determine the number of topics, thereby avoiding the need for manual adjustments and hyperparameter tuning typically associated with LDA topic modeling. BERTopic identifies the most significant words within content clusters, known as topic keywords, to describe the essence of each topic. This process results in coherent and interpretable topics, offering a substantial improvement over conventional topic modeling techniques like LDA (Grootendorst, 2022b). Technically, the algorithm works by first creating a document-term matrix using a class-based variation of TF-IDF. Next, it uses BERT transformers to create dense clusters of similar corpus documents and assign them to topics. To do so, HDBSCAN, a density-based clustering algorithm well-suited for high-dimensional data, is applied. The resulting clusters are then used to identify the most representative words for each topic, which are called topic keywords (Devlin et al., 2019).

CorEx. The Correlation Explanation (CorEx) model is a semi-supervised approach designed to identify "maximally informed topics" within a textual corpus. It operates on the principles of total correlation and mutual information. In contrast to LDA, it does not require hyperparameter tuning (Gallagher et al., 2017). CorEx is particularly relevant for its interactive anchoring system. It allows the integration of user-defined “anchor words” to direct the model's focus toward specific themes, enhancing mutual information between the corpus and these topics (Gallagher et al., 2017; Vincent, 2021).

This semi-supervised learning capability has advantages when certain terms or subjects occur infrequently in the corpus. In contrast to LDA's probabilistic approach, CorEx assigns topics to documents in a binary fashion, leading to more decisive categorization. Moreover, the model allows the user to adjust the anchor strength, modulating the influence of anchor words on topic generation. This level of intervention and control over thematic emphasis is not possible with unsupervised algorithms.

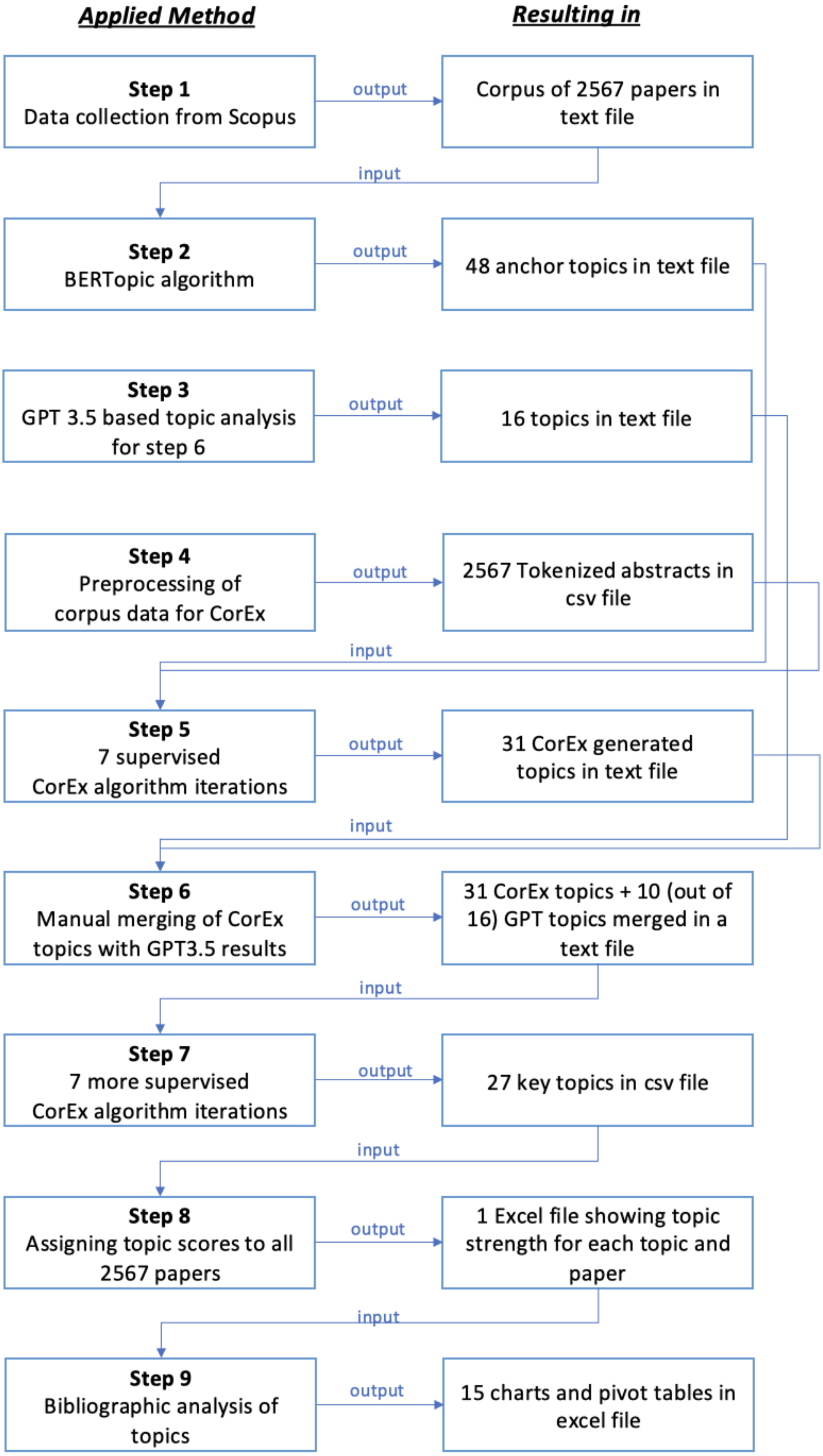
GPT-based Topic Modeling. GPT, or Generative Pre-trained Transformer, is a large language model developed by OpenAI that uses a transformer neural network for deep learning. It works by analyzing patterns and relationships in the text data it has been trained on, using unsupervised learning techniques. It generates responses based on probabilities, predicting the most likely next word or phrase in a sequence, thereby crafting coherent and contextually appropriate text. It learns from vast internet text datasets, generating human-like text (Radford & Narasimhan, 2018). GPT 3.5, a sophisticated iteration, is notable for its learning ability (called "few short learning") to adapt to various language tasks without specific training (Brown et al., 2020).

In this paper, we combine the capabilities of llama_index, an indexing system, with GPT 3.5 to conduct an additional iteration of unsupervised topic identification. The combination of llama_index and GPT 3.5 allows for Retrieval Augmented Generation (RAG). The retrieval component is responsible for fetching relevant information from a collection of documents. When a prompt is received, this retrieval system searches the documents to find the most relevant pieces of information. The generation component will use the retrieved information to generate a response. At its core, the method transforms a large set of textual data - in our case, the corpus of abstracts into vector representations and indexes those vectors for efficient storage and prompting. Subsequently, we can store this information in a vector database, which can be subsequently prompted with models like GPT 3.5.

2.3. METHODOLOGY

In the following, we explain the integrated topic modeling approach we developed to leverage the benefits of novel natural language processing breakthroughs fully. To illustrate this approach, we have selected 2567 paper abstracts on the topic of innovation ecosystems.

Figure 2.1. Steps in the Integrative Topic Modeling Approach



With a cut-off data of 23.05.23, as **a first step,** we have applied the following filter to Scopus to limit publications from the year 2015 (01.01.2015) to the year 2022 (31.12.2022) in business and management publications. We chose to start in 2015 as it was a significant year when the number of publications on innovation ecosystems surpassed 100, indicating a notable increase in research and interest in this field.

> TITLE-ABS-KEY (INNOVATION AND ECOSYSTEMS) AND LIMIT-TO (PUBYEAR , 2022) OR LIMIT-TO (PUBYEAR , 2021) OR LIMIT-TO (PUBYEAR , 2020) OR LIMIT-TO (PUBYEAR , 2019) OR LIMIT-TO (PUBYEAR , 2018) OR LIMIT-TO (PUBYEAR , 2017) OR LIMIT-TO (PUBYEAR , 2016) OR LIMIT-TO (PUBYEAR , 2015)) AND (LIMIT-TO (DOCTYPE , "AR") OR LIMIT-TO (DOCTYPE , "CP") OR LIMIT-TO (DOCTYPE , "CH")) AND (LIMIT-TO (SUBJAREA , "BUSI"))

Using the exporting functionality of Scopus, we downloaded the abstracts into a single csv file for further exploration. With 2567 papers, the dataset likely captures a wide spectrum of research — from foundational studies to papers that delve into the latest advancements and challenges.

In the **second step**, we started analyzing the corpus of abstracts using BERTopic in a python script that we ran in Jupyter Notebook. With BERTopic being an unsupervised neural topic modeling algorithm, it does not require any user-specified guidance to execute the topic modeling process (Leonnardo Benjamin & Derwin, 2022).

After several trials, we found to receive the most diverse and clear set of topics with the following configuration:

```
representation_model = MaximalMarginalRelevance(diversity=0.5)
topic_model = BERTopic(
  representation_model=representation_model,
  n_gram_range = (1,2),
  min_topic_size = 10,
  top_n_words = 10,
  verbose = True,
```

It is not advised to pre-process data before feeding it into BERTopic due to its native ability to handle raw textual data (Grootendorst, 2022a). To get more precise topics, however, we removed the words “innovation” and “ecosystem” via find-and-replace with “ “ from the abstracts to improve the quality of the results.

BERTopic provided us with a list of 48 topics (see Table 2.1). While the unsupervised topic modeling step has indeed produced a number of interesting and well-defined topics, it has also resulted in topics that lack coherence or are challenging to decipher.

Table 2.1. Overview of topics emerging out of BERTTopic

Anchor #	**Topic words**
anchor0	['research', 'abstract', 'study', 'business', 'digital', 'development', 'new']
anchor1	['platform', 'platforms', 'complementors', 'digital', 'developers', 'complementary', 'product', 'study', 'software', 'abstract']
anchor2	['service', 'value', 'cocreation', 'actors', 'logic', 'institutional', 'creation', 'research', 'servicedominant', 'study']
anchor3	['entrepreneurial', 'regional', 'policy', 'entrepreneurship', 'local', 'industrial', 'economic', 'region', 'clusters', 'development']
anchor4	['financial', 'fintech', 'mobile', 'payment', 'banking', 'payments', 'services', 'money', 'banks', 'technology']
anchor5	['knowledge', 'firms', 'performance', 'smes', 'study', 'collaboration', 'results', 'collaborative', 'export', 'research']
anchor6	['smart', 'city', 'cities', 'urban', 'mobility', 'public', 'governance', 'citizens', 'open', 'development']
anchor7	['tourism', 'hospitality', 'smart', 'hotel', 'destination', 'destinations', 'service', 'research', 'tourists', 'wine']
anchor8	['entrepreneurship', 'university', 'entrepreneurial', 'students', 'education', 'universities', 'academic', 'graduates', 'educational', 'research']
anchor9	['china', 'chinese', 'chinas', 'cruise', 'development', 'government', 'industrial', 'firms', 'industry', 'technology']
anchor10	['green', 'environmental', 'energy', 'water', 'sustainable', 'climate', 'sustainability', 'natural', 'carbon', 'development']
anchor11	['blockchain', 'decentralized', 'technology', 'energy', 'information', 'trust', 'distributed', 'paper', 'dlt', 'applications']
anchor12	['digital', 'transformation', 'servitization', 'business', 'digitalization', 'technologies', 'benefits', 'management', 'new', 'convergence']
anchor13	['healthcare', 'health', 'care', 'patients', 'patient', 'digital', 'service', 'mhealth', 'telehealth', 'value']
anchor14	['social', 'entrepreneurship', 'se', 'entrepreneurs', 'enterprise', 'partnerships', 'enterprises', 'study', 'chapter', 'entrepreneurial']
anchor15	['entrepreneurship', 'entrepreneurial', 'countries', 'entrepreneurs', 'economic', 'growth', 'pillars', 'startup', 'africa', 'women']
anchor16	['business', 'model', 'models', 'chain', 'supply', 'logistics', 'dts', 'barriers', 'study', 'literature']
anchor17	['circular', 'economy', 'ce', 'environmental', 'business', 'model', 'sustainability', 'economic', 'transition', 'circularity']

anchor18	['data', 'big', 'datadriven', 'apis', 'hadoop', 'business', 'agricultural', 'applications', 'open', 'science']
anchor19	['open', 'oi', 'openness', 'research', 'organisational', 'strategy', 'taxonomy', 'competency', 'management', 'source']
anchor20	['brazil', 'brazilian', 'incubation', 'incubators', 'sas', 'startups', 'tbis', 'army', 'companies', 'government']
anchor21	['covid19', 'pandemic', 'crisis', 'insurance', 'healthcare', 'citizen', 'response', 'statistical', 'nsds', 'lockdown']
anchor22	['value', 'networks', 'business', 'firms', 'network', 'paper', 'em', 'management', 'mtsa', 'diamond']
anchor23	['universities', 'university', 'regional', 'ubc', 'universityindustry', 'innovative', 'development', 'role', 'research', 'talents']
anchor24	['venture', 'startups', 'capital', 'startup', 'investors', 'funding', 'sdss', 'new', 'support', 'exit']
anchor25	['helix', 'triple', 'quadruple', 'regional', 'model', 'democracy', 'qh', 'quintuple', 'social', 'society']
anchor26	['learning', 'education', 'educational', 'teachers', 'students', 'school', 'pedagogical', 'design', 'social', 'schools']
anchor27	['iot', 'business', 'things', 'resistance', 'internet', 'ecommerce', 'models', 'agrifood', 'companies', 'paradigm']
anchor28	['pharmaceutical', 'drug', 'medicine', 'vaccines', 'diy', 'drugs', 'biotech', 'gs', 'personalized', 'cros']
anchor29	['digital', 'entrepreneurship', 'dee', 'digitalization', 'entrepreneurial', 'netizens', 'maker', 'organizing', 'rri', 'research']
anchor30	['sustainable', 'entrepreneurship', 'entrepreneurial', 'sustainability', 'environmental', 'upscaling', 'literatures', 'section', 'policy', 'green']
anchor31	['intellectual', 'property', 'ip', 'ic', 'index', 'output', 'countries', 'gii', 'country', 'performance']
anchor32	['crowdfunding', 'crowdsourcing', 'crowd', 'crowds', 'backers', 'projects', 'campaign', 'platform', 'psos', 'detes']
anchor33	['russian', 'development', 'regional', 'russia', 'textile', 'innovative', 'assessment', 'clusters', 'economic', 'industrial']
anchor34	['european', 'union', 'eu', 'policy', 'national', 'countries', 'forests', 'failures', 'europe', 'level']
anchor35	['food', 'waste', 'consumer', 'production', 'rice', 'farmers', 'consumers', 'systems', 'responsibilisation', 'afn']
anchor36	['transfer', 'offices', 'ttos', 'universities', 'technology', 'tto', 'university', 'entrepreneurial', 'knowledge', 'research']
anchor37	['disruptive', 'disruption', 'disruptions', 'incumbents', 'entrants', 'new', 'forking', 'technologies', 'disruptors', 'commercialization']
anchor38	['network', 'contexts', 'analysis', 'keywords', 'armory', 'research', 'km', 'springfield', 'ttns', 'interlocks']
anchor39	['3d', 'printing', 'manufacturing', 'printed', 'technologies', 'crowdfunding', 'new', 'products', 'firms', 'product']
anchor40	['sport', 'sports', 'esports', 'football', 'economics', 'mses', 'diplomacy', 'marketing', 'audience', 'measurement']
anchor41	['labs', 'living', 'lab', 'lls', 'diy', 'setting', 'open', 'user', 'reallife', 'users']
anchor42	['energy', 'bipv', 'business', 'buildings', 'building', 'market', 'dsm', 'ladder', 'lowcarbon', 'switching']

anchor43	['spinoffs', 'usos', 'university', 'spinoff', 'uso', 'innovative', 'academic', 'universities', 'capital', 'creation']
anchor44	['ai', 'intelligence', 'performance', 'capabilities', 'business', 'value', 'scf', 'artificial', 'organizations', 'cocreation']
anchor45	['accelerators', 'accelerator', 'corporate', 'agtech', 'builder', 'programs', 'entrepreneurial', 'corporations', 'zealand', 'micros']
anchor46	['india', 'startup', 'startups', 'indian', 'expectancy', 'millennial', 'trends', 'triz', 'lifecycle', 'indias']
anchor47	['global', 'international', 'mnes', 'multinational', 'ib', 'mne', 'vc', 'internalisation', 'born', 'mnc']

With the **third step**, our goal was to already prepare for Step 6 and 7’s semi-supervised topic modeling using the CorEx algorithm: CorEx relies on user guidance via input of topics. To generate further user-guided topics for CorEx, our employed method combined the capabilities of llama_index, an indexing system, with GPT 3.5, one of OpenAI's most advanced language models. This represents an approach to information retrieval and document analysis. At its core, the method transforms vast textual data - in our case, the corpus of abstracts – into vector representations using GPT 3.5 and then indexes those vectors with llama_index for efficient storage and querying.

The reason this combination is particularly powerful stems from GPT 3.5's semantic understanding of text. Instead of relying on mere keyword matches, the model discerns underlying meanings, themes, and contexts, ensuring that document indexing and subsequent retrieval are based on genuine semantic relevance. This semantic depth means that when researchers query the system, they are not just retrieving documents with matching keywords; they are accessing content that genuinely aligns with their inquiry's intent (OpenAI, 2023).

GPT3.5 generated vector representations of our corpus, capturing their essence in a format suitable for large-scale indexing. Once these vectors were created, llama_index took over in our python script, organizing them in a manner that facilitates rapid querying.

We queried the index using a specific GPT3.5 prompt to identify topic families and their corresponding topics.

PROMPT = "YOUR ONLY KNOWLEDGE IS THE MANY ABSTRACTS OF SCIENTIFIC PAPERS YOU HAVE JUST BEEN TRAINED ON. DONT HALLUCINATE ANY INFORMATION. YOUR JOB IS TO IDENTIFY 20 TOPIC FAMILIES. PLEASE GIVE ME LINE BY LINE THE TOPIC FAMILIES AND THEIR CORRESPONDING TOPICS (MAX 3 TOPICS PER TOPIC FAMILY)"

Starting with the phrase, "YOUR ONLY KNOWLEDGE IS THE MANY ABSTRACTS OF SCIENTIFIC PAPERS YOU HAVE JUST BEEN TRAINED ON," the prompt sets the context for the model. It reminds the model to focus solely on the information from the scientific abstracts it has been trained on and to avoid pulling information from elsewhere.

The directive, "DONT HALLUCINATE ANY INFORMATION," is crucial. Machine learning models, especially large language models, can sometimes generate or "hallucinate" information that is not strictly based on their training data (Lin et al., 2022). By stating this explicitly, the prompt aims to ensure that the model sticks to factual and accurate information derived from the abstracts.

Then, the primary task is delineated: "YOUR JOB IS TO IDENTIFY 20 TOPIC FAMILIES." This instructs the model to categorize or group the information into 20 distinct overarching themes or "topic families." Interestingly, GPT3.5 could only identify 16 topics, suggesting that it could not discern more subjects from the dataset.

Lastly, "PLEASE GIVE ME LINE BY LINE THE TOPIC FAMILIES AND THEIR CORRESPONDING TOPICS (MAX 3 TOPICS PER TOPIC FAMILY)" provides a format for the output. It not only asks for the identified topic families but also wants a breakdown of up to three specific topics within each of those families, all presented in a structured, line-by-line format. Table 2.2 shows the 16 topics and topic constructs that emerged out of this prompt.

Table 2.2. Overview of topics emerging out of GPT based topic modeling

Topic Family #	Topic Name	Corresponding Topics
Topic Family 1	Corporate Entrepreneurship	Technological Applications, Business and Operations Management, Motivation Principles
Topic Family 2	Innovation	Human-Computer Interaction, Value Co-Creation, Business Models
Topic Family 3	Climate Change	Carbon Dioxide Emissions, Energy Usage of Bitcoin, Climate Impact
Topic Family 4	Sustainable Development	Public Policies, Quality of Life, Social Change
Topic Family 5	Pharmaceutical Business Ecosystem	Framework Conditions, System Conditions, Hospital Coordination
Topic Family 6	Open Science	Co-Creation, Open Access Publishing, Open Source Software
Topic Family 7	Innovation Management	Ecosystems, Innovation, Business Ecosystems
Topic Family 8	Human-Computer Interaction	Corporate Entrepreneurship, Employee Engagement, Innovation
Topic Family 9	Cloud Computing	Service Platforms, Resource Agglomeration, Service Pattern Innovation
Topic Family 10	Sustainable Development	Indicators, Innovation, Public Policies
Topic Family 11	Pharmaceutical Business Ecosystems	Relational Governance, Contractual Governance, Coordination
Topic Family 12	Open Science	Framework Conditions, System Conditions, Outcomes
Topic Family 13	Internet of Things	Value Proposition, Architecture, Reward Model
Topic Family 14	Cross-Sectoral Collaboration	Socio-Technical Environment, Organizational Dynamics, Ecosystem Strategies
Topic Family 15	Women-Led Businesses	Attitude, Personality Traits, Awareness
Topic Family 16	COVID-19 Restrictions	Co-Presence in Organizations, Dynamic Temporary Co-Presence, Local 'Buzz'

Step four. Preprocessing is seen as a crucial step in Natural Language Processing and Machine Learning because it can significantly improve the performance of a model (Camacho-Collados & Pilevar, 2017; Jurafsky & Martin, 2023). However, it was intentionally skipped for BERTopic due to its native ability to handle raw textual data (Grootendorst, 2022a). Also, OpenAI's GPT 3.5 language model does not require preprocessing of data due to its ability to understand "how a human would" (Romero, 2021).

Conversely, the original corpus was preprocessed for analysis with CorEx. This preprocessing is deemed vital as removing noise, and standardizing the dataset ultimately enhances the precision of the resulting topics, thereby enabling the extraction of more meaningful insights (Gallagher et al., 2017).

Therefore, via python scripts, we performed a sequence of operations to clean and prepare our set of abstracts for further analysis.

In the initial preprocessing step, we converted the entire corpus to lowercase to ensure consistency in word representations, treating words like 'The' and 'the' as identical. Moreover, we removed recurring words such as “Abstract” or “Purpose of this paper”. To further refine the text, we removed punctuation marks as they do not provide meaningful information for topic modeling. Additionally, we eliminated numbers from the corpus as they often lack substantial relevance for topic analysis. We also removed URLs that were added by the automatic Scopus Export as they would not benefit the topic analysis.

Next, we employed tokenization to break down the text into individual words or "tokens," enabling a more granular analysis at the word level. In order to streamline the analysis and focus on the most meaningful words, we then eliminated common stop words that often carry little semantic value and can introduce noise to the results.

To further enhance the quality of the tokens, we applied lemmatization, a technique that reduces words to their base or root form. By reducing words like 'running', 'runs', and 'ran' to their lemma 'run', we effectively consolidate different forms of the same word, enabling better clustering and understanding of the underlying topics within the corpus.

In the **fifth step,** we employed the unsupervised topic results of BERTopic from Step 2 to initiate the search for significant topics within the corpus. In other words, the BERTopic-generated topics served as initial anchors for the semi-supervised learning phase with CorEx.

Initially, we set the anchor strength to the level of 2, allowing for some flexibility in deviating from the initial list of anchor words for generating topics.

The CorEx procedure is iterative, allowing for ongoing refinement. While it has automated components, human judgment of the topics is essential. During the process, we adjusted the anchor words based on our expertise and understanding of the topic of innovation ecosystem to align results with real-world context. During the seven iterations, we removed unrelated words that would have no connection to the other words (and would not represent the topic) from the anchor lists and tried to add words that, based on our understanding of the topic, represented the topic better. We also proactively scoured for expected topics within the corpus, adding completely new anchor words during the iterations. These words were added to the starting list of anchor words for the next iteration. For instance, in several iterations, we added anticipated anchor words like ‘Innovation Hub’, ‘Collaboration’, ‘Innovation metrics’, ‘Disruption’, ‘Monopoly’. After seven iterations, CorEx presented a set of 31 topics alongside their representing words for the topics.

It is worth noting that CorEx chose to omit certain topics presented by BERTopic due to their suboptimal anchor strength and, thus, low interpretability out of the CorEx topic models. For illustration, among the dropped topics were “'3D', 'PRINTING', 'MANUFACTURING', 'PRINTED', 'TECHNOLOGIES', 'CROWDFUNDING', 'NEW', 'PRODUCTS', 'FIRMS', 'PRODUCT' as well as 'SPORT', 'SPORTS', 'ESPORTS', 'FOOTBALL', 'ECONOMICS', 'MSES', 'DIPLOMACY', 'MARKETING', 'AUDIENCE', 'MEASUREMENT', indicating CorEx would not find sufficient topic strength of these topics within the initial set topics, provided by Step 2 – BERTopic.

Occasionally, places like cities and countries, such as "Lisbon", “California”, "Brazil", and "Russia", appeared as a result within anchors strings. While we attempted to categorize these locations into distinct topic anchors, none stood out as a dominant topic within the corpus.

Additionally, we couldn't form a topic using multiple locations, suggesting a lack of prominent discussion about Innovation Ecosystem locales. Consequently, we designated these locations as stop words to exclude them from topics.

We decided to stop after 7 iterations as formal evaluation metrics (total correlations explained) and the quality of the generated topics no longer demonstrated significant improvement.

To further enhance the richness of our analysis, we combined in **step 6 and 7** the topics list with the GPT results from Step 3. Out of GPT's generated 16 topics, 10 were clearly distinguishable from our list of topics generated before. Therefore, we injected 10 new topics and engaged in additional CorEx iterations. Although CorEx discarded some topics in subsequent iterations due to their limited strength (for instance, Human-Computer Interaction), GPT3.5 insights proved beneficial to yield new topics such as "Internet of things". This topic has been confirmed by CorEx via the addition of topic-describing words to the new anchor. In this way, the inputs from GPT3.5 helped identify additional topics that CorEx, as well as our human understanding of the overall topic, might have missed. Moreover, GPT3.5 also helped confirm topics identified already, such as Covid-19 and Sustainable Development, proving their relevancy and CorEx's correct functioning.

During our iterative process, we incrementally enhanced the anchor strength parameter. This adjustment meant the algorithm progressively emphasized the predefined anchors in topic generation. Furthermore, we also calculated the total correlation score, which indicates the quality of the topic model. When the total correlation score no longer improved, we proceeded with the next step.

In total, 14 CorEx iterations were executed, each helping to sharpen distinct thematic nuance. Throughout all cycles, we merged related topics and separated conflated ones.

Table 2.3 presents the final selection of topics that the CorEx algorithm crystallized in its 14th iteration.

Table 2.3. Final selection of topics based on CorEx 14 iterations

Anchor #	**Key Words**	**Topic Name**
anchor0	['software', 'developer', 'software ecosystem', 'generativity', 'platform', 'digital platform', 'thirdparty developer', 'software service', 'complementors', 'app']	Software Ecosystem
anchor1	['fintech', 'payment', 'financial', 'mobile', 'mobile payment', 'financial service', 'banking', 'bank', 'mobile money', 'money']	Financial Service
anchor2	['city', 'smart city', 'smart', 'urban', 'urban innovation', 'urban development', 'city project', 'smart city project', 'city development', 'smart city development']	Smart City
anchor3	['tourism', 'destination', 'hotel', 'hospitality', 'smart tourism', 'tourism destination', 'heritage', 'tourist', 'tourism ecosystem', 'tourism hospitality']	Tourism
anchor4	['university', 'student', 'education', 'higher education', 'entrepreneurship education', 'educational', 'teacher', 'school', 'graduate', 'course']	Education
anchor5	['entrepreneurial', 'entrepreneurial ecosystem', 'entrepreneurial intention', 'intention', 'entrepreneurial activity', 'relationship entrepreneurial', 'graduate entrepreneurial', 'urban entrepreneurial', 'urban entrepreneurial ecosystem', 'entrepreneurial intention student']	Entrepreneuship
anchor6	['blockchain', 'blockchain technology', 'decentralized', 'dlt', 'ledger', 'ledger technology', 'smart contract', 'blockchainbased', 'blockchains', 'distributed ledger']	Blockchain
anchor7	['business model', 'business model innovation', 'model innovation', 'bmi', 'model', 'business', 'servitization', 'digital servitization', 'innovation bmi', 'ecosystem business model']	Business Model
anchor8	['healthcare', 'health', 'care', 'patient', 'health care', 'digital health', 'medical', 'medicine', 'hospital', 'healthcare ecosystem']	Healthcare
anchor9	['social innovation', 'social', 'social entrepreneurship', 'social enterprise', 'social innovation ecosystem', 'social entrepreneur', 'social impact', 'social problem', 'dsi', 'social innovator']	Social Innovation
anchor10	['iot', 'thing', 'internet thing', 'internet', 'thing iot', 'internet thing iot', 'iot platform', 'attack', 'cyber', 'iot architecture']	IoT
anchor11	['green', 'environmental', 'sdgs', 'sustainable', 'sustainable development', 'green innovation', 'sustainable development goal', 'green economy', 'sustainability', 'environmental regulation']	Sustainability

anchor12	['crowdfunding', 'crowdsourcing', 'crowd', 'crowdfunding platform', 'backer', 'campaign', 'crowdfunding ecosystem', 'crowdsourcing process', 'crowd innovation', 'open science']	Crowdfunding
anchor13	['food', 'waste', 'food system', 'agrifood', 'food ecosystem', 'food waste', 'food sector', 'innovation food', 'global food', 'farmer']	Food
anchor14	['energy', 'carbon', 'clean energy', 'clean', 'renewable', 'renewable energy', 'energy system', 'electricity', 'offshore', 'wind']	Renewable Energy
anchor15	['accelerator', 'corporate accelerator', 'corporate', 'ecosystem builder', 'business accelerator', 'accelerator program', 'role accelerator', 'builder', 'zealand', 'new zealand']	Accelerator
anchor16	['helix', 'triple helix', 'triple', 'quadruple', 'quadruple helix', 'helix model', 'triple helix model', 'quadruple helix model', 'helix innovation', 'quintuple']	Innovation Helix
anchor17	['supply chain', 'supply', 'chain', 'logistics', 'chain ecosystem', 'procurement', 'chain management', 'supply chain ecosystem', 'supply chain management', 'scm']	Supply Chain Management
anchor18	['smes', 'enterprise smes', 'sme', 'small', 'small medium', 'smes innovation', 'medium enterprise smes', 'smes innovation performance', 'manufacturing smes', 'dihs']	SME
anchor19	['coopetition', 'orchestration', 'ecosystem orchestration', 'ecosystem coopetition', 'coopetition strategy', 'resource orchestration', 'ally', 'shared goal', 'mne', 'competition']	Orchestration
anchor20	['service', 'service innovation', 'service ecosystem', 'value cocreation', 'cocreation', 'sd', 'servicedominant', 'sd logic', 'logic', 'resource integration']	Cocreation
anchor21	['index', 'country', 'gei', 'gii', 'output', 'dea', 'innovation output', 'efficiency', 'global entrepreneurship', 'aspiration']	Innovation Performance
anchor22	['intellectual property', 'property', 'intellectual', 'ipr', 'ip', 'appropriability', 'protection intellectual', 'ip index', 'property right', 'intellectual property right']	Intellectual Property
anchor23	['circular', 'circular economy', 'circular business', 'circular business model', 'circularity', 'waste', 'circular ecosystem', 'economy', 'towards circular', 'transition circular']	Circular Economy
anchor24	['covid', 'pandemic', 'crisis', 'insurance', 'covid pandemic', 'covid crisis', 'lockdown', 'response', 'post covid', 'emergency']	Covid
anchor25	['open innovation', 'open innovation ecosystem', 'open', 'open innovation strategy', 'knowledge transfer', 'openness', 'understanding open innovation', 'understanding open', 'within open innovation', 'within open']	Open Innovation
anchor26	['ai', 'big data', 'big', 'artificial', 'artificial intelligence', 'artificial intelligence ai', 'intelligence ai', 'data', 'ai innovation', 'ai innovation ecosystem']	Artifical Intelligence

The output of the CorEx iterations can be used to calculate topic scores for the abstracts in the database. Therefore, in **Step 8**, we relied on the outcome of our final CorEx iteration to generate a list for each topic containing the 10 most relevant words as well as the topic correlation score of each representative word. The topic correlation score indicates how relevant the particular word is for the topic. Subsequently, we used this information to calculate scores on each topic for each abstract in our database. For every abstract, we applied the following procedure. First, we generated a list of all words that were part of the cleaned abstracts. Subsequently, we also created lists containing all possible bigrams and trigrams for the cleaned description. For each word, bigram, and trigram, we subsequently checked whether it was part of the 10 most relevant words, bigrams, or trigrams of a particular topic. If this was the case, the abstract received the topic correlation score of that particular word. If more than one word, bigram, or trigram in the abstract was part of the relevant word, bigram, or trigram of a particular topic, the abstract received the sum of the topic correlation scores of these particular words. To make the scores comparable across different topics, we normalized the final topic scores (Lakshmanan, 2019). Following this procedure, every abstract from our corpus received a normalized topic score for each topic. The higher the value on a particular topic score, the more representative the particular topic for the paper's abstract.

An example of this can be seen in Table 2.4. This particular abstract shows strong anchor strength for the identified topic "Business Model". There is some correlation with the topics "Software Ecosystem" and "Value Co-Creation". However, there is no correlation at all with the exemplary topics "Smart City", "Blockchain" or "Healthcare".

Table 2.4: Assigning topic scores to abstracts

Authors	Heshmatisafa S.; Seppänen M.
Title	Exploring API-driven business models: Lessons learned from Amadeus's digital transformation
Year	2023
Abstract Tokenized	['digital', 'transformation', 'compels', 'company', 'seek', 'new', 'strategy', 'business', 'model', 'allowing', 'firm', 'find', 'new', 'way', 'create', 'capture', 'value', 'apis', 'widely', 'used', 'improving', 'internal', 'process', 'opened', 'recently', 'thirdparty', 'developer', 'enable', 'innovation', 'digital', 'ecosystem', 'however', 'study', 'documented', 'success', 'story', 'value', 'generation', 'employing', 'public', 'apis', 'thus', 'enabling', 'new', 'apidriven', 'business', 'model', 'single', 'case', 'study', 'examined', 'digital', 'transformation', 'one', 'leading', 'organization', 'travel', 'industry', 'amadeus', 'corporation', 'publicly', 'available', 'document', 'gray', 'literature', 'used', 'data', 'content', 'analysis', 'result', 'show', 'amadeus', 'transformed', 'business', 'model', 'apidriven', 'business', 'model', 'study', 'contributes', 'literature', 'apidriven', 'business', 'model', 'demonstrating', 'implication', 'apis', 'strategic', 'tool', 'value', 'cocreation', 'identified', 'six', 'different', 'value', 'generated', 'exposing', 'digital', 'asset', 'api', 'consumer', 'apidriven', 'business', 'model', 'selfservice', 'platform', 'commercialization', 'ecosystem', 'big', 'data', 'market', 'marketplace', 'entrepreneurship', 'rd', '©', 'author']
Abstract Original	Digital transformation compels companies to seek new strategies and business models allowing firms to find new ways to create and capture value. APIs have been widely used in improving internal processes and have been opened up recently to third-party developers to enable innovations in digital ecosystems. However few studies have documented success stories of value generation through employing public APIs and thus enabling new API-driven business models. This single case study examined the digital transformation in one of the leading organizations of the travel industry Amadeus Corporation. Publicly available documents and other gray literature were used as data for content analysis. The results show how Amadeus has transformed its business model into an API-driven business model. This study contributes to the literature on API-driven business models by demonstrating the implication of APIs as a strategic tool in value co-creation. We identified six different values that were generated by exposing digital assets to API consumers through the API-driven business models of self-service platforms: commercialization ecosystem big data market and marketplace entrepreneurship and R&D. © 2023 The Authors
Software Ecosystem	0,03473933
Smart City	0
Blockchain	0
Business Model	0,74926249
Healthcare	0
Value Co-Creation	0,04765974

In Step 9, we subsequently used the topic scores to engage in a bibliographic analysis of the database of innovation ecosystem abstracts. The results will be presented in the following chapter.

2.4. FINDINGS

2.4.1. Methodology

We employed the set of algorithms discussed earlier to perform a bibliometric analysis of the literature on innovation ecosystems. For further data analysis, we started with the output of the CorEx algorithm. It provided a matrix in an Excel format, where the rows corresponded to the 2567 papers ingested, and columns represented the 27 identified topics.

We added the papers' titles and author's or authors' names. Given the temporal nature of our database, the year of publication was added, allowing us to trace the evolution of papers and themes over time. Further delving into the academic context of each paper, we added the hosting journal's name. This information would later help to identify which academic platforms predominantly catered to which topics. Recognizing the importance of a paper's reception in the academic community, we also added the number of citations each paper had received. The citations can be seen as an indication of a paper's impact and relevance, thus importance in subsequent analyses. Lastly, to understand the content and context, we incorporated the original abstracts of the papers into our enriched dataset. The abstracts offered quick insights into the papers' core themes and helped with manually validating our findings and analysis.

2.4.2. Data Analytics

Number of papers published annually. Our data shows a steady growth in the number of papers published annually on the topic of "innovation ecosystem." The continuous growth indicates an increasing relevance and interest in innovation ecosystems. There is an active and thriving community of scholars, researchers, and practitioners contributing to and drawing from this body of knowledge. However, the sheer volume also shows that traditional, manual literature review methods may struggle to follow up with the body of knowledge due to its large volume.

Figure 2.2. Number of publications

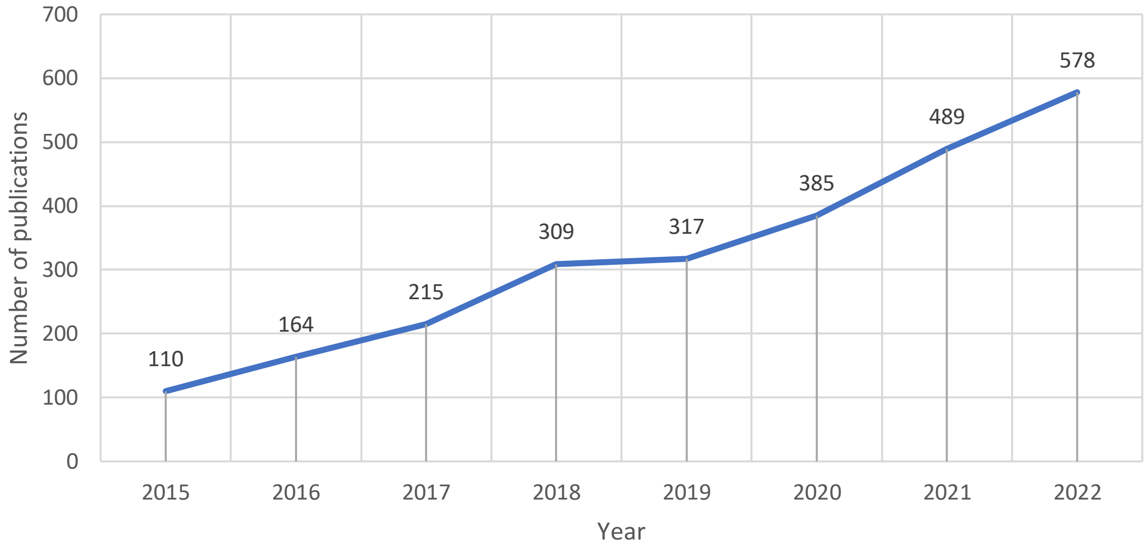
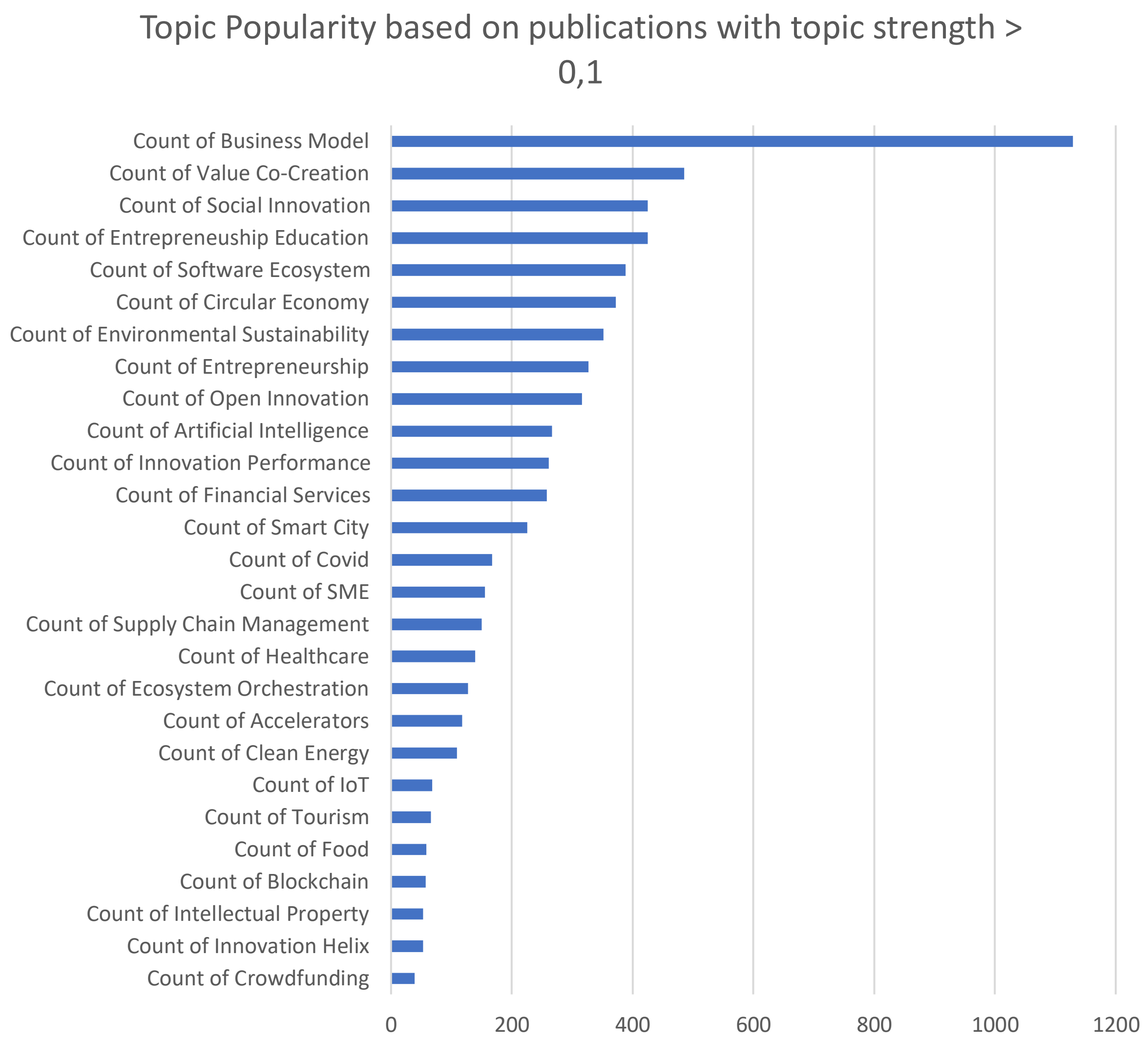


Table 2.5. Innovation Ecosystem Paper Growth

Year	Number of Papers
2015	110 papers
2016	164 papers, representing a 49.1% increase from 2015
2017	215 papers, a growth of 31.1% from 2016
2018	309 papers, indicating a 43.7% increase from 2017
2019	317 papers, a modest 2.6% rise from 2018
2020	385 papers, marking a 21.5% increase from 2019
2021	489 papers, showing a 27% growth from 2020
2022	578 papers, an 18.2% increase from 2021

Based on the number of publications matching the topics, the most relevant topics are as follows in descending order

Figure 2.3. Topic Popularity

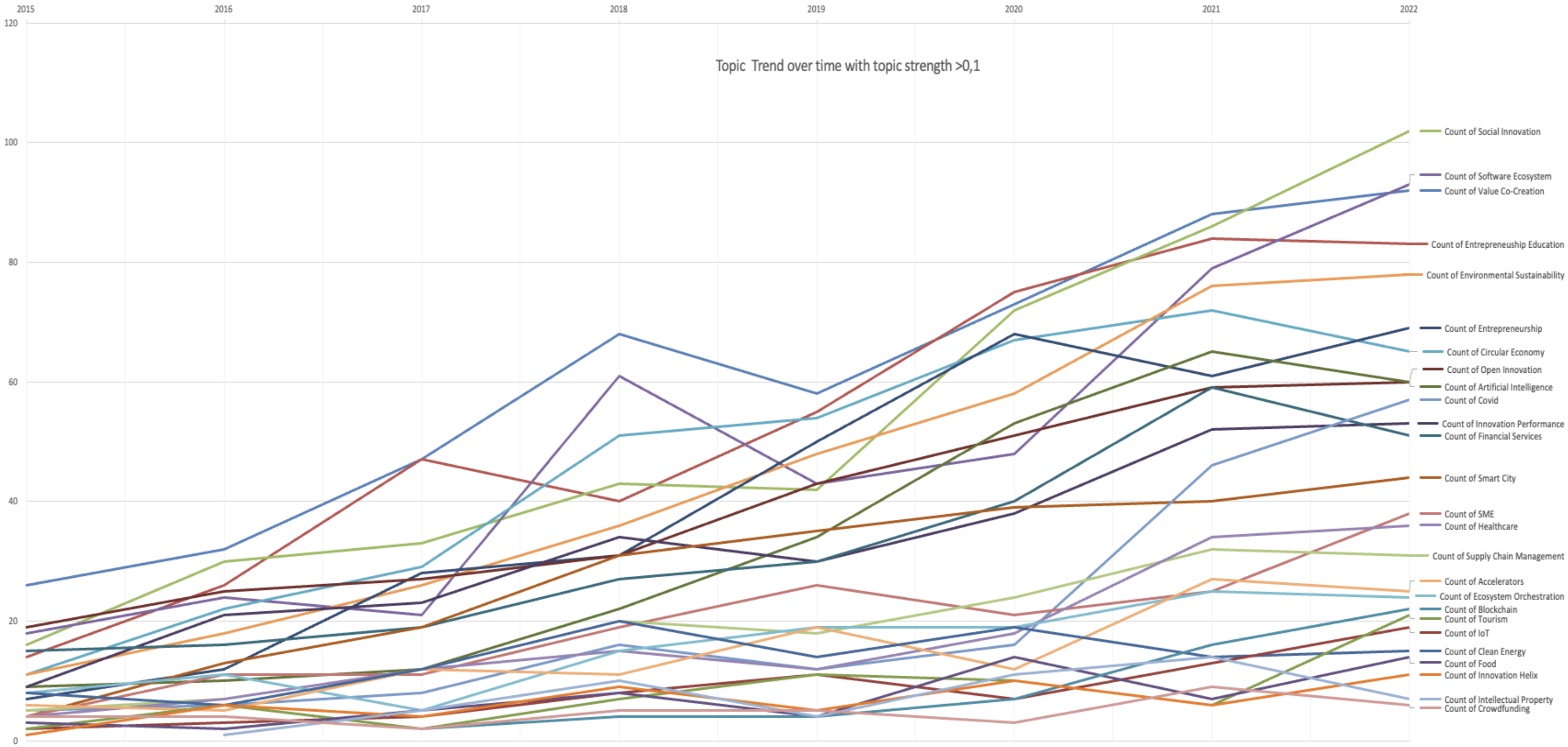


The by-far most popular topic is "Business Model". On the contrary, the least popular topic is "Crowdfunding".

It's important to mention that our analysis employed a topic anchor strength threshold of 0.1 to sieve out weak associations between papers and their corresponding topics. We followed this approach to focus on papers where the topic was substantially represented, thereby offering a clearer view of the topic's significance.

We further analyze how the **prevalence of each topic has changed over time**. To do this, in our initial analysis, we count the number of papers each year that are associated with each topic and plot these counts over time. This allows us to see if certain topics are becoming more or less popular.

Figure 2.4. Evolution of topic popularity (excl. positive outlier Business Model)



It is visible that most topics have grown substantially, emphasizing the dynamic nature of research in innovation ecosystems. It also appeared that none of the topics in the dataset have experienced a downward trend over the provided timespan of 2015-2022. This suggests that the relevancy of the topics in the dataset has either remained stable or seen growth in terms of research interest.

As a next step for our trends analysis for each topic, we aimed to take into account the general growth of papers in the domain “Innovation Ecosystem”. Therefore, we followed a heuristic approach based on the change in the proportion of papers published on each topic from the beginning to the end of the available data. This required calculating the normalized topic value each year, which was done by dividing the count of papers for the topic by the total count of papers across all topics for each year. This normalization helps to understand the relative focus or interest in each topic compared to others over time. We then looked at the initial (the first year, 2015) and final (the last year, 2022) normalized values for each topic. These values represent the starting and ending points of the trend within the available data. If the final value was significantly higher than the initial value (specifically, more than 20% higher), we categorized it as an "Upward Trend." This indicates growing interest or focus on that topic over the years. Conversely, if the final value was significantly lower than the initial value (specifically, less than 80% of the initial value), we categorized it as a "Downward Trend." This suggests decreasing interest or focus on that topic. If the delta between the initial and final values did not meet these thresholds, we categorized the trend as "Stable/Fluctuating." This indicates that the topic has either maintained a consistent level of interest, or it has varied over the years without a clear direction of increase or decrease.

Figure 2.5. Normalized Evolution of topic popularity

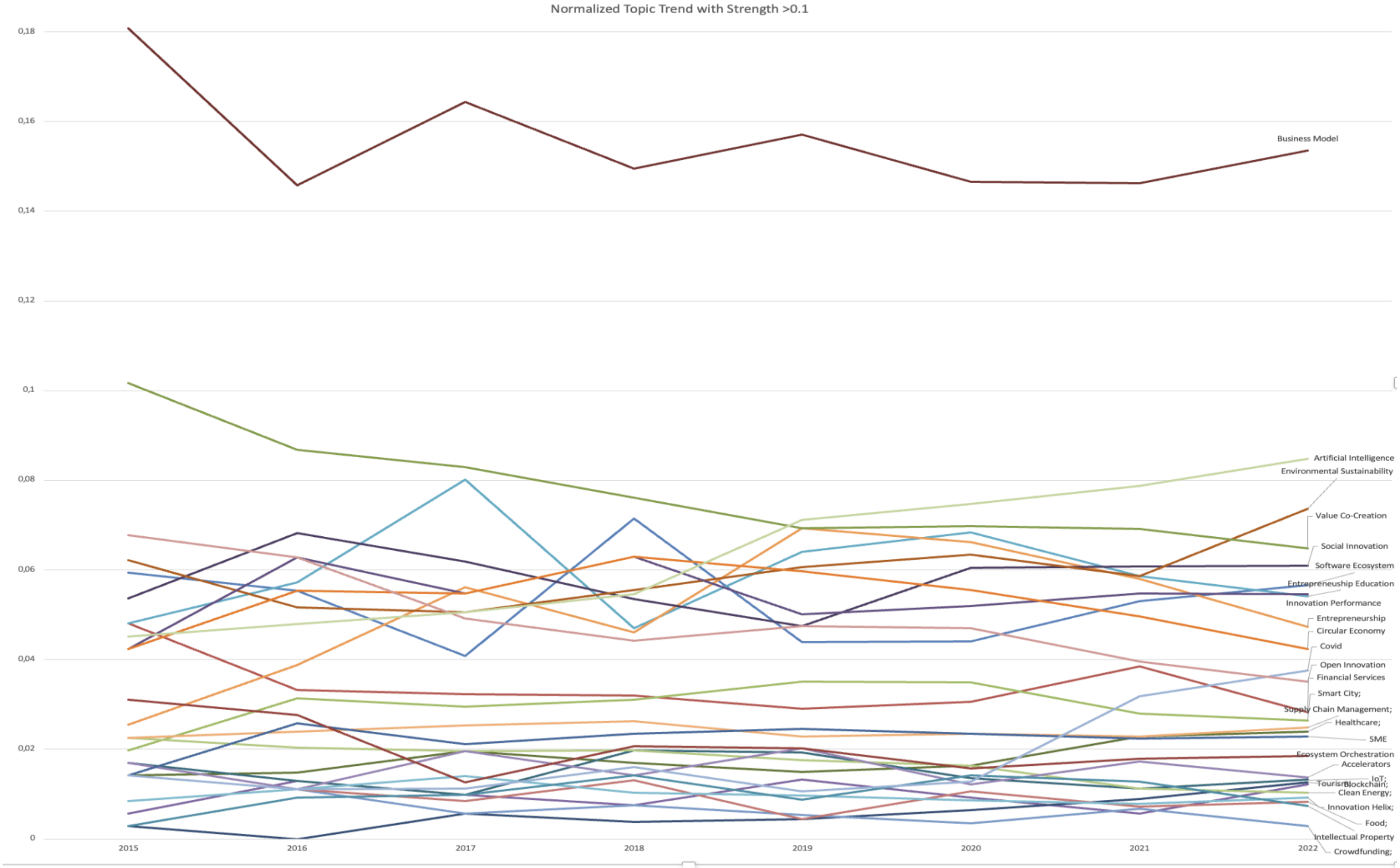


Table 2.6. Normalized topic trends in descending order

#	Topic	Trend	Initial Value	Final Value	Delta
1	Value Co-Creation	Downward Trend	0,099	0,064	0,035
2	Open Innovation	Downward Trend	0,072	0,042	0,03
3	Financial Services	Downward Trend	0,057	0,035	0,022
4	Entrepreneurship	Upward Trend	0,027	0,048	0,021
5	Covid	Upward Trend	0,019	0,04	0,021
6	Clean Energy	Downward Trend	0,03	0,01	0,02
7	Smart City	Upward Trend	0,015	0,031	0,016
8	Ecosystem Orchestration	Downward Trend	0,03	0,017	0,013
9	Environmental Sustainability	Upward Trend	0,042	0,054	0,012
10	Crowdfunding	Downward Trend	0,015	0,004	0,011
11	SME (Small and Medium-sized Enterprises)	Upward Trend	0,015	0,026	0,011
12	Blockchain	Upward Trend	0,004	0,015	0,011
13	Business Model	Stable/Fluctuating	0,194	0,184	0,01
14	Healthcare	Upward Trend	0,015	0,025	0,01
15	Social Innovation	Stable/Fluctuating	0,061	0,071	0,01
16	Artificial Intelligence	Upward Trend	0,034	0,042	0,008
17	Tourism	Upward Trend	0,008	0,015	0,007
18	Accelerators	Downward Trend	0,023	0,017	0,006
19	Entrepreneurship Education	Stable/Fluctuating	0,053	0,058	0,005
20	IoT (Internet of Things)	Stable/Fluctuating	0,008	0,013	0,005
21	Software Ecosystem	Stable/Fluctuating	0,068	0,064	0,004
22	Innovation Helix	Stable/Fluctuating	0,004	0,008	0,004
23	Circular Economy	Stable/Fluctuating	0,042	0,045	0,003
24	Innovation Performance	Stable/Fluctuating	0,034	0,037	0,003
25	Supply Chain Management	Stable/Fluctuating	0,019	0,021	0,002
36	Intellectual Property	Stable/Fluctuating	0,003	0,005	0,002
26	Food	Stable/Fluctuating	0,011	0,01	0,001

The upward trend in areas like Covid, Smart City, Blockchain, and AI suggests a growing academic focus on these areas, possibly due to their increasing real-world significance and potential for future development. Topics like Value Co-Creation and Open Innovation, despite being significant in the business and academic world, are seeing a decline in interest, which might indicate a shift in focus to newer or more pressing issues.

Another perspective on our topics offers the **average distribution of pages**. This distribution of page counts could provide insights into the complexity or depth of discussions on various topics. For instance, topics with longer average page counts might indicate comprehensive methodologies, extensive literature reviews, or complex theoretical discussions. On the other hand, topics with shorter average page counts might lean towards concise, or empirical research. It is essential to note that while page count can provide an indication of the depth or breadth of a paper, it is not a definitive measure of its quality or impact.

Table 2.7. Average Page Count per topic

Topic	**Average Page Count**
Crowdfunding	24
Entrepreneuship Education	18
Supply Chain Management	18
Financial Services	17
SME	17
Intellectual Property	17
Social Innovation	17
Entrepreneurship	17
Circular Economy	17
Smart City	17
Covid	17
Innovation Performance	17
Ecosystem Orchestration	17
Accelerators	17
Business Model	17
Environmental Sustainability	17
Open Innovation	17
Artificial Intelligence	17
Value Co-Creation	17
Food	17
Software Ecosystem	16
Healthcare	16
Innovation Helix	16
Clean Energy	16
Blockchain	15
IoT	15
Tourism	14

Assessing the page counts for papers across different topics, we observed trends. The topic of "Crowdfunding" leads with an average page count of approximately 24 pages. Conversely, "Tourism" tends to have shorter papers, with an average page count of 14. However, while longer papers may provide thorough discussions, the quality and impact of a paper are not necessarily correlated with its length.

To determine which topics are more relevant than others, we look at the **number of citations each paper has received**. We then calculate the average and the median number of citations for papers associated with each topic. Topics with a higher average and median number of citations could be considered more relevant.

Figure 2.6. Relevance of topics in terms of citations

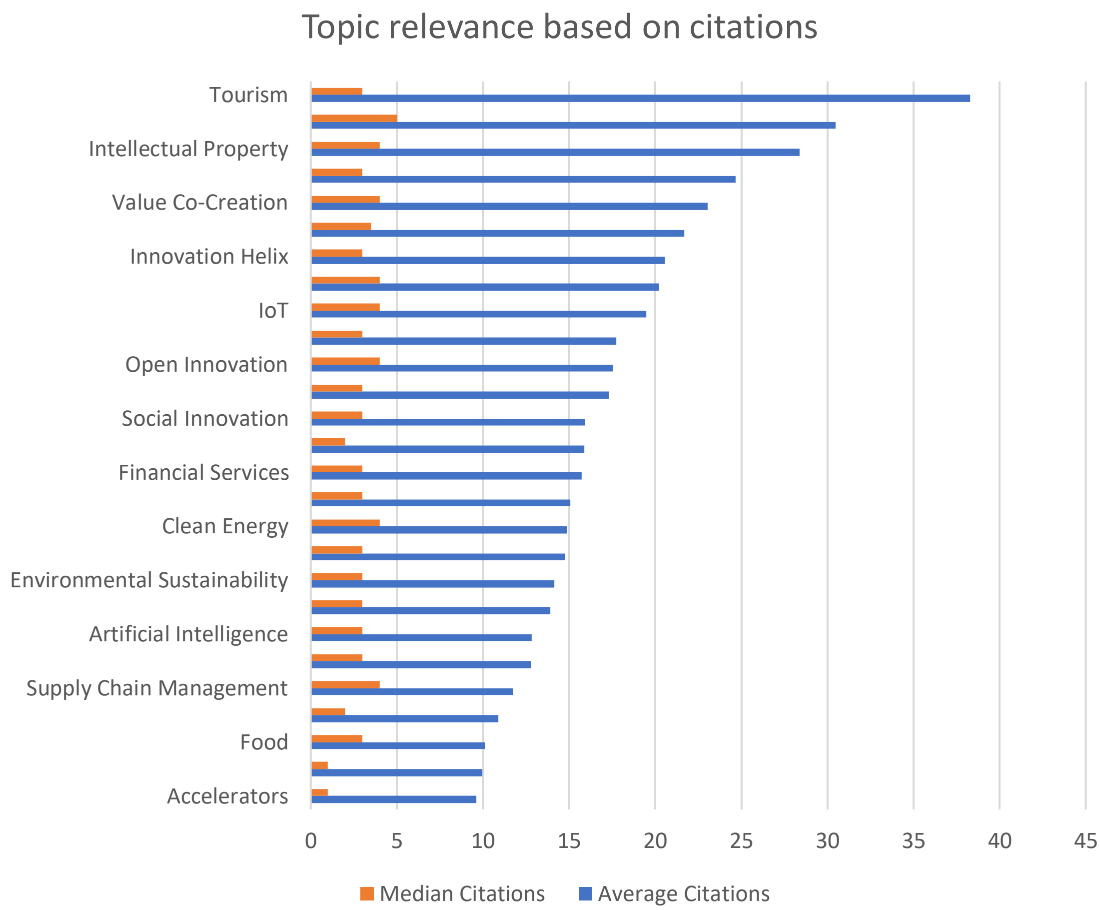


Table 2.8. Average and Median Citation count per topic

Topic	Average Citations	Median Citations
Tourism	38	3
Ecosystem Orchestration	30	5
Intellectual Property	28	4
Smart City	25	3
Value Co-Creation	23	4
Entrepreneurship	22	4
Innovation Helix	21	3
Software Ecosystem	20	4
IoT	19	4
Business Model	18	3
Open Innovation	18	4
Crowdfunding	17	3
Social Innovation	16	3
Blockchain	16	2
Financial Services	16	3
Innovation Performance	15	3
Clean Energy	15	4
Entrepreneurship Education	15	3
Environmental Sustainability	14	3
SME	14	3
Artificial Intelligence	13	3
Circular Economy	13	3
Supply Chain Management	12	4
Healthcare	11	2
Food	10	3
Covid	10	1
Accelerators	10	1

Our findings revealed that the topic of Tourism exhibited the highest average citation count, with a score 38. Interestingly, in our previous analysis, we also observed that the topic of Tourism is at the same time the topic with the lowest average page count. One possible interpretation could be that the shorter average page count might make the papers more accessible and digestible to a broader audience, possibly leading to higher citation counts as more researchers can read and reference the work.

The median count of citations ranges from 5 citations for the topic Ecosystem Orchestration to only one citation for the topics Covid and Accelerators. The lower median citation count for these topics does suggest that the academic discourse around them may still be evolving. Especially for COVID, being a recent global event, the research around it is likely still in its nascent stages, and it will take time for the academic community to engage with and cite this body of work fully. Another perspective could be that our dataset might not encompass some of the foundational papers in these areas, leading to the observed citation dynamics.

The median citation count of three for Tourism places it in the middle field, reflecting a balanced and moderate level of engagement from the academic community, contrasting with its highest average citation count. This suggests the presence of outliers—papers with exceptionally high citation counts—that are influencing the average. Thus, the rather balanced median, coupled with a high average, might indicate diverse engagement levels within the field, with some papers receiving widespread recognition while others have a more moderate citation count. Our subsequent analysis involved two distinct yet complementary strategies to provide an overview of academic contributions on our identified top five topics.

Top Authors Analysis: Given that multiple authors often collaborate on a single paper, we initiated our process by isolating each author. This approach resulted in an expanded list of papers, where each line uniquely attributed a paper and its corresponding topic strengths to a singular author. Consequently, every paper appeared in our refined list as many times as there were contributing authors. We then focused on individual authors, analyzing their contributions across different papers. This approach involved aggregating the topic strengths for each author, accounting for their involvement in multiple research works. By summing the topic strengths, we could effectively measure an author's overall contribution to a particular topic. This analysis highlights the key authors in each field, recognizing those who have significantly influenced the development and understanding of specific topics through their scholarly work.

Top Papers Analysis per topic: To accurately assess the impact of individual papers on our identified topics, we implemented – in contrary to our previous author analysis – a strategy that ensures each paper is counted only once, irrespective of the number of contributing authors. This approach mitigates the potential overrepresentation of papers with multiple authors. We determined the significance of each paper by considering the maximum topic strength it received. This method offers a clear and unbiased view of the most influential papers in each topic, based on their thematic relevance and impact.

In the following, we provide tables for the top five relevant topics as per topic relevance: Business Model, Value Co-Creation, Social Innovation, Entrepreneurship Education, Software Ecosystem

Table 2.9. Most relevant paper analysis for the top five topics (based on topic strength)

Topic	**Top Paper #**	**Title**	**Author full names**
Business Model	1	The evolution of facility management business models in supplier-client relationships	Nardelli Giulia; Rajala Risto
Business Model	2	Empirical research on business model innovation alignment with social relationship affect firm performance.	Guan Yujuan; Zhang Li
Business Model	3	Business model innovation and decision making: uncovering mechanisms for coping with uncertainty	Schneckenberg Dirk; Velamuri Vivek
Business Model	4	Exploring business model innovation for competitive advantage: a lesson from an emerging market	Saqib Natasha; Satar Mir Shahid
Business Model	5	Managing business model innovation: an innovative approach towards designing a digital ecosystem and multi-sided platform	Hoch Nino Bernd; Brad Stelian
Value Co-Creation	1	Incremental and radical open service innovation	Myhren Per; Witell Lars

Value Co-Creation	2	The influence of platform service innovation on value co-creation activities and the network effect	Fu Wenhui; Wang Qiang
Value Co-Creation	3	Process reference frameworks as institutional arrangements for digital service innovation	Iden Jon; Eikebrokk Tom Roar
Value Co-Creation	4	Technology-Based Self-Service (TBSS) Innovations in B2B Settings: An Abstract	Ekman Peter; Raggio Randle
Value Co-Creation	5	An Examination of the Effect of Knowledge Utilization on Service Innovation: The Moderating Roles of Performance-Orientation Culture and Competitiveness Culture	Engelseth Per; Zhang Wenjing
Entrepreneurship Education	1	Life sciences and health in South West Wales: A sub-regional innovation ecosystem	Davies Gareth Huw; Clement Robert
Entrepreneurship Education	2	Personnel development in Chinese innovation-active companies	Zavyalova Elena; Alsufyev Artem
Entrepreneurship Education	3	Development of innovation ecosystem framework for successful adoption of industry 4.0 enabling technologies in Indian manufacturing industries	Pasi Bhaveshkumar Nandanram; Maheshwari Prateek
Entrepreneurship Education	4	Researching Design Policy Ecosystems in Europe	Mortati Marzia; Maffei Stefano
Entrepreneurship Education	5	System Dynamics Modeling of Innovation Ecosystem	Yung Kai Leung; Jiang Zhong-Zhong
Software Ecosystems	1	Grassroots Resistance to Digital Platforms and Relational Business Model Design to Overcome It: A Conceptual Framework	Ricart Joan Enric; Snihur Yuliya
Software Ecosystems	2	Tight and loose coupling in evolving platform ecosystems: The cases of airbnb and uber	Hein Andreas; Böhm Markus
Software Ecosystems	3	The effects of technology standards on complementary innovations	Wen Wen; Forman Chris
Software Ecosystems	4	How do firms meet the challenge of technological change?	Gao Yuchen; Liu Xielin
Software Ecosystems	5	Startup trust model: The role of trust in successful fintech startup ecosystems	Assyne Nana; Adjei Joseph
Social Innovation	1	Building capabilities through social innovation: Implications for the economy and society	Alijani Sharam; Luna Alvaro
Social Innovation	2	Building a new third construction sector through social enterprise	Loosemore Martin
Social Innovation	3	Hybridity of social enterprise models and ecosystems	Neverauskiene Laima Okuneviciute

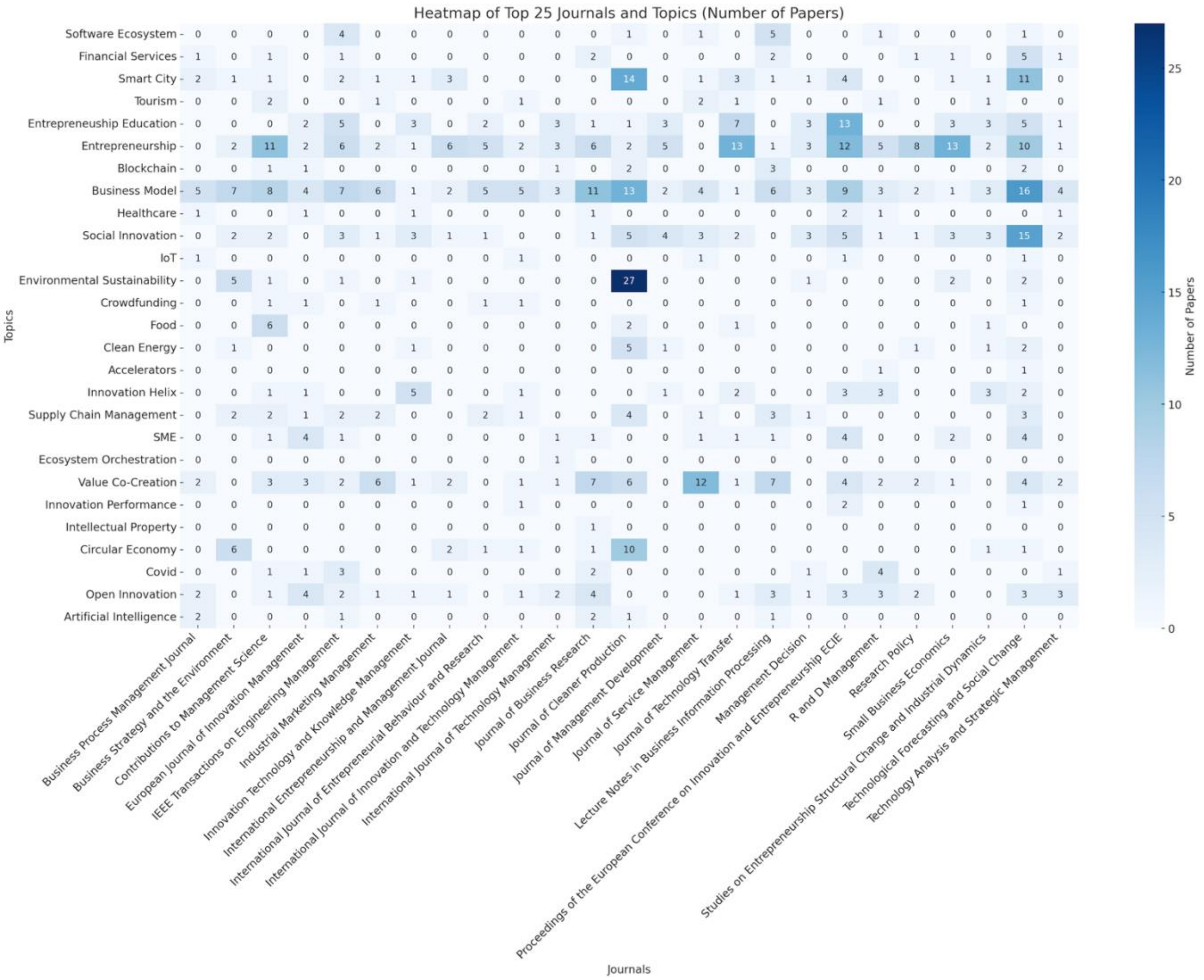
Social Innovation	4	Hybridity of social enterprise models and ecosystems	Macías-Prada John; Vargas-Sáenz Alfredo
Social Innovation	5	Why do some social enterprises flourish in Vietnam? A comparison of human and ecosystem partnerships	Nguyen Minh Hieu Thi; Carr Stuart

Table 2.10. Most relevant authors for the top five topics (based on cumulative topic strength)

Topic	**Top Author #**	**Author**
Business Model	1	Parida V.
Business Model	2	Carayannis E.G.
Business Model	3	Wincent J.
Business Model	4	Chen Y.
Business Model	5	Zhang L.
Value Co-Creation	1	Edvardsson B.
Value Co-Creation	2	Parida V.
Value Co-Creation	3	Koskela-Huotari K.
Value Co-Creation	4	Tronvoll B.
Value Co-Creation	5	Vargo S.L.
Entrepreneurship Education	1	Secundo G.
Entrepreneurship Education	2	Miller K.
Entrepreneurship Education	3	Fischer B.
Entrepreneurship Education	4	Carayannis E.G.
Entrepreneurship Education	5	Meissner D.
Software Ecosystems	1	Krcmar H.
Software Ecosystems	2	Böhm M.
Software Ecosystems	3	Tsujimoto M.
Software Ecosystems	4	Cennamo C.
Software Ecosystems	5	Salerno M.S.
Social Innovation	1	Carayannis E.G.
Social Innovation	2	Guerrero M.
Social Innovation	3	Akaka M.A.
Social Innovation	4	Unceta A.
Social Innovation	5	Luna A.

Delving deeper into the scholarly contributions, our heatmap provides an overview **of the top journals for each topic** based on the number of papers they've published in that domain. To avoid an overwhelming visualization, we'll consider the top journals for each topic based on the number of papers they've published in that domain.

Figure 2.7. Heatmap of journals and topics (topic strength >0,1)



From this heatmap, we can discern which journals are the primary contributors to specific topics. For instance, "Journal of Cleaner Production" appears frequently, highlighting its broad coverage across various topics.

Finally, the following table shows the most relevant journals for the top five topics.

Table 2.11. Most relevant journals for the top 5 topics (based on cumulative topic strength)

Topic	Journal #	Source title
Business Model	1	Technological Forecasting and Social Change
Business Model	2	Journal of Cleaner Production
Business Model	3	Proceedings of the European Conference on Innovation and Entrepreneurship
Business Model	4	Journal of Business Research
Business Model	5	European Journal of Innovation Management
Value Co-Creation	1	Journal of Service Management
Value Co-Creation	2	Technological Forecasting and Social Change
Value Co-Creation	3	Journal of Business Research
Value Co-Creation	4	Journal of Service Research
Value Co-Creation	5	Journal of Cleaner Production
Social Innovation	1	Technological Forecasting and Social Change
Social Innovation	2	Journal of Cleaner Production
Social Innovation	3	Journal of Service Management
Social Innovation	4	International Journal of Entrepreneurial Behaviour & Research
Social Innovation	5	Encontros Bibli
Entrepreneurship Education	1	Proceedings of the European Conference on Innovation and Entrepreneurship
Entrepreneurship Education	2	California Management Review
Entrepreneurship Education	3	IEEE Transactions on Engineering Management
Entrepreneurship Education	4	Journal of Technology Transfer
Entrepreneurship Education	5	Business Horizons
Software Ecosystem	1	Lecture Notes in Business Information Processing
Software Ecosystem	2	Technological Forecasting and Social Change
Software Ecosystem	3	IEEE Transactions on Engineering Management
Software Ecosystem	4	International Journal of Production Economics
Software Ecosystem	5	PICMET (Portland International Conference on Management of Engineering and Technology)

2.5. DISCUSSION

Leveraging novel NLP algorithms. The expansion of academic literature is understood as a challenge for researchers dealing with large research communities. Traditional literature reviews, anchored primarily in manual human effort, are increasingly stretched beyond their efficacy (Bawden & Robinson, 2009; Edmunds & Morris, 2000). In this rapidly evolving academic landscape, staying current requires tools and methods to match the pace and depth of this literature increase. Topic modeling is a text-mining tool for the discovery of hidden semantic structures in a text body. The conventional method for topic modeling is latent Dirichlet allocation (LDA), introduced by Blei, Ng, and Jordan (2003). Although this technique has allowed researchers to identify and use topics in a wide variety of research contexts, it also has important limitations.

Comparison to traditional LDA approaches. For demonstration, we employed the Latent Dirichlet Allocation (LDA) from the python Gensim library to understand its performance in topic modeling within our research domain. Gensim's LDA is particularly important because it is optimized for text corpora, making it a widely adopted and cited tool in academic research for topic modeling tasks (Řehůřek & Sojka, 2010). Using Gensim's LDA provides a benchmark for evaluating new NLP techniques. By comparing its output with that from more sophisticated models, researchers can demonstrate improvements over this baseline, justifying the development and use of advanced algorithms that offer deeper contextual insights (Hoffman et al., 2010).

As shown below, the results obtained through LDA revealed significant limitations when compared to our novel combination of NLP-driven approach.

Table 2.12. LDA Topic Modeling

Topic #	Key Words
1	sustainable, sustainability, environmental, development, economy, economic, policy, green, circular, food
2	social, entrepreneurial, chapter, new, community, research, networks, policy, entrepreneurs, crowdfunding
3	research, paper, study, approach, implications, open, purpose, findings, methodology, analysis
4	service, research, knowledge, industry, based, study, university, approach, universities, design
5	smart, study, cities, city, data, urban, service, business, research, based
6	tourism, technology, india, new, well, world, challenges, future, development, case
7	business, technology, technologies, new, data, systems, blockchain, services, models, information
8	entrepreneurial, start, university, startups, ups, entrepreneurs, development, support, universities, startup
9	firms, knowledge, platform, study, firm, research, open, based, capabilities, management
10	value, business, service, creation, co, model, actors, process, study, models
11	data, smes, enterprises, performance, medium, ie, study, small, sme, user
12	health, healthcare, care, quality, cultural, services, life, development, design, innovative
13	helix, triple, projects, openness, generativity, project, cruise, electric, park, model
14	digital, platforms, transformation, platform, new, industry, fintech, technologies, digitalization, services
15	china, education, chinese, students, learning, educational, chapter, mode, higher, school
16	sector, economic, financial, energy, companies, climate, development, global, new, services
17	platform, enterprises, leadership, medical, analytics, developers, firms, options, government, israel
18	regional, policy, entrepreneurial, countries, development, regions, analysis, high, clusters, results
19	industry, development, industrial, innovative, economic, technology, national, also, based, countries
20	learning, management, business, new, case, crowdsourcing, strategic, phase, firm, organization

In our comparative analysis, the LDA model from the Gensim library missed several nuanced topics that our advanced NLP approach successfully identified. Furthermore, it displays signs of contamination and lack of coherence in the keyword clusters. For example, topics such as 'sustainable, sustainability, environmental, development, economy, economic, policy, green, circular, food' (Topic 1) and 'tourism, technology, india, new, well, world, challenges, future, development, case' (Topic 6) show a wide range of terms that, while loosely related, do not necessarily constitute a coherent thematic narrative. This suggests a dilution of thematic focus, which is a common challenge with LDA, where keywords can be grouped together based on their statistical occurrence rather than their contextual or thematic relevance.

The introduction of transformer-based models, like those in this study, helped to address these limitations by improving interpretability: The major strength of our approach lies in its ability to generate more contextually rich and thematically coherent topics. Unlike LDA, which primarily relies on the frequency and distribution of words, our model integrates advanced NLP techniques like transformer-based algorithms, BERTopic, GPT-based categorization, and CorEx, enabling it to discern deeper contextual meanings and relationships between words. This results in topics that are not only richer in its variety are also seem more representative of the underlying content of the literature.

However, the field of NLP is evolving rapidly, requiring researchers ongoing updates and adaptations to these models to maintain their efficacy and relevance.

Development of an alternative topic modeling approach. Our study explores novel NLP developments to enhance bibliometric analysis in management research. We introduce an approach combining BERTopic, GPT-based categorization, and the semi-supervised CorEx model, allowing for a semi-automated process that incorporates human judgment in the topic-modeling process. The use of these techniques on a corpus of 2567 journal paper abstracts from innovation ecosystems has led to the identification of 27 distinct themes, showcasing an enhanced ability to process and interpret large sets of textual data in the management domain.

Our methodology allows for replicability. Given its machine-based nature, it enables consistent replication across diverse studies, ensuring reliable outcomes. This feature facilitates accurate comparisons across different topics or temporal periods and adds a new layer of depth and rigor to the analysis by directly addressing the interpretability and context-aware limitations often found in traditional topic modeling. Moreover, the extensibility of our approach also bears mention. While this study focused on the domain of business and innovation management, the foundational techniques employed are versatile. Researchers across disciplines can adapt and apply the methodology to their specific domains.

Bibliometric review of innovation ecosystems. Our methodology yields results that possess both interpretability and application potential. The outcomes of our bibliographic analysis provide multifaceted implications for the academic community. Our insight into 27 innovation ecosystem topic trends over time can act as a compass for researchers to understand the evolutionary trajectory within the domain. By recognizing emerging themes, academics can align their research focus with areas that promise both contemporary relevance and future potential. Furthermore, the analysis of authors and journals helps to navigate influential contributors and platforms in the field.

Thus, our paper demonstrates the utility of its novel topic modeling approach in conducting a bibliometric review of innovation ecosystems. This approach has enabled a detailed and scalable analysis of a large volume of text, providing a quantitative perspective on the various thematic structures within a broad body of scholarly work on innovation ecosystems. Ultimately, our application shows the approach's capability to handle extensive research domains efficiently and effectively.

Yet, despite the strengths of this NLP-based approach, we recognize that it cannot completely replace the depth and context provided by qualitative literature reviews.

Our topic modeling approach predominantly hinges on analyzing paper abstracts. Given the vast amount, breadth, and depth of this corpus, it underscores the importance and relevance of the subject matter in both academic and practical contexts. However, this methodology has an inherent risk. It is likely that some exceptional papers might have abstracts that don't adequately represent the depth or scope of the research. Conversely, papers of lesser significance might possess well-crafted abstracts that overstate their contributions. Such inconsistencies can influence the accuracy of the topic modeling, potentially leading to missed insights or overemphasis on certain areas. Therefore, while abstracts provide a concise overview, they may not always capture the full essence of a paper, posing a limitation to our approach.

Further limitations of our approach. It is essential to recognize our methodology's limiations, which also stems from its evolutionary nature. As with any scientific – and especially NLP-based – endeavor, continuous refinement based on feedback, testing, and new technological approaches is crucial. The intersection of machine learning and human judgment in our approach is not an end but a dynamic process that will inevitably evolve as we gather more insights from its applications and NLP methodologies advance in maturity.

One of the acknowledged limitations of this approach is the tension between unsupervised and semi-supervised topic modeling. When using semi-supervised methods, such as providing anchor words, there's a risk of guiding the analysis too strongly towards expected outcomes, thereby potentially overlooking novel or unexpected insights.

Another challenge is the difficulty in detecting new or emerging topics that have not been widely represented in existing literature. This limitation points to the need for further refinement of the methodology to enhance its ability to uncover and categorize such underrepresented themes.

While our research leveraged the capabilities of GPT-3.5 for topic modeling, it's worth noting that newer models, such as GPT-4, have since been released. These newer iterations often have improved training data, refined architectures, and enhanced capabilities. By utilizing GPT-4 or other advanced models, there's potential for achieving more granular and diverse topic categorizations. This could result in identifying nuanced sub-topics or emerging research areas that may not have been as prominently captured using GPT-3.5. Thus, while GPT-3.5 offered valuable insights, transitioning to more recent models could further elevate the quality and diversity of the topics identified.

Implications for academics. The implications of our methodology extend beyond the scope of our immediate study. We underscore the valuable role of bibliographic analyses in guiding and understanding the academic landscape. In essence, integrating automated NLP tools like BERTopic, Corex and GPT-3.5 into the literature review process can be a transformative shift. By enabling more efficient, comprehensive, and objective explorations, it equips scholars with deeper insights, setting the stage for more impactful research in the realm of business management.

While we deem our research methodology at the forefront of utilizing advanced NLP techniques, our previously described limitations need to be addressed. Awareness of these limitations allows us to approach the results with a balanced perspective. It also offers direction for enhancing the robustness and comprehensiveness of future similar studies.

As our methodology establishes its efficacy, several exciting directions for future research emerge. Exploring variations of our utilized algorithms, adapting GPT prompts to extract specific types of topics, and investigating the impact of different constraints are all avenues that can contribute to further refinement and optimization. Moreover, extending the methodology to handle diverse forms of textual data, such as full-text articles, offer promising prospects for advancing topic modeling across various domains.

These challenges underscore the importance of developing more intuitive and user-friendly models that can better align with human comprehension and judgment.

For instance, the process of integrating the findings from both BERTopic and GPT3.5 in our approach has not been an automated task. While our intention was to carefully consolidate insights, this integration can introduce challenges. Managing redundancies and handling overlapping topics can be complex and require careful scrutiny. Although offering a more holistic view, this fusion process can add layers of complexity that need additional interpretation and reconciliation.

Preprocessing is a crucial step in many NLP tasks. Our methodology involved preprocessing data for CorEx, but not for BERTopic. This differential treatment can introduce disparities in how each model interprets the data. While preprocessing often aids in refining the dataset, eliminating noise, and standardizing inputs, there's an implicit trade-off. The potential variations between raw and preprocessed data might yield different topic clusters or interpretations between the two utilized models.

The final phase of our research, analyzing the bibliometric results, involved a significant manual component. While this allowed for a human touch, ensuring a level of qualitative assessment, it also introduced potential biases and inconsistencies. Automating this process could produce more standardized, reproducible, and scalable results. Moreover, automation could facilitate the identification of intricate patterns, trends, or relationships within the data that might be challenging to discern manually. While our manual approach ensured a detailed review, integrating more automated and algorithm-driven methods could further refine the analysis, making it less complex to create.

Finally, our choice of using only Scopus abstracts has its own set of constraints. While Scopus is an expansive database, capturing a vast swath of academic literature, it doesn't necessarily include all relevant publications in a given field. By restricting our source, we might inadvertently miss critical insights or thematic evolutions present in non-Scopus literature.

CHAPTER THREE:

ECOSYSTEM GOVERNANCE DYNAMICS: NAVIGATING THE TENSION BETWEEN CENTRALIZATION AND DECENTRALIZATION IN ENVIRONMENTAL SUSTAINABILITY-CENTERED ECOSYSTEMS

ABSTRACT

Existing ecosystem governance literature has emphasized the advantages of centralized orchestration and a minimal viable ecosystem configuration. Engaging in a longitudinal case study of the development of an ecosystem to facilitate data exchange on carbon emissions, we, however, observed that this kind of governance design led to initial ecosystem rejection. Triggered by this surprising observation, we engaged in an inductive analysis of governance dynamics in sustainability-centered ecosystems. Our findings show how both fully centralized and fully decentralized governance approaches triggered important challenges that hampered the growth of the ecosystem. At the same time, we identify two important ecosystem governance principles – i.e., democratic, multi-body governance, and give-first mentality – which can help to address these challenges and navigate the tension between centralization and decentralization in sustainability-centered ecosystems.

Unpublished Working Paper. An adjusted version of this chapter has received Revise & Resubmit at *Long Range Planning* Weinhold, M., Faems, D., & Beitinger, G. (2024). Ecosystem governance dynamics: Navigating the tension between centralization and decentralization in sustainability-centered ecosystems. *Long Rangle Planning*. Ecosystem Governance Dynamics: Navigating the tension between Centralization and Decentralization in Sustainability-Centered Ecosystems

3.1. INTRODUCTION

Ecosystems have become an increasingly important research topic in the discussion of business strategies across organizations, providing a perspective that complements the established schools of thought on competition and cooperation (Adner, 2016; Autio, 2022; Thomas & Autio, 2019). Ecosystems refer to multilateral constellations of partners that jointly aim to develop a value proposition (Adner, 2012). Despite their increasing popularity, the failure rate of ecosystems is high, with studies estimating that up to 85% of ecosystems are terminated without reaching their anticipated objectives (Pidun et al., 2020; Reeves et al., 2019).

Ecosystem governance has been identified as a core driver of ecosystem success (Daymond et al., 2022; Jacobides et al., 2018; Pidun et al., 2020; Stonig & Müller-Stewens, 2019). Ecosystem governance refers to the systems and processes that are put in place to manage and regulate the interactions and activities within an ecosystem. This includes the rules, policies, and institutions that govern the behavior of all the actors within the ecosystem and ensure the ecosystem's attractiveness and sustainability (Adner, 2016; Jacobides et al., 2018). Existing ecosystem governance literature has identified centralized ecosystem leadership and minimal viable ecosystems as governance principles that can foster ecosystem success and stability. Ecosystem leadership refers to the ability of a central actor to effectively establish and manage interactions, relationships, partnerships, and alliances within an ecosystem (Adner & Kapoor, 2010; Daymond et al., 2022). A minimal viable ecosystem refers to a governance approach where the ecosystem orchestrator initially keeps the ecosystem as small as possible to minimize different types of ecosystem risks (Adner, 2021).

In parallel, the sustainability literature has started discussing the pressing need for firms to collaborate in ecosystems to address grand societal challenges such as climate change (Aarikka-Stenroos et al., 2021; Hopkinson et al., 2020; Papalexopoulos & Klein, 2022). In this literature, it is emphasized that interdisciplinary collaboration between organizations (Fuso

Nerini et al., 2019) can enable more sustainable business models (Aarikka-Stenroos et al., 2021; Rajala et al., 2018; Snihur & Bocken, 2022) to achieve “higher disclosure quality” and “help protect the entire planet and ecosystems as well” (Tettamanzi et al., 2022, p. 1). Despite the acknowledged importance of ecosystems in the sustainability literature, the particular topic of governance has remained relatively underexplored in this field.

However, when we engaged in a case study of a sustainability-focused ecosystem, we swiftly identified the need for a better understanding of ecosystem governance in this context. Analyzing the initiation of an ambitious ecosystem to facilitate data exchange on carbon emissions, we observed how the ecosystem initiator planned to embrace the governance principles of centralized orchestration and minimal viable configuration. However, we observed how these governance principles were fundamentally rejected by potential ecosystem partners. These observations made us question the generalizability of some of the governance principles that are highlighted in the extant ecosystem literature and highlighted the need for exploring the governance of ecosystems in the setting of sustainability-centered ecosystems.

To better understand the governance of sustainability-centered ecosystems, we engaged in an inductive and longitudinal analysis of our case study. We observed how the ecosystem, over time, shifted between more centralized and decentralized governance approaches to address the particular challenges that were faced. Based on our analysis, we demonstrate how fully centralized as well as fully decentralized governance approaches can hamper ecosystem growth. In addition, we inductively derive two principles, contributing to a more hybrid governance approach: (i) democratic, multi-body governance, and (ii) give-first mentality.

Based on our findings, we theorize ecosystem governance as a pendulum swing where different challenges force ecosystems to move away from fully centralized and decentralized approaches. Moreover, we highlight how sustainability-centric ecosystems might require a leadership approach where, instead of claiming a centralized orchestrator role, the potential leader gives away decision-making power and core assets to the ecosystem in order to establish the necessary conditions for ecosystem growth.

3.2. ECOSYSTEM GOVERNANCE: STATE-OF-THE-ART

3.2.1. Ecosystem Governance: Core Principles

Ecosystem governance is acknowledged as a critical driver of success for ecosystems. Adner (2006), for instance, postulates that effective ecosystem governance helps to ensure that the interests of all members of the ecosystem are taken into account and that the ecosystem is able to adapt and respond to changing circumstances. Looking at the emerging ecosystem literature, a number of governance principles have steadily emerged. Two core principles are (i) the relevance of a centralized ecosystem leader and (ii) the need for a minimal viable configuration.

Centralized Ecosystem Leadership. Centralized ecosystem leadership refers to an organization's ability to lead and manage interactions and relationships within an ecosystem effectively. This includes managing partnerships, strategic alliances, and other relationships with other companies, organizations, and stakeholders. Therefore, the ecosystem leader typically manages access to the ecosystem, controls the interdependencies (Adner & Kapoor, 2010; Daymond et al., 2022), and aligns the incentives across the participant. Masucci et al. (2020) point out that the leader often has a vested interest in leading the ecosystem by removing bottlenecks from the ecosystem to yield further efficiency and reinforce their own core business. It is also the ecosystem‘s leader who advocates the value proposition to other actors within and outside the ecosystem, similar to an institutional entrepreneur (Thomas & Ritala, 2022). The ecosystem leader addresses technological shortcomings in business ecosystems that typically hinder the integration of ecosystem partners of all sizes (Adner & Kapoor, 2010), often by providing (Thomas & Ritala, 2022) or even controlling (Masucci et al., 2020) core resources such as technology upon which complementors build to offer their products and services. In turn, and as a reward for their efforts, it is often the ecosystem leader who can capture “above-average rents” (Masucci et al., 2020, p. 1) or “reap the lion’s share of gains of the ecosystem” (Adner, 2016, p. 10) of ecosystem value creation for themselves.

It must be noted that there is a key challenge ecosystem leaders are confronted with: Ecosystems are often characterized by a lack of formal contracts and hierarchies between ecosystem participants and the ecosystem leader (Jacobides et al., 2018). This means that, in contrast to alliances, networks, and supply chain relationships, ecosystems often rely on voluntary inputs to co-produce value (Autio, 2022; Thomas & Autio, 2020). A lack of value creation within the ecosystem, however, may lead to ecosystem stagnation and, eventually, ecosystem failure (Adner & Kapoor, 2016). Addressing this challenge, Autio (2022) discusses the concept of ecosystem orchestration. They find that, in order to achieve voluntary inputs from ecosystem participants without formal hierarchies, the ecosystem orchestrator constantly needs to persuade others of the ecosystem’s overarching technological, economic, behavioral, and institutional offerings while at the same orchestrating collaboration between the stakeholders.

Minimal Viable Ecosystem Configuration. When companies engage in ecosystem-based strategies, they are confronted with unique risks that are absent in traditional business models. Two of the most prominent risks are co-creation risk, which pertains to the challenges of jointly creating value with partners, and adoption chain risk, which relates to the sequential adoption of innovations across the ecosystem (Adner, 2012). These risks introduce an added layer of complexity in ecosystem governance. Adner (2012) posits that to govern these risks effectively, companies should adopt a minimal viable ecosystem (MVE) approach. The MVE approach suggests that the orchestrator should deliberately limit the number of participating partners in the initial stages. This constraint serves as a mechanism to manage and mitigate co-creation and adoption-chain risks. Only after the ecosystem has demonstrated its viability in this constrained configuration, the orchestrator should consider expanding by integrating additional partners. This phased approach not only provides a structured pathway for ecosystem evolution but also ensures that foundational governance mechanisms are robustly tested before the ecosystem scales. Such an approach is crucial in ensuring that the inherent complexities of ecosystem-based strategies do not overwhelm the participating entities and that the ecosystem remains resilient and adaptive in the face of challenges (Adner, 2021).

3.2.2. The Role of Ecosystems for Environmental Sustainability

Sustainability challenges represent one context that has been highlighted as a very relevant context for ecosystems (Adner, 2016; Adner & Kapoor, 2010; Daymond et al., 2022; Jacobides et al., 2018). In the sustainability literature, a number of arguments are provided to explain why ecosystems are a vital strategy to address important sustainability topics. Moreover, this literature has identified challenges of ecosystem collaboration in this particular setting.

Need for Ecosystems to Address Sustainability Challenges. Business ecosystems are expected to play a vital role in helping firms align their individual business goals with sustainability targets (Aarikka-Stenroos et al., 2021; Papalexopoulos & Klein, 2022; Peltola et al., 2016). To achieve these sustainability targets, an intertwined network of companies (Korhonen et al., 2018) and "diverse stakeholders" (Peltola et al., 2016, p. 2) is needed. For instance, buyers and suppliers need to establish complex collaborative constellations to coordinate on emission targets to "reduce their emissions in line with climate science" (Farsan et al., 2018, p. 5).

In addition, alignment within ecosystems is expected to foster the creation of new interfirm business models that contribute positively to environmental sustainability (Rajala et al., 2018). Especially the context of closed-loop systems such as circular economy models (Korhonen et al., 2018; Teerikangas et al., 2021) with circular business models as an elementary building block (Hopkinson et al., 2020) offers an opportunity for a more sustainable future as they include the use of "new resources and include more service elements in their business model” (Teerikangas et al., 2021, p. 3). Aarikka-Stenroos et al. (2021) highlight that the notion of “ecosystem” has become crucial to transforming traditional economy linear models into circular ones to achieve resource efficiency and manage interdependencies among supply chain actors.

Finally, firms collaborating within ecosystems can achieve “higher disclosure quality” of environment-related data which should subsequently “help protect the entire planet and ecosystems as well” (Tettamanzi et al., 2022, p. 1). A crucial example is that collaboration in ecosystems can help to pinpoint responsibility for emissions in the supply chains (Farsan et al., 2018) through the exchange of data along supply chains. This, however, can only be achieved through collaboration from within business ecosystems, requiring firms to think beyond their own value chains (Erhard et al., 2019; Farsan et al., 2018; Labutong, 2018).

Challenges of Building Ecosystems for Sustainability. Examining the existing sustainability literature, we can identify two particular challenges that are likely to complicate ecosystem collaboration of sustainability challenges: (i) need for interdisciplinary collaboration and (ii) need for complex technological solutions.

Interdisciplinary collaboration is crucial for the development and maintenance of sustainability-centric ecosystems. The multifaceted nature of sustainability issues, which span environmental, social, and economic dimensions, often necessitates a holistic approach that draws upon diverse fields of knowledge (Fuso Nerini et al., 2019; Lu et al., 2021)

For instance, addressing climate change requires insights from environmental science, urban planning, economics, and social sciences. Each discipline offers unique perspectives and tools that, when integrated, can lead to more comprehensive and effective solutions. However, the pursuit of interdisciplinary collaboration in ecosystems is challenging (Bocken & Geradts, 2020). Differences in terminologies, methodologies, and epistemologies among disciplines can lead to misunderstandings and misalignments. In sum, while the imperatives of sustainability demand a breaking down of silos, the practicalities of achieving meaningful interdisciplinary collaboration remain a significant hurdle.

The use of novel technologies is often seen as an important enabler of sustainability-centric ecosystems (Papalexopoulos & Klein, 2022). In our case, for instance, blockchain was seen as a relevant technology to address the challenge of data exchange on carbon emissions. However, the implementation of these technologies is fraught with challenges. Novel technologies often lack standardized protocols, leading to potential interoperability issues (Biswas & Misra, 2021). In addition, the integration of these technologies requires significant capital investment, skilled human resources, and a paradigm shift in traditional business models, which many organizations might find challenging (Nowiński & Kozma, 2017). In other words, ecosystem partners might lack both the motivation and ability to embrace such technologies, hampering ecosystem success.

In sum, addressing sustainability challenges often requires the implementation of an ecosystem. Moreover, we know from existing literature that such ecosystems often entail interdisciplinary collaboration and technological innovation. However, the sustainability literature has remained relatively silent on the actual governance of such ecosystems. Moreover, as we engaged in our focal case, we quickly noticed how core governance principles, which have been emphasized in extant ecosystem governance literature (i.e., centralized orchestration and minima viable configuration), were rejected by potential ecosystem partners. Based on these initial observations, we decided to engage in an in-depth inductive analysis of the governance of sustainability-centered ecosystems. In the next section, we discuss our methodological approach to address this research objective.

3.3. METHODOLOGY

3.3.1. Case Setting

For this study, we explore the journey of a novel sustainability-centric ecosystem, "SUSTUM[3]". SUSTUM serves as an organizational forum between market-leading organizations for collaboration and technological innovation. Its members faced the challenge of presenting sustainability information to their customers, which required unprecedented data sharing beyond firm-centric boundaries. Therefore, the ecosystem participants aimed to innovate together towards common data sharing standards, technical interoperability, and provisioning of technology with the goal to enable interfirm sharing of data across complex supply chains while maintaining data confidentiality.

The studied ecosystem SUSTUM was initiated by its architect firm Factory-Worx[1]. Factory-Worx is a global multinational company that specializes in producing industrial goods for various sectors, such as transportation, energy, and factory automation equipment. The company caters to a diverse range of customers, including major corporations, public-private partnership corporations, and governments. These customers share a common concern about the environmental sustainability of the assets they purchase. To address these requirements, Factory-Worx started a sustainability initiative aimed at understanding, evaluating, and minimizing the product carbon footprints of its products manufactured worldwide.

Product carbon footprints generally consist of two major categories: The first category is “own manufacturing emissions”, also known as “Scope 1&2” emissions, such as those emissions from fuel burned in company vehicles and energy used in production. The second category is the “upstream and downstream supply chain emissions”, also referred to as “Scope 3” emissions. Those emissions arise from the production of the product and its subcomponents along the supply chain, while downstream emissions stem from the product's distribution to the

[3] In this study, all names have received pseudonyms to guarantee confidentiality.

customer, usage, and disposal (Damert et al., 2020). Data from the Carbon Disclose Project (CDP) and Lloyd et al. (2022) showed that supply chain emissions typically make up 75%-90% of total product emissions in the manufacturing industry. This finding was consistent with the results of Factory-Worx's own Life Cycle Assessments of their products' carbon footprints.

The novel sustainability initiative to assess Factory-Worx’s products' carbon footprints was launched for a few pilot products in Germany with the goal of measuring and reducing their carbon footprint. While Factory-Worx was able to assess their own factories’ “gate-to-gate” emissions relatively easily, it turned out to be much more difficult to gather information about emissions from their suppliers and sub-suppliers in the supply chain. To create carbon-neutral products, it is essential to focus on supply chain emissions. However, to select and award the most sustainable suppliers, it is paramount to receive reliable sustainability data from them. Factory-Worx recognized that obtaining precise information on a product's supply chain emissions is primarily a data-collection and data-comparison issue. To overcome the challenge of collecting data on emissions beyond their direct control, the company made significant efforts to investigate ways to facilitate the exchange of emission data with its suppliers and customers.

Searching for software solutions to share and request sustainability data with their supply chain partners, Factory-Worx concluded that none of the existing options available in the market met their requirements. In particular, none of the available solutions provided for synchronized data exchange, making it difficult to identify carbon hotspots in the supply chain and compare the data coming from the vast pool of suppliers.

As no solution was available, Factory-Worx decided to create a proprietary solution for evaluating the carbon footprint of their industrial products. Originally, the company intended to build a centralized platform-based ecosystem, requiring all its suppliers to input their specific emission contributions based on Factory-Worx’s requirements. However, as we will explain in our findings section, this initial design was challenged by the targeted partners, forcing Factory-Worx to move to an alternative governance approach. In April 2023, when we finalized data collection for this case, SUSTUM has been transformed into an active ecosystem, consisting of close to 20 member organizations coming from various industries, certification bodies, and academic institutions. The ecosystem produced tangible outcomes such as technology and whitepapers and has been featured in prominent places such as the World Economic Forum as a best practice example for sustainability-centered ecosystems.

3.3.2. Data Collection

For this study, we relied on three data collection pillars. The first pillar of data is built upon semi-structured interviews that Factory-Worx employees conducted prior to the initiation of the ecosystem in order to understand the design requirements and concerns of potential participants. The second pillar of data is built upon semi-structured interviews conducted post-initiation of the ecosystem by consultants. Third, we had access to Factory-Worx's primary data stemming from the period of ecosystem initiation.

It is important to highlight that the first and third authors of this paper had privileged access to the case as employees of Factory-Worx. As a Head of Product at Factory-Worx, the first author was in charge of the internal development of the software solution, of which parts Factory-Worx contributed later on to the SUSUM ecosystem. The third author is Senior Vice-President, Manufacturing and Head of Factory Digitalization, at Factory-Worx. The active involvement of the two authors allowed for a rich and deep understanding of the case. At the same time, this involvement triggers the risk of the researchers actively shaping the outcomes of their research question and increases the risk of biased interpretations. Several steps were taken to mitigate these risks. First, the inductive analysis of the case only started in May 2022. At this point, the SUSTUM ecosystem had already reached its final governance design, implying that the insights of the research could not have influenced the observed governance processes. Second, whereas the first author has been actively involved in conducting the second batch of interviews, we made sure to collect complementary data to maximize opportunities for triangulation and check for potential biases in the interview data.

First Batch of Interview Data. The first batch of interview data was collected when Factory-Worx decided to explore the mechanics of how collaboration with customers and suppliers should be designed to collaborate beyond their firm-centric boundaries. The interviews took place between February 2021 and November 2021 and were conducted by Factory-Worx in-house consultants for the purpose of exploring the business potential in the arena of sustainability data sharing. Interviewed organizations were potential actors within the ecosystem, i.e., manufacturing companies of all sizes from varying industries and certifiers. The positions of the interviewees varied: Interviews were conducted with top-level executives, sustainability officers, account managers, and procurement personnel to find out their motivation, but also concerns about joining the platform ecosystem. We received this qualitative semi-structured interview-based data from 23 companies from nine countries. 15 large multinational enterprises, five small and medium-sized enterprises, and three globally acting certifiers were interviewed (see Table 3.1). The interviews ranged from 15 – 45 minutes and were conducted by account managers of Factory-Worx, who held good relations with the interviewed companies. One interview with a large manufacturing company was terminated during the interview by the interviewed company due to confidentiality concerns. A standardized interview template was used for these interviews. After some introduction questions, the interviewer started by sharing the objective of the ecosystem and Factory-Worx' intended design of the ecosystem. The interviewees were asked if they would be willing to accept the envisioned ecosystem as a potential solution to their challenge. Further, they were asked to explain their motivation to join the ecosystem as well as the most important requirements for such an ecosystem. Interviewees were also asked to highlight conditions that would make it impossible for them to join the case ecosystem. A total of 11 hours of interviews were conducted and transcribed, resulting in 56 pages of text, with German interviews being translated into English.

Second Batch of Interview Data. The second batch of interview data was collected in the second half of 2022 (see Table 3.2). The interviews were initiated by Factory-Worx as it wanted to push for more technology solution providers within the ecosystem and further scale the number of ecosystem members. The goal of these interviews was to understand how the ecosystem could become more attractive for ecosystem participants, specifically solution providers. While these interviews were mainly conducted by external consultants, the first author participated in two interviews. By this time, the first author's research interest had already been triggered. Consequently, the interviews in which they participated served dual purposes. The author recognized the significance of these interviews both for their professional objectives and for their academic relevance.

Table 3.1. Overview of the first batch of interviews

#	Company Type	Headquarter	Interviewed On	Interview Duration	Role
1	Large sized Company	US	25.10.21	30min	Factory Head for 2 Factories
2	Large sized Company	Denmark	22.07.21	15min	Senior Manager "Mold Maintenance"
3	SME	Germany	01.04.21	45min	Head of Production & Health and Safety Manager
4	SME	Germany	29.03.21	40min	Director Quality and Environmental Management - responsible for 30 factories
5	SME	Germany	30.03.21	25min	Coordinator for EHS topics
6	Large sized Company	Germany	05.10.21	20min	Research associate with the Life Cycle Assessment Research Group

7	Large sized Company	Germany	26.03.21	25min	Lifecycle Management Manager - Metrics and Implementation
8	Large sized Company	UK	13.04.21	30min	Lead Sustainability
9	Large sized Company	Germany	15.04.21	30min	Sustainability Officer
10	Large sized Company	Switzerland	18.11.21	30min	Global Project Lead Climate and Energy
11	Large sized Company	Germany	03.05.21	30min	Head of RD Materials Research and Energy Conversion
12	Large sized Company	Germany	26.03.21	30min	ESG Officer
13	Large sized Company	US	26.05.21	30min	Lead global initiative SCM Sustainability
14	Large sized Company (Auditing	Germany	15.06.21	30min	Lead Advanced Manufacturing group
15	Large sized Company (Auditing	Germany	10.06.21	45min	Key Account Manager
16	Large sized Company (Auditing	Switzerland	13.10.21	30min	Product Manager
17	SME	Austria	01.09.21	30min	CIO
18	Large sized Company	Sweden	13.10.21	30min	Head of Communications
19	Large sized Company	Belgian	17.02.21	30min	Procurement Sustainability
20	Large sized Company	Germany	19.02.21	45min	Head Innovation Center
21	Large sized Company	Dublin	31.03.21	30min	Head of Strategic Purchasing
22	SME	Germany	28.05.21	15 min	Sustainability Officer
23	Large sized Company	US	28.05.21	Aborted during interview	Program Lead, Supplier Clean Energy Program

The interviewers collected data from actual ecosystem participants regarding their capabilities, motivations, actions, and satisfaction with the ecosystems. We received access to the complete set of interview data. Five representatives of ecosystem participants were interviewed through a semi-structured interview. All interviewees were asked to introduce themselves and their motivation to join the ecosystem briefly. Subsequently, the interviewees were asked to describe their activity within the case ecosystem and whether the ecosystem was on the right track. Lastly, the interviewees were asked about their vision for the ecosystem and how the ecosystem could be improved. In total, this second batch of interviews encompasses 21 pages of transcriptions and 6 hours of recordings.

Table 3.2. Overview of the second batch of interviews

	Company Type	**Interviewed On**	**Duration**	**Role**
1	NGO	13.05.22	30min	Fellow WEF, Director Factory-Worx
2	Large Enterprise	15.06.22	45min	Head of Decentralized Architecture
3	Large Enterprise	05.08.22	45min	VP Sustainability
4	NGO	12.05.22	60min	Automotive Alliance Head of Treasury, Director at Factory-Worx
5	SME	13.07.22	60min	Co-Founder

Field-level access to data. Lastly, we had access to SUSTUM's internal digital document storage from January 2022 through September 2022 (see Table 3.3). This includes access to original documents such as emails and PowerPoint presentations. The data collection contains more than ten letters regarding the association's formal foundation and its operation principles, including minutes of foundation meetings and legal documents. Our data also entails multiple discussion papers of the board of directors regarding their assessment of the state of the association. Furthermore, the document storage entails communication presentations that were to be shown to potential new members and shown at exhibitions and fairs. We analyzed multiple positioning papers regarding the working group's direction and more than 100 email conversations between ecosystem participants.

Table 3.3. Archival data collection

Type	**Quantity**	**Date**
Mail Conversations	109	14.04.2023 - 04.10.2022
LinkedIn Posts	4	10.09.2022 - 19.04.2023

Factory-Worx Internal presentations	12	20.04.2022 - 17.11.2021
SUSTUM Founding Documents	15	01.03.2020 - 15.04.2021
SUSTUM operational documents (PDF, presentations, working group documents)	11	01.03.2020 - 30.09.2022

3.3.3. Data Analysis

For our inductive analysis, we conducted several iterations where we gradually moved from case-specific observations to more abstract insights on the governance dynamics of sustainability-centered ecosystems (Brattström & Faems, 2020). As a starting point, we identified critical events that have shaped the journey of the SUSTUM ecosystem. Based on this initial step, we composed a 13-page case report, providing a description of major events from initiation through the evolution of SUSTUM from the beginning of 2020 to the end of 2022.

Once we had a comprehensive understanding of the ecosystem's journey, our next endeavor was to craft a data structure encapsulating the governance nuances of our case ecosystem, SUSTUM (see Table 3.4). Our foundation for this endeavor lay in the pillars of our collected interview and field data as described before. Through rigid examination of these data sources, we discerned notable features and characteristics unique to SUSTUM.

An evident pattern began to surface during our analysis: the governance of SUSTUM was not static but evolutionary. This realization allowed us to compartmentalize our data into evolutionary stages that closely mirrored SUSTUM's governance journey: (1) Phase 1 - Centralized Approach, (2) Phase 2 - Decentralized Approach, and (3) Phase 3 - Hybrid Approach. Applying the traditional inductive coding approach (Gioia et al., 2013), we identified for each stage core characteristics as well as core challenges (see Figure 3.1). As a final step of our analysis, we confronted the emerging findings with extant insights on ecosystem governance and sustainability-centric ecosystems to generate a more abstract understanding of the core implications of our findings.

Table 3.4. Supporting evidence for First Order concepts

Characteristics of SUSTUM changed governance over time			
Phase	**Feature**	**Theme**	**Supporting evidence**
Phase 1 Centralized Approach	Characteristics	Centralized infrastructure with own solution at the core	"With this initiative, we want to build a shared approach to enable effective climate action. Connecting to this network will allow companies of all sizes to exchange emissions data without the need to invest in infrastructure themselves" (Management presentation from 2020) "Purpose driven monetarization: SUSTUM provides our company a platform for goods and services based on and secured by transparent standards" (Project Document from Nov 2020)
		One firm, aiming to initiate and dominate a "beyond own firm" approach	"Based on standards of the platform ecosystem, Factory-Worx can offer products to access the network and leverage production data in its software suite" (Project Document from Oct 2020) "I have to admit we were interested in all the data that would have been able to collect. But we synthesized major learnings and realized quickly that nobody would collaborate with us. Control of the ecosystem would not have led to its success" (Manager of Factory-Worx in a mail)
	Challenges	Numerous partners are needed for legitimacy reasons	"The key issue for me is how to build a large enough ecosystem to increase transparency across the industry. We have to manage to motivate the industry to join in. A real ecosystem push is needed, until eventually there will be a pull and the necessary players will join in on their own" (Head of innovation in 2020)

			“What I am questioning: Wouldn’t it be hard to find companies to take part? But also I am convinced, there are companies. But we need to activate them. We really need not just a bunch, but a lot of participating companies to make this a credible success” (Business Development Manager in an interview in 2020)
		Smell of one firm makes ecosystem look like marketing tool	“The Ecosystem must not be misused as a marketing channel for products” (Founding member in a mail) “Let me say it clearly, we do not want to become dependent on a provider” (Interviewed Manager of potential ecosystem partner in 2021)
		Centralization of confidential data leads to ecosystem rejection	“What if we share data that would be used against us- We must make sure that the data is actually comparable. How can we control who sees the data?” (Product Manager of an interviewed supplier company in April 2021) “Yes, definitely data sovereignty is important. Who owns the data, and who makes the rules? It has to be a collaborative effort, almost altruistic, but by the industry for the industry“ (Manager of an interviewed potential ecosystem partner in 2021)
Phase 2 Decentralized Approach	Characteristics	Decentralized data storage as opposed to centralized infrastructure	“The focus is on calculating, exchanging, reducing, and offsetting carbon footprints along the supply chain through decentralized trust technology and self-sovereign data management” (Whitepaper, 2023) “SUSTUM builds on the technology of decentralized trust to avoid high cost, maintain data sovereignty and enable quick expansion with trustworthy information that can be shared along the supply chain with verifiable credentials” (Whitepaper, WEF, 2023)

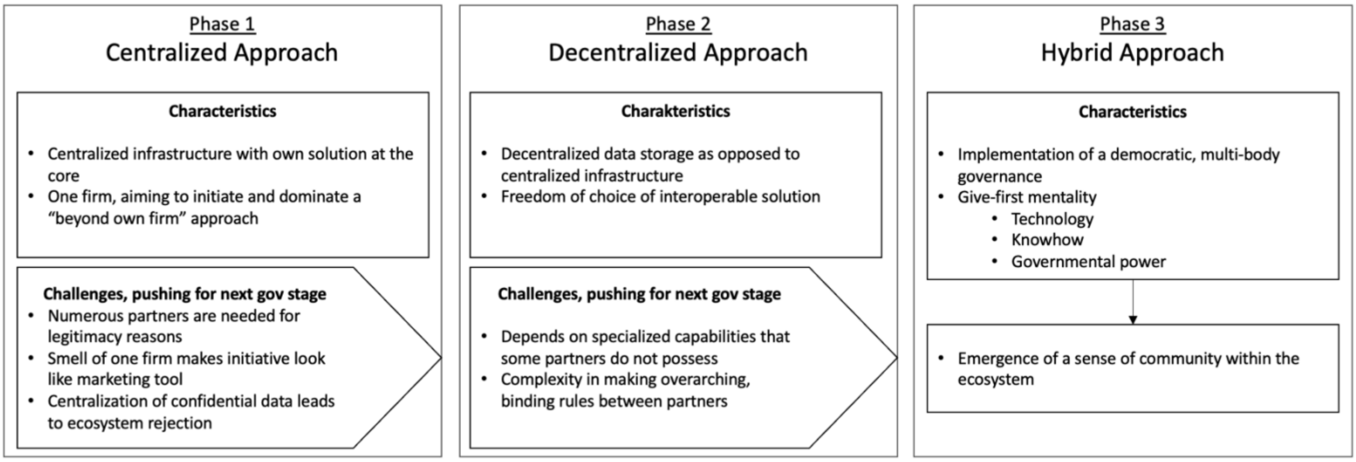
		Freedom of choice of interoperable solution	“I like the idea that this is not a single-company thing. In no case should the ecosystem be a ‘one company solution’, I don't think that would be accepted if it comes to a larger scale” (Innovation Manager of an interviewed potential ecosystem partner in 2021) “Instead of taking a share from the pie, a big-enough new pie will be baked” (Manager of Factory-Worx in a mail in 2020)
	Challenges	Depends on specialized capabilities that some partners do not possess	“We cannot drive such an initiative alone, it is technologically way too complex for us and too far from our core business. We are not google or SAP, we have other tasks” (Interview with a Head of Production & Health and Safety Manager in 2021) “The solution must be supported by technology that we can reuse. We can't get this backbone up and running on our own. We are grateful to be able to fall back on this technology” (Solution provider in mail in 2022)
		Complexity in making overarching, binding rules between partners	“How do we manage to come from A to B without getting lost in the diversity of opinions?” (SUSTUM MoM Document from 2021) “I have to be honest, sometimes it all still feels a bit like a party with new friends who have arranged to go hiking, but everyone has a bit of a different destination in mind” (Manager of an interviewed Factory-Worx Manager in June 2021)
Phase 3 Hybrid Approach	Characteristics	Give-first mentality • Technology • Knowhow • Governmental power	“Furthermore, to ensure constant improvement and ideal adaptation to user requirements, the adherence to principles of open-source software development should be essential to SUSTUM’s work whenever possible” (Internal SUSTUM positioning paper from 2022)

			“Therefore, the association is also founded and thus galvanically separated from the software: SUSTUM serves the altruistic goals and our software serves the business interests of Factory-Worx” (Manager via Mail in August 2023)
		Democratic, multi-body governance	“I would like to welcome you to the founding meeting of our association SUSTUM (…) I would take over the chair of our founding meeting today (…) We want to found an association with the name SUSTUM. The association should therefore apply for recognition as a non-profit organization (…) I would like to invite you to discuss this draft Articles of Association (…) As chairman of the meeting, I call the vote by acclamation. (…) Thus the association SUSTUM is now founded, it is given the articles of association displayed on our screens and the represented members of the association are the founding members of SUSTUM” (MoM of Foundign meeting of SUSTUM from 2022) “The probation comes for me when the association has to speak for the first time against the activities of members because of its statutes. This will happen, we will critically review the activities of members as a commonality” (Founding member of the association in 2022)
		Emergence of a sense of community within the ecosystem	“At SUSTUM, four different working groups are working on the realization of the project to make the world a better place” (Quote from the homepage from March 2023) “SUSTUM is the a global platform for our research body and joint knowledge” (Quote from a member’s press statement from June 2022)

3.4. FINDINGS

In this section, we describe how the SUSTUM ecosystem adjusted its governance design over time, facing particular challenges and concerns from ecosystem partners. Moreover, we explain how balancing the needs for centralization and decentralization finally resulted in a hybrid governance approach that is characterized by what we call ‘Democratic, multi-body governance’ and ‘Give-first-mentality’. Figure 3.1 shows the three governance stages, capturing the evolution of the SUSTUM ecosystem.

Figure 3.1. Different governance stages in the SUSTUM ecosystem



3.4.1. Phase 1: Centralized Approach

(Q3 2021 to Q1 2022)

Core Characteristic Centralized Infrastructure with its Own Solution at the Core. Factory-Worx took the initiative to invest in and develop a novel digital infrastructure for sustainability-related data exchange. This centralized approach was supposed to eliminate fragmented and incompatible systems across the supply chain actors. Unifying various data formats and facilitating data exchange was meant to harmonize and aggregate sustainability and carbon-emission data provided by suppliers, customers and auditors.

> *"Our goal is to bring the global economy behind a common approach to create transparency, set standards, and enable effective climate action by ensuring an auditable, secure, efficient, and standards-based data exchange" (Project goal statement from April 2020)*

Factory-Worx's platform was anticipated to evolve into a comprehensive industry-wide solution that other industry firms could also use, acting as a digital infrastructural layer for sustainability. The firm further planned to allow the integration of third-party services in the form of micro apps. Nonetheless, the centralized control and management of the platform were exclusively retained by Factory-Worx, emphasizing the centrality of its solution within the infrastructure.

Core Characteristic One firm aiming to Initiate and Dominate a "Beyond Own Firm" Approach. The intention to retain unequivocal ownership underlined a commitment to providing focused leadership. Factory-Worx understood that impactful change towards more exchange of data required directed and unwavering leadership to navigate the complexities of industry dynamics. They intended to control the terms under which other entities could participate. The objective was to enforce transparency and standardization, and control the platform to moderate its growth.

> *Originally, we did think about steering the ecosystem's direction, that could have been a very efficient approach (Manager of Factory-Worx in a mail in 2020)*

By doing so, Factory-Worx aimed to facilitate an environment that fosters collaboration, driving the industry toward more cohesive and integrated sustainability data exchange. However, Factory-Worx's intention to retain unequivocal ownership and directional control over the platform was also about steering the industry with a firm and visionary hand. The directive authority was imperative to ensure that the platform's evolution and growth were aligned with Factory-Worx's strategic objectives.

Core Challenge Numerous Partners are needed for Legitimacy Reasons. Factory-Worx conducted extensive consultations with several prospective member companies to evaluate the feasibility of their proposed solution. The interviewed organizations confirmed the critical need to address supply chain transparency issues and the lack of software solutions that would ease data exchange of carbon emissions based on standardized protocols. However, a couple of challenges with Factory-Worx's original approach were brought to light. First, the success of the data-sharing platform ecosystem would be contingent upon the active participation of a sufficiently large number of organizations. Potential users, however, expressed reluctance to invest resources in a platform that lacked legitimacy due to an insufficient number of participants.

> *There needs to be a scale achieved, I mentioned it earlier, a larger number of partners are needed to participate in our data-sharing approach" (Head of sustainability interviewed in 2021)*

Some organizations expressed the need for a "green industry movement" to be established to inspire users to adopt and legitimize the platform.

Core Challenge Smell of one Firm makes Ecosystem look like Marketing Tool. Organizations were hesitant to join a platform ecosystem that relied on a single technology provider who facilitated the exchange of sensitive sustainability data exclusively. The interviewed firms emphasized that the platform ecosystem by any means must not be used to collect, analyze potentially, and market sensitive ESG information while entrapping participants within a centralized solution. This was underpinned by comments indicating a "smell" of the one firm dominating the other actors. This smell made the ecosystem look like a tool to market Factory-Worx's solutions instead of supporting the above-mentioned green industry movement.

> *"You should remove the smell of your company to be accepted as a neutral authority" (Manager of an interviewed supplier company in June 2021)*

***Core Challenge Centralization of confidential data leads to ecosystem rejection*.** Finally, potential members raised concerns about centralizing the data exchange. The utilization of a centralized infrastructure raised questions about who had access to perform operations such as data reading, writing, and updating. Interviewees emphasized that confidential sustainability data could be potentially misused by competitors if they became publicly available, leaked, or sold.

> *What if we share data that would be used against us? How can we control who sees it? Product Manager of an interviewed supplier company in April 2021)*

Consequently, companies were unwilling to share sensitive sustainability data via a centralized infrastructure. Data-sharing firms needed to retain ownership and have complete control over their data, even after sharing it with downstream customers.

3.4.2. Phase 2: Decentralized Approach

(Q2 2022 to Q3 2022)

Core Characteristic Decentralized Data Storage as opposed to Centralized Infrastructure. To secure the trust and cooperation of the numerous partners needed for the data-sharing ecosystem to function, Factory-Worx recognized the need for a different governance approach. Instead of relying exclusively on its own technical carbon tracking solution, the company acknowledged the importance of allowing a coopetive ecosystem of multiple solution providers with their individual software solutions that utilized common data-sharing interfaces to ensure interoperability. To address concerns about data centralization, Factory-Worx promoted a model that empowered participants to select their preferred solution provider and ensured that shared data was not centrally stored or controlled by any entity. Instead, an already existing, publicly funded, and open blockchain network was employed to distribute data between participants and safeguard data ownership.

> *The trustworthiness of the shared information is ensured by the innovative distributed ledger technology (DLT), by means of which cryptographic certificates can be created and exchanged. This verifies the data provided and thus enables the trustworthy aggregation of a carbon footprint across the entire supply chain - without the participating companies having to disclose strategically relevant information, for example about their supply chains. (Description of SUSTUM Working group 3 in an internal document)*

The fundamental idea was that data shared between supply chain partners would be exchanged solely on a peer-to-peer basis between the data sender and receiver, using their software solution of choice. Blockchain technology cryptographically ensured that only authenticated and accurate data could be exchanged, connecting the software solutions with one another.

Core Characteristic Freedom of Choice of Interoperable Solution After proposing this framework publicly in industry networks, Factory-Worx was able to motivate two startups and one large enterprise to become solution providers that followed this design and developed interoperable solutions for the purpose of exchanging data with Factory-Worx. Moreover, Factory-Worx received confirmation from several firms to participate and employ these solutions in a pilot for exchanging sustainability-related data. The solution providers and the users established communication channels and working groups to meet and discuss progress regularly. These groups ensured that the singular interests of members could be addressed and satisfied. One working group, for instance, discussed and aligned requirements on common data standards between the prospect solutions.

> *"I like the idea that this is not a single-company thing. In no case should the ecosystem be a 'one company solution'" (Innovation Manager of an interviewed potential ecosystem partner in 2021)*

Encouraged by this momentum, Factory-Worx developed a functional software solution for assessing the carbon footprint of their products. This software could collect "gate-to-gate" emissions from the factory shopfloor, request carbon data from suppliers based on pre-defined data standards, and then share the carbon content of their manufactured products with customers using the decentralized infrastructure. Once a supplier received a data request from Factory-Worx's software, the party could choose the software of their choice to answer the data request. In sum, whereas Factory-Worx initially envisioned a fully centralized solution, where it would be an exclusive provider of the core technological platform and would fully control this platform, it now proposed a fully decentralized governance solution, where different providers could offer their software solution, and participants kept full control of their data based on decentralized blockchain infrastructure.

Core challenge Depends on specialized Capabilities that some Partners do not possess. Unexpectedly, it became clear that other potential solution providers, who initially expressed their willingness to become interoperable with Factory-Worx and followed the proposed solution design, lacked both the technical proficiency and the sustainability expertise needed to develop their own software solutions.

> *We cannot drive such an initiative alone; it is technologically way too complex for us and too far from our core business. We are not Google or SAP, we have other tasks"(Interview with a Head of Production & Health and Safety Manager in 2021)*

In this way, Factory-Worx's solution ended up being the only available solution in the ecosystem to assess, request, and share carbon data within the supply chain. This, however, hindered the effective implementation of the envisioned decentralized governance design, which was based on the interoperability of multiple solutions that relied on common data-sharing interfaces, as well as the decentralization of shared data to address concerns regarding data centralization. The unavailability of alternative solutions, therefore, significantly hampered the ecosystem's ability to ensure data integrity and created a dependency on Factory-Worx's software solution.

This dependence on a single solution raised concerns from users about the reliability and trustworthiness of the entire system. At the same time, it increased the risk of a single point of failure, as any malfunction or security breach in the Factory-Worx solution could potentially compromise the integrity of the entire ecosystem. Thus, despite Factory-Worx's ambition for a more decentralized, open, and interoperable design, SUSTUM now faced a situation where a decentralized infrastructure was established, but the ecosystem was still dominated by one single partner because of the lack of alternative technological solutions which were required to establish a coopetive ecosystem. Because of this lack of alternative solutions, users expressed fear and concerns about Factory-Worx’s potential to dominate sensitive information flows, monetize complementors' data, and manipulate rules of collaboration or use.

Core challenge Complexity in making overarching, Binding Rules between Partners. Another concern of participants was that, given the focus on decentralization, the ecosystem faced a lack of binding, overarching agreements to safeguard software implementation principles and guide collaboration practices between partners.

We need an organizer independent of business interests
(Manager of an interviewed supplier company in June 2021)

In the current situation, the ecosystem lacked a governance mechanism to agree on binding rules, nor were there any agreed regulations in place. Users reportedly worried about whether the ecosystem was intended to be a forum to market ideas or products from a single firm. An overarching aligned governance framework was missing to provide stability and contractually reassure users of the ecosystem's actual intention. The ecosystem, one manager remembered, felt like a "Company of new friends, starting to go hiking without a target in mind”.

In sum, SUSTUM had shifted to a decentralized governance approach in response to the concerns of potential participants with the initial centralized approach. However, this radical shift triggered new challenges that were hampering the growth potential of the ecosystem.

3.4.3. Phase 3: Hybrid Approach

(Q3 2022 to Q4 2022)

Democratic, Multi-Body Association to address Lack of Orchestration. Ecosystem partners realized that the rather loose network asked for more structure and leadership. To address this issue, the existing ecosystem partners decided to initiate a non-profit association, which would become the core governance body of the ecosystem. The association was dedicated to advancing the use of technology to reduce carbon emissions within the industrial supply chain. Its primary objective was to champion the transition towards more sustainable and eco-friendly industrial practices by leveraging innovative technological solutions. The decision to establish a non-profit association was motivated by the idea of separating independent ecosystem governance from business interests as much as possible.

> *We were told in a mail the non-profit status helped to get support and funding from international organizations and the government, further increasing adoption.*
> *(Manager of Factory-Worx via Mail in 2023)*

Leadership in the association was established by the democratic voting principle, for which the vehicle of an association provided the right tools. The association's board consisted of representatives from multiple participating firms.

> *Thus, the association SUSTUM is now founded, it is given the articles of association displayed on our screens and the represented members of the association are the founding members of SUSTUM"*
> *(MoM of Founding meeting of SUSTUM from 2022)*

A chairperson, a deputy chairperson, a treasurer, a secretary, and spokespersons for the working groups were appointed to represent the ecosystem and organize the decision-making processes. The governance structure was designed to ensure a fair and transparent decision-making process that allowed all members to be democratically represented. It also implied that a partner such as Factory-Worx, which contributed significant resources to the ecosystem, agreed to be underrepresented in this governance body, demonstrating its intention of not dominating the ecosystem.

By establishing the association as a legal entity, it was empowered to communicate with external institutions like banks, think tanks, and non-governmental organizations, which helped to build trust in the ecosystem's legitimacy. The non-profit character of the organization also reassured external institutions that the association was not driven by commercial interests but rather by the common goal of promoting sustainable practices.

Give-first Mentality to address the Challenge of lack of Coopetition. To address the lack of solution providers within the ecosystem, Factory-Worx decided to donate core technology to the SUSTUM ecosystem. This technology allowed other solution providers within the ecosystem to develop their own interoperable data-sharing solutions based on their needs and thus participate in the data-sharing infrastructure. Factory-Worx provided crucial but technologically complex libraries and modules from their data-sharing solution, which were then put under a SUSTUM inner-sourcing license. This donation formed the initial tech stack of SUSTUM. A dedicated SUSTUM working group became the new maintainer of the donated and open-sourced libraries. The working group identified and discussed participants' requirements on data sovereignty and confidentiality.

This donation by Factory-Worx sparked a "give-first" mentality amongst SUSTUM's members. Members gave specialized capabilities to support members who were lacking those, trusting that their donation would serve the ecosystem as a whole.

> *"We have to share our knowledge with all members of the association"*
> *(Statutes of the association)*

Another manifestation of this mentality is that the investment of ecosystem members in the ecosystem was not limited to pure software. To fulfill the ecosystem's purpose of enabling the sharing of sustainability data, research was conducted by its members in a total of four working groups. Research projects included, for instance, determining which data file formats would be used between the members and how the data formats would be interpreted.

Emergence of a Sense of Community. The emergence of a communal sense became evident as members of SUSTUM directed significant resources, time, and effort toward collaborative research projects. The attribution of research findings to SUSTUM, along with the omnipresent presence of its logos and designs, was pivotal in establishing a cohesive identity, portraying SUSTUM as a unified collective with shared goals and interests.

The endorsement of the joint slogan

“A race WE must win”

(Quote from SUSTUM homepage from February 2023) signified a critical juncture, marking the birth of a community. It symbolized more than a mere union of organizations; it represented a fusion of aspirations and a common ethos. This harmonious unity was cultivated through joint research initiatives and reciprocal investments, forging a shared vision that evolved continuously. Regular member meetings facilitated the orchestration of collective ideas. The unified promotion of ventures across various media platforms reinforced mutual ownership and a shared identity.

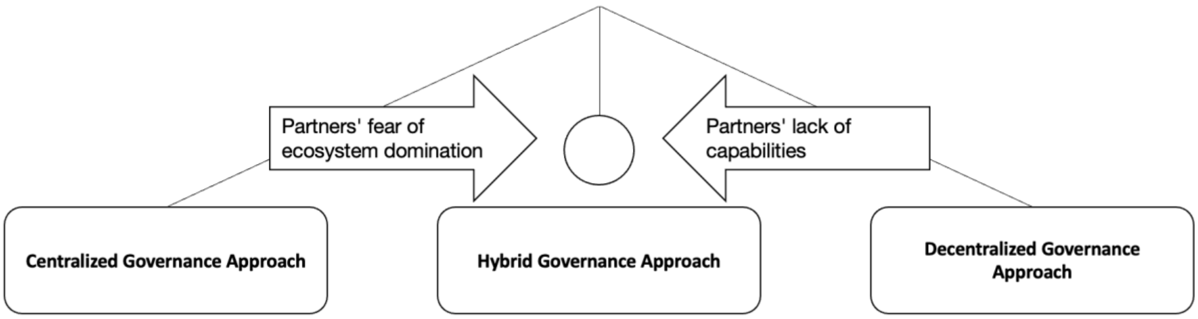
This conscious cultivation of collaborative synergy emerged as a fundamental element of the ecosystem, embodying the diverse shared experiences and cooperative ventures relished by the participants within the ecosystem.

3.5. DISCUSSION AND CONCLUSION

3.5.1. The Ecosystem Governance Pendulum

We have described how the governance approach of the SUSTUM ecosystem has shifted over time from a highly centralized approach to a fully decentralized and finally stabilizing in a more hybrid governance structure. As illustrated in Figure 3.2, we depict this kind of dynamics as a pendulum swing where opposing forces seem to push the governance approach away from fully centralized and decentralized approaches toward a hybrid approach.

Figure 3.2. The ecosystem governance pendulum



Centralized ecosystem orchestration by one single firm has been described as a viable approach for ecosystems, encompassing several governance advantages (Adner, 2016; Adner & Kapoor, 2010; Autio, 2022). Initially, Factory-Worx wanted to apply this approach to the SUSTUM ecosystem. However, they quickly found out that the motivation of potential partners to join such a centralized ecosystem was rather low. Among the potential partners, there was a substantial fear that such centralized ecosystem approach would lead to ecosystem domination by one single firm, which was seen as an unacceptable risk in this particular setting. For instance, participants feared being dominated, controlled, and locked in a proprietary system. Moreover, participants were concerned that they would not have the ability to partake in decisions or that their views would not be considered when key governmental decisions were made. Given this fear of ecosystem domination by one single firm, the SUSTUM ecosystem was pushed toward a more decentralized governance design.

By distributing data storage and decision-making authority to multiple actors and minimizing focal technological choke points, decentralized governance allows for greater inclusivity and diversity of perspectives (Tapscott & Tapscott, 2016; Zwitter & Hazenberg, 2020). However, our findings illuminate that a push toward decentralization triggers practical challenges related to participants' capabilities. From a technological perspective, decentralized governance requires a level of technical expertise that participants often do not possess. From an organizational perspective, it can be difficult to ensure that decision-making processes are understandable, transparent, and equitable. In our case, these capability challenges generated a strong force to move away from a fully decentralized approach.

In the end, the tension between the forces of decentralization and centralization led to the emergence of hybrid governance, which combined elements of both centralized and decentralized models. In this hybrid approach, the decision-making authority was distributed among multiple actors within a democratically elected board of directors. However, the association was a centralized authority ensuring efficiency and effectiveness. This helped to ensure that key decisions were made in a coordinated and efficient manner while also allowing for greater flexibility and autonomy among individual participants. The effectiveness of "semi-decentralization" has already been suggested by studies on the governance of cryptocurrency ecosystems (Chen, Pereira, et al., 2020; Zachariadis et al., 2019). Our findings illuminate that, in the context of sustainability-centered ecosystems, such a hybrid governance approach is likely to be the final result of a dynamic shift between more centralized and decentralized approaches.

3.5.2. Theoretical Implications

In this study, we contribute to the growing body of literature on ecosystem governance in two specific ways. First, our findings highlight the inherently dynamic nature of ecosystem governance. Based on our findings, we propose that, in sustainability-centric ecosystems, different challenges are likely to generate a pendulum swing between more centralized and decentralized governance approaches over time. This implies that, when scholars want to study ecosystem governance, it will be important to move away from research approaches that only provide a snapshot of the governance approach. Instead, longitudinal approaches that allow for measuring and capturing the evaluation of ecosystem governance over time seem to be crucial.

Second, our findings contribute to a richer understanding of the notion of ecosystem leadership. In most studies, ecosystem leadership is associated with a central orchestrator that takes the lead in managing and governing the ecosystem. In exchange, this central orchestrator will often generate above-average returns. Our data, however, suggest that, in particular circumstances, the predestined leader of the ecosystem might need to intentionally downplay its own role. In our case, Factory-Worx realized the need to establish a democratic governance body at the core of the ecosystem to address the participants' fear of ecosystem domination. In addition, Factory-Worx gave away strategically important technologies to the ecosystem to address the capabilities problems that were associated with the decentralized infrastructure. These strategic actions reduced the individual rent-generating capacity of Factory-Worx but increased the motivation and ability of other partners to contribute to the ecosystem. In sum, ecosystem leadership should not be automatically associated with one company taking a centralized position in the ecosystem. Ecosystem leadership might also imply taking a more modest approach to generate the boundary conditions for the ecosystem to flourish.

3.5.3. Managerial Implications

This study offers several key insights for managers seeking to develop and govern sustainability-centered ecosystems. First, we point to the potential relevance of an alternative ecosystem leadership style, where instead of claiming a centralized leadership role, the potential orchestrator relinquishes their dominant position. While centralized governance might look attractive in terms of individual rent-generating opportunities, too much domination by one company can breed distrust, which hampers ecosystem growth. In this respect, democratic multi-body governance with shared decision-making and the cultivation of a give-first mentality seems viable alternatives to secure partner buy-in in sustainability-centric ecosystems.

Our findings also clearly highlight that ecosystem managers should remain flexible and adapt governance approaches over time to address emerging challenges. Imposing full centralization or decentralization risks ecosystem instability or rejection. Overall, an empowering, adaptive leadership style coupled with inclusive governance and voluntary contributions appears vital for ecosystem success.

3.5.4. Limitations and Future Research

While providing important insights, this study has some limitations that suggest avenues for future research. First, although our single case study approach allowed for a deep and rich understanding of ecosystem governance dynamics in one single setting, we need additional research to understand the boundary conditions of our findings better. For instance, in our setting, the implementation of democratic, multi-body governance helped to address the need for a central orchestration entity without generating the fear of domination by one single partner. However, examples are present where such a multi-body governance approach did not work. When Facebook launched its Libra ecosystem to build a digital currency payment system, it also installed a multi-body governance system, involving multiple partners of the ecosystem. Nevertheless, this multi-body governance system could not change the perception of the public that Facebook would dominate this ecosystem, triggering a substantial regulatory backlash. This example illustrates that we need a deeper understanding of when and how this kind of alternative orchestration mechanism can be effective.

Our study also highlights the potential for research that delves deeper into the dynamics of the tension between value generation and value appropriation in ecosystems. To address the capability challenges of decentralization, the SUSTUM ecosystem established a give-first mentality, where ecosystem partners were willing to hand out technologies, information, and know-how to the ecosystem to ensure value creation. Given the timeframe of this study, it was too early to study the implications of this give-first mentality, on appropriate-later dynamics. The question of who will benefit most from this ecosystem remains open at this moment. Research, studying the long-term value appropriation implications of early-stage governance mechanisms that are aimed at stimulating ecosystem value creation, could therefore be very interesting.

CHAPTER FOUR:

UNDERSTANDING CHALLENGES AND INTRICACIES SURROUNDING THE SCALABILITY OF BLOCKCHAIN APPLICATIONS WITHIN LARGE INDUSTRIAL CORPORATIONS

ABSTRACT

Based on an in-depth analysis of two projects in a large industrial corporate, we inductively derive three organizational challenges, explaining why blockchain applications struggle to scale. Our findings contribute to the growing management literature on blockchain in different ways. First, whereas current management literature has positioned blockchain as a substitute for existing interfirm governance structures and as a third-party replacement, our study suggests that blockchain is rather a complementing governance structure. In particular, our findings highlight that traditional collaborative governance frameworks are required before blockchain-based governance can fully realize its promise. Second, as opposed to extant assumptions that blockchain can be seen as a trust-enabling silver bullet, we find that the successful implementation of blockchain actually relies on preexisting trust, provided by sound governance structures.

An adjusted version of this chapter was presented at *EURAM Conference 2022 Winterthur* Weinhold, M., & Faems, D. (2022). *Understanding Challenges and Intricacies surrounding the Scalability of Blockchain Applications within Large Industrial Corporations* EURAM Conference 2022 Winterthur Understanding Challenges and Intricacies surrounding the Scalability of Blockchain Applications within Large Industrial Corporations

4.1. INTRODUCTION

Blockchain is a digital technology that refers to a group of software programs that run on tamper-proof and certified records of transactions between participants. Both practitioners (Deloitte, 2020; PwC, 2021) and academics (Hellwig et al., 2020; Lumineau, Schilke, et al., 2020) highlight the potential of blockchain to fundamentally transform how business interact. Blockchain's "architecture of trust" (Schmeiss et al., 2019, p. 126; Werbach, 2018) has the potential to mitigate opportunistic behavior (Saberi et al., 2019) and even disrupt existing organizational collaboration capabilities (Mahtab et al., 2019; van Rijmenam et al., 2017). Lumineau, Wang, et al. (2020, p. 3) therefore argue that "blockchains may be considered a critical turning point in organizing collaborations and can be viewed as a type of governance mechanism that is distinct from traditional contractual and relational governance."

Despite the business potential of blockchain, transitioning corporate blockchain initiatives from a prototype to a scalable commercial activity appears to be difficult. A recent PwC (2021) study, for instance, indicates that only 10 to 25% of all corporate blockchain projects make it from a technical proof of concept to a productive application. This raises the fundamental question of why blockchain projects struggle to scale up to full-fledged company operations. Lumineau et al. (2020, p.20) therefore conclude that: "Further study is needed to find out if factors, such as management support, government intervention, and industry trends, affect the decision-making process, whether to transform existing business models to BT-based applications or not."

The core purpose of this paper is therefore to explore how and why blockchain projects applications fail to scale in a corporate setting. Whereas most management research on blockchain is conceptual and/or relies on publicly available case studies, we leverage our access to two different blockchain-based projects that were started by the same multinational corporation. In both projects, a technologically viable solution was developed, indicating the technological potential and feasibility of the blockchain-based application. However, both projects failed to evolve into a viable commercial activity for the corporation. In this way, this empirical setting provided a unique opportunity to study the scaling challenges of blockchain projects in a corporate setting. Moreover, whereas extant management research on blockchain-designed projects has focused on a limited set of blockchain applications (i.e., cryptocurrency), we rely on in-depth observations on two projects in a different setting (i.e., industrial supply chain management), where the introduction of this technology is still nascent.

Based on an inductive analysis of both projects, we identify three organizational challenges, explaining why blockchain-based companies struggle to scale: (i) Lack of suitable governance structure (ii) Lack of trust to successfully execute blockchain, and (iii) Lack of knowledge and skills. Jointly, our findings add to the emerging management literature on blockchain in two important ways. First, whereas recent management literature has positioned blockchain as a substitute for existing interfirm governance structures and third parties (Smits & Hulstijn, 2020), our findings point to blockchain as a complementing governance framework. In particular, we argue that traditional collaborative governance institutions still need to be implemented before blockchain-based governance can fully realize its potential. Finally, as opposed to popular belief that blockchain can be seen as a trust-enabling silver bullet, we find that the successful implementation of blockchain actually relies on preexisting trust, provided by sound governing structures.

4.2. BLOCKCHAIN IN MANAGEMENT RESEARCH

In this section, we first illustrate the most crucial ideas of blockchain and decentralization. Subsequently, we describe the current debate on blockchain in management and organizational literature.

4.2.1. Blockchain Technology in a Business Setting

Blockchain is a concept used to describe cryptographically protected software programs that authenticate and produce tamper-resistant records of verified, chained transactions between network users (Bennett, 2021; Treiblmaier, 2019). Blockchains always include append-only, time-stamped transactional data on some type of distributed database or ledger for value exchange and shared record-keeping, independent of their precise implementation form (Lacity & Van Hoek, 2021).

Blockchain-based applications, at their core, do not rely on centralized governance by delegating decision-making authority to the participants. The entities that connect to the blockchain are known as nodes in most blockchain variations. They usually save a full copy of all blockchain transactions, establishing a distributed database. This database of transactions is commonly referred to as "the single source of truth" and is one of the most important aspects of blockchain applications between usually non-trusting entities. Peers from each blockchain node (also known as an organization) approve or validate blocks of transactions and form consensus, indicating a multi-actor agreement. A consensus is created when enough peers agree on the results of a transaction, and the block is added to the blockchain. This also implies that transactions that contradict the consensus, such as those that have been tampered with or run on an out-of-date ledger, will be rejected. As the consensus is replicated across all entities, vulnerable spots are eliminated, resulting in a more secure system at the expense of efficiency. To sign transactions and manage read and write access, asymmetric key cryptography and elliptic curve encryption are utilized. Data based on the ledger, or agreements based on the data, cannot be readily altered since all other entities in the network regularly check their data sets with recent modifications and confirm if the majority of peers agree (Chen, Pereira, et al., 2020; Gorkhali et al., 2020).

Some blockchains also support machine-based automation, labeled as “Smart Contracts," in which the output of a transaction is linked to conditions embedded in digital code (Iansiti & Karim, 2017). The concept of Smart Contracts was introduced and developed by Nick Szabo (Ganne, 2018). Smart Contracts are also stored in the blockchain ledger. Its execution is triggered based on pre-defined conditions by the nodes within the network. All nodes that execute the Smart Contract jointly agree on a common set of conditions and rules. To successfully record the smart contract on the blockchain, all nodes must derive the same results from the execution. The most common Blockchain variants for Smart Contracts are Ethereum and Hyperledger Fabric[5] (Alkhudary et al., 2020; Buterin, 2015; Casino et al., 2019). As stated in Bennett (2021): "Smart contracts are neither smart nor contracts. But they are an excellent vehicle for automating processes that span corporate boundaries. It is an 'advanced' form of process automation that goes across corporate boundaries”

The general differentiating factors for blockchain are mostly focused on the management and permissions levels. Anyone can join and participate in public blockchains. In addition, all parties have the ability to carry out actions and activate contracts. Most public blockchains allow an endless number of participants to engage anonymously with one another without the need for a central intermediary (Gaur & Gaiha, 2020). To reach consensus in public blockchains, high-energy-consuming techniques known as "proof-of-work" are frequently used. The most well-known blockchain implementations, such as Bitcoin and Ethereum, are public permissionless blockchains. Because of its incentive mechanism, bitcoin consumes a lot of energy. Virtually anybody can join in and begin mining bitcoin to win a pay-out. Private blockchain (also known as permissioned blockchain) is not run for the purpose of earning cryptocurrencies. They exist because members want to utilize their nodes to help safeguard the network. A whitelist of permitted actors specifies each person's specific authorization for

[5] In Hyperledger Fabric, Smart Contracts are called Chaincode

network operations. In other words, in a private network, a small number of known parties engage with one another. Well-known implementations of private blockchains are the TradeLense or FoodTrust chain (Hellwig et al., 2020; Morabito, 2017; Treiblmaier, 2019). Private blockchain networks are mainly used in firm-to-firm commercial contexts, which are often strictly controlled and regulated (Alkhudary et al., 2020; Gaur & Gaiha, 2020; Lacity & Van Hoek, 2021). Anti-money laundering and know-your-customer requirements, for example, may not apply to public blockchains. Because this paper focuses on the use of blockchain in an industrial setting, we use the term "blockchain" to refer to permissioned blockchain throughout the rest of the paper.

4.2.2. Blockchain in Management Research: State-of-the-art

In management literature, blockchain is typically framed as an emerging topic that may be a turning point in how economic, social, political, and business organizations collaborate in the future. Business leaders therefore need to strategically prepare for blcokchain-driven disruption (Iansiti & Karim, 2017; Lacity & Van Hoek, 2021). Fenwick, Kaal, and Vermeulen (2018) state that digital technologies have already disrupted centralized, hierarchical corporate organizations by facilitating platforms. This process of disruption is said to only continue as new blockchain-based technologies emerge. However, because of its disruptive potential, regulators need to be even more attentive to these changes and their economic and social effects (Fenwick et al., 2018).

One of the most prominent a reasons for blockchain's expected unique way of joint value creation is its potential to create new ecosystems (Malhotra et al., 2021) and business models (Holotiuk et al., 2017). In a blockchain-based ecosystem, companies bring their distinctive capabilities and skillsets to create joint value. At the same time, blockchain takes out “disintermediates and trusted third parties” (Lacity & Van Hoek, 2021).

Lumineau and his co-authors have set the scene for this discussion, highlighting that blockchain-based collaboration may be distinct from traditional contractual and relational governance (Lumineau, Schilke, et al., 2020); Lumineau, Wang, et al. (2020). This is mainly attributed to blockchain's inherent "trust mechanism." Blockchains are described as a form of inherent controlling power that is expected by their users to be immutable (Meijer & Ubacht, 2018) and replace an authority that was previously assigned to established governance bodies, which are deemed untrustworthy. For this reason, blockchain is often called a "trustless" technology, acting as a "confidence machine" that increases confidence in the operations of a computational system based on its underlying governance structure (De Filippi et al., 2020).

Focusing on the advantages of blockchain, one often stated use case is that it might make collaboration and coordination between corporate partners more simple and effective (Gaur & Gaiha, 2020). For instance, blockchain might aid in the handling of recent crises such as Covid-19 and Climate Change by guaranteeing that all parties follow the rules. Furthermore, in these crisis conditions, blockchain is supposed to speed up operational processes like locating and onboarding of new suppliers and detecting and preventing fraud (Bennett, 2021; Lumineau et al., 2021; Malhotra et al., 2021). In this way, businesses that use blockchain technology may have a competitive edge over competitors that are slow to adopt this new technology (Felin and Lakhani, 2018).

Whereas this existing research stream has provided interesting insights in the potential advantages and opportunities of blockhain, it remains relatively silent on the potential implementation challenges and risks. Next to technological obstacles, such as transaction speed, storage capacity, and high energy consumption issues, organizational challenges might complicate the implementation of blockchain projects. (Alkhudary et al., 2020, p. 19). Vogel et al. (2019) therefore highlight the need for "further research that is focusing on practical evidence and implementation". In a similar vein, Saberi et al. (2019) recommend investigating post-implementation success and failure factors to understand the broader applicability of blockchain to business purposes. In this paper, we aim to contribute to addressing this research gap by exploring the implementation challenges in two blockhain-based projects in a single corporate setting.

4.3. METHODOLOGY

4.3.1. Case Selection

The goal of this study is to understand why and how blockchain projects fail to scale in a corporate setting. As our research topic is relatively under-investigated and the constructs of interest were not known in advance, we rely on an inductive approach as it allows us to gain contextual and in-depth know-how and induce novel theory from the data (Eisenhardt, 1989; Pelz, 2021).

For this study, we rely on two blockchain-based projects – i.e. NewProtect and NewCert[6] - that were situated in one of the largest industrial corporations in the world. From a technological perspective, both projects were successful as in both cases a working blockchain solution has been generated. From a business perspective, however, both projects can be considered as failures as both initiatives failed to transition into a mature business activity for the corporation. By collecting primary data on two blockchain projects that failed to produce business value, we address the criticism that: "Not many cases of actual implementations of blockchain technologies are known, and even less that have successfully developed in beyond a pilot stage. Moreover, those applications that have been published are likely to be successful ones." (Smits & Hulstijn, 2020, p. 6).

[6] In this study, all names have received pseudonyms to guarantee confidentiality

4.3.2. Data Collection and Analysis

For this study, we conducted 18 interviews with informants. Five of our informants had senior management roles, and 13 of our informants held senior expert positions. We interviewed four informants specifically to NewProtect. An additional six informants were interviewed, particularly on NewCert. All other informants were experts in the firm with broader, general management or engineering positions located in business units and corporate research functions. We engaged our informants in semi-structured interviews, inquiring about the their core expertise in the field of blockchain. All informants were asked about their position and relationship towards the projects to minimize bias. Whereas managers and experts related to our projects were specifically asked for their post-project recount and specific anecdotes of particular interest for our research project, we asked all other experts for their assessment of the blockchain projects described based on their wealth of experience. The interviewing process lasted a year, from June 2020 to June 2021. All of the interviews were conducted remotely and lasted between 60 and 100 minutes. A total of 21 hours of interviews were taped and transcribed resulting in 59 pages of text.

To analyze structure the vast qualitative data, we engaged in a set of first and second-order coding analyses (Gioia et al., 2013). As a first step, first-order notions were gathered and categorized as impediments and issues that our agents reported. An exemplary first-order notions is “Users (employees) lacked IT skills to fully grasp and use the blockchain system”. All first-order notions are related to at least one of the two trajectories. The variety of seven first-order hurdles was then rigorously scrutinized, and further condensed into second-order themes. All second-order themes contain at least one first-order notion with applicability on both trajectories, demonstrating broader applicability of the second order themes.

Interestingly, our data shows that while the first order notions played out in a logical sequence, they had to be classified within different second-order themes. For instance, “Blockchain governance structure was not designed for necessary post-hoc changes” has been categorized within “Lack of suitable governance structure”. The effect of the former notion, “Actors lost ‘inside trust’ in the system, which led to decreased participation”, however, is a first-order notion being classified within the second bucket “Lack of trust to successfully execute blockchain”. This is owed to the strong interconnection of trust and governance and will be further discussed in this paper. It was also surprising to us that all interviewees spoke with some excitement and generally positivity about blockchain technology with high expectations for the future, even though the cases were framed as unsuccessful and discontinued.

Our combination of first-order observations as expressed by informants and the broader second-order challenge themes allowed us to theorize on the obstacles of scaling blockchain projects. In the next section, we first provide a brief description of the observed projects. Subsequently, we discuss the core second-order challenges that emerged out of our inductive analysis. Finally, we reflect on the core implications of our findings.

4.4. FINDINGS: CASE DESCRIPTION

4.4.1. NewProtect

The goal of the NewProtect project was to upend a century-old business paradigm for asset insurance. Traditionally, policyholders and insurance companies have had two distinct responsibilities. NewProtect’s innovative concept was to combine these roles with more cost-effective insurance policies. The decentralized, Smart Contract-based solution was used to pilot activities that were previously performed by an insurance firm. The community’s rules and decisions were used to execute this Smart Contract. The scope of this application comprised onboarding new users to the system, payments, claim management, conflict resolution, and voting processes, as well as the entire digitalized user journey for asset insurance. The idea was that a united community of “insurance providers and takers” would vote on claim judgments in a fair and non-profit-oriented way. Furthermore, it was expected that the Smart Contract-based algorithm would be able to perform insurance activities more effectively than a human-operated system. Traditional trust in a single third party, the insurer, was shifted to the community as a collective of insurers making decision together.

Realizing the disruptive nature of this project, the NewProtect team opted to start with a small company segment and investigate the "learning by doing" process of decentralized insurance. The team searched for an intriguing product that could attract the attention of future stakeholders and participants. In the end, the team developed an intra-company insurance program for the company's employees. Workers' personal cellphones and tablets were specifically chosen as insurance assets to be safeguarded.

The NewProtect team viewed mobile phone insurance as a "safe zone" in the event that the insurance procedure failed due to unforeseeable occurrences or even fraud. The researchers hypothesized that in a community of firm employees, individuals would be less motivated to commit fraud due to the low value of mobile phones. The implementation team envisioned expanding the notion of decentralized risk-sharing to large-scale industrial assets if the decentralized decision-making logic, technological implementation, and general acceptability level proved effective.

NewProtect's operations were set up similar to those of traditional insurance companies. Policyholders paid a monthly fee into an insurance pool out of which the cost of repairing or replacing participants' devices would be paid. In addition, operational and IT costs were covered from this budget pool. One remarkable and unique feature was that excess money in the insurance pool was not retained and kept as profits but returned to those paying in. The official website read, "We have found a good balance between usability and security. Every transaction will be transparently documented. Once information is put into the system, it can no longer be changed due to the immutability of the blockchain".

Furthermore, a Smart Contract handled the process of handling claims alongside organizing community decisions to the approval of damage or loss. Randomly chosen members from the community were appointed as a jury once a user handed in a claim. The individual jury members were guided through a digital multi-step evaluation process where the jury decided whether the claim had been valid or not. If none of the reviewers objected to the compensation for the claim within five days after the request, the claim got approved and paid out. In the case that at least one of the reviewers objected to the claim in the initial round, a second voting round began with a new jury. The second randomly chosen review group was presented with the same information as the first group, plus the reason for denial from the first voting round. Now, the process followed the principle of a simple majority decision. A quorum of more than fifty percent within the second reviewer group approved or disapproved the claim. Using this methodology, the financial risk of damage or loss of a mobile phone was shared among all community participants from the pool. In this way, the implementation team created a system that took decision-making authority out of their hands from the start, allowing the community to decide on every part of the process. Voting was not only required for claiming decisions, but it was also necessary for modifications to how the system operated. There was just one non-consensus decision: the NewProtect team's only remaining authority was to shut down the system.

The technological implementation of this project was successful. The team was able to create a working blockchain-based system and a significant number of employees were willing to participate in this insurance system. Nevertheless, the application was pulled down after a 12-month period. The remaining money was returned to the participants and the system was never operational again, with no intentions for a second phase.

The decision to shut down the system was driven by a dramatic decrease in users during the first year, which was fostered by a number of anecdotes that started circulating within the company. A first anecdote refers to a situation in which a participant emailed the NewProtect implementation team, questioning why his/her claim was refused only seconds after submitting the claim. In the email, the user stated that (s)he submitted all documentation in the right format and that (s)he expected the legitimate claim to be recognized. However, the real reason for the denial was that the user had replaced the phone throughout the insurance term, resulting in a change of the phone identification number. Unfortunately, the user forgot to change the identification number in the system. Before holding a community vote on the claim, the Smart Contract always verified whether the applicant's identification number was saved in the system's database, ensuring that only users who have signed up for the service could use. As the new identification number was not kept in the system, the claim was denied by the Smart Contract. Given that the post-hoc adjustments were not possible in the blockchain-based application, there was no way to repair this mistake of the user. Even though the customer paid a monthly charge, the team was unable to amend or override the Smart Contract's refusal. While the Smart Contract performed as expected this inability to repair the mistake of the user was seen as a significant flaw of the system by many employees.

A second anecdote was related to the inability of Apple phone users to make claims at a particular point in time. When Apple phone customers wished to make new claims, the Smart Contract was programmed to retrieve the current market price of the iPhone from the official Apple website, which was used to determine the payout amount. However, because to an ongoing litigation with Samsung, the original iPhone's website pricing was removed, and the Smart Contract refused to accept a price from any other website that supplied a current market price. As a result, claims about Apple phones could not be resolved. Moreover, the NewProtect team could not reverse the conclusion because of the preprogrammed restrictions. The only option the team would have is to completely shut down the system.

Both anecdotes went viral in the case firm and received a lot of attention from the people. As a result, participants lost trust in the system and the user base has shrunk dramatically. Less paying users meant dwindling cash that eventually made it impossible to manage the payment of legitimate claims. The system had to be shut down in the end.

4.4.2. NewCert

This project took place in an industrial setting where heavy machinery assets for the energy and rail industries are created. Designing, manufacturing, engineering, commissioning, and decommissioning of industrial assets requires comprehensive documentation to demonstrate that compliance was not violated, either intentionally or accidentally. As a result, a substantial amount of paperwork needs to be passed back and forth between the actors that are engaged in creating and checking compliance documents. Traditionally, this means that the auditor and the auditee have to verify the paper records manually. Managing and overseeing physical papers for modifications, anomalies, and technically impossible additions was one of the most tiresome and error-prone responsibilities for operating staff. Furthermore, as the quantity of manufactured items, suppliers, and client orders expanded within a particular business unit, so did the amount of data inputs to be regulated and maintained. As a result, the risk of mistakes grew exponentially.

The objective of the NewCert project was to find out whether applying blockchain could ease compliance by introducing digital certificates attached to physical assets and harmonizing operations between stakeholders. Specifically, a first use case covered scenarios where owners of critical infrastructure assets must demonstrate their assets' conformity to customers and external bodies, such as supervisory authorities or insurance companies. Suppliers are obliged to require these certificates for safety-critical, externally procured components. Suppliers confirm with these certificates that an independent audit agency has been involved to test critical components for safety and that the asset meets all the needed safety requirements. Instead of piling up and sending manually signed paper documents, the new blockchain-based process created a process where confirmation of the asset compliance would be immutably saved on the blockchain and made accessible to all parties involved. In other words, the blockchain generated a digital twin of compliance, covering conformity checks of all steps in the planning, production, commissioning, testing, and acceptance phase:

The responsible team was able to design a functioning system, where users first entered the entire certification-relevant information and corresponding data in a simple web-frontend. Subsequently, a Smart Contract hashed the data from the backend to create a digital, encrypted time and content stamp, which was then communicated to all network participants. In this way, the certification results were ‘frozen in a non-editable manner and could no longer be detached from the chain. Furthermore, the rules based on which the digital certificates are generated and distributed could be reviewed by every party that reviewed the Smart Contract.

After the blockchain architecture was created, it was necessary to integrate the supplier’s legacy IT systems into the decentralized blockchain architecture. Suppliers should run and operate their own nodes to fully benefit from the blockchain’s features and be considered equal actors with equal rights. To avoid the impression of one-sided disadvantage or mistrust in the data, the tracking of data changes should not be managed centrally by one party. Instead, a decentralized data management approach was pursued. It was even envisioned that, in the future, contractual penalties for late delivery could be automatically calculated based on the new blockchain-based single source of truth, allowing for reduced transactional costs. Despite these potential advantages, interviewees highlighted that it was challenging to convince suppliers to join NewCert:

> *“The concept technically works very well. The problem was, frankly, that suppliers were not interested in the blockchain and its security. It is a blockchain managed by our company. Suppliers want a lean process above all else” (Project Manager)*

Some suppliers did participate and shared the required data. However, enthusiasm for generating their own blockchain node could not be generated. Blockchain was supposed to be a technology that fosters trust in the data on all sides. Yet, there was a fundamental difference between what stakeholders expected when implementing the new application and the actual outcome.

> *"We see the supplier's desire and interest in the pilot blockchain. (..) However, suppliers often still lack the vision, and so far, the supplier has seldom come up with their own ideas." (Head of Procurement)*

A NewCert manager told us that they tried to promote blockchain to the suppliers with the promise of non-modifiable, secure data if they ran their node. The answer the team got was

> *"We don't know what blockchain is, but If this improves the overall process, then we'll join your project" (Supplier). But it must not lead to additional efforts."*

Interestingly, one project team member told us that, when a supplier was asked to join the blockchain network, the supplier researched the background of blockchain and asked if the case company did not trust the supplier before. This supplier was puzzled that the team wanted to come up with a solution that enforced trust. A case company member remembered a conversation where the supplier stated

> *"We don't have a trust problem because we have a contract." (Supplier)*

After presenting the project to another supplier, their IT department sent a high invoice for "customer-specific adjustments in the IT system" to participate and operate their node. Another supplier, a large logistics company, showed no interest in joining the project at all. This company said they had established processes with all its customers that could not be changed because of one single project. If data was to be tapped, then their APIs and databases had to be used.

NewCert provided the technological baseline for an automated and secure digital certification process. However, because of suppliers' resistance to actively participate in this system as an actual blockchain node, the idea of establishing an ecosystem was given up and the remaining blockchain parts of the application were shut down. In the end, top management of the corporation concluded that the cost of operation was too high compared to the modest value provided by the blockchain.

4.5. FINDINGS: CASE ANALYSIS

In this section, we discuss the main challenges that emerged out of our inductive analysis. Figure 4.1 provides an overview of our data structure, whereas Table 4.1 provides interview quotes for the different first-order concepts.

Figure 4.1. Data Structure

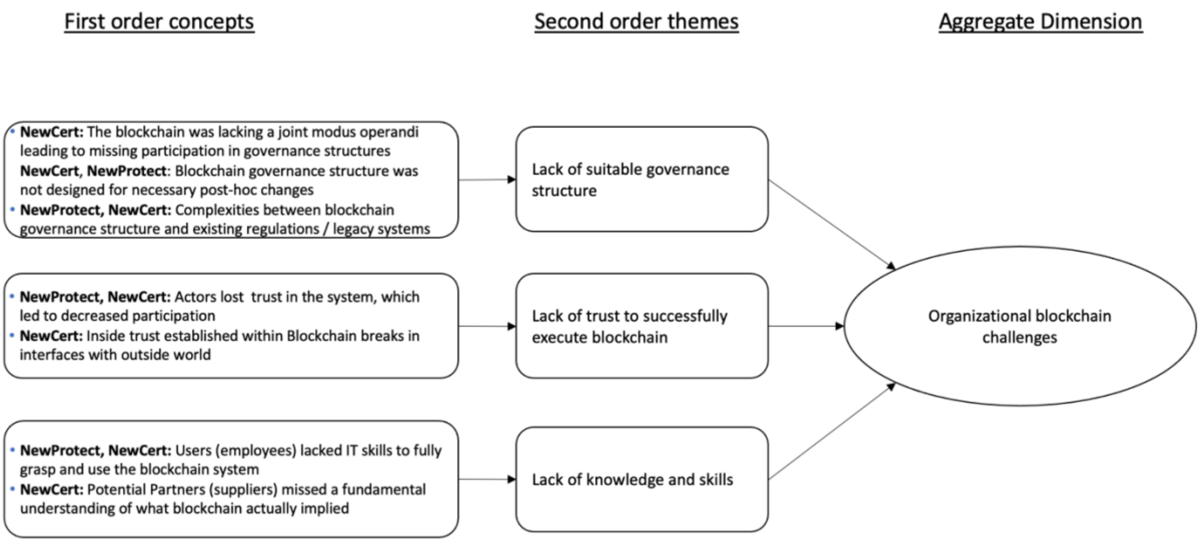


Table 4.1. Supporting evidence for First Order concepts

Second order theme	**First order concept**	**Further supporting evidence**	
		NewCert	**NewProtect**
Lack of suitable governance structure	*The blockchain was lacking a joint modus operandi leading to missing participation in governance structures*	“The head that thinks about the overall governance concept is missing”	“We didn't consider, who's going to host this, who's going to run this?”
		"Who programs, who owns the code in a truly decentralized solution? It is very complicated!“	
		“Nobody looks at the big picture. That is what's missing”	
		“Perhaps we are not yet able to properly convey the medium- and long-term potential here. Our suppliers often still lack the vision”	
	Blockchain governance structure was not designed for necessary post-hoc changes	"At the end of the day, though, we have to be honest. We deal with complex business relationships that have evolved over many years, and you need to have a good feel for your counterpart. My relationship with my supplier is more important to me than enforcing my rights at any price via a Smart Contract"	“The users were not amused when they found that they would not receive any money from the Smart Contract”
	Complexities between blockchain governance structure and existing regulations / legacy systems	"I would also like to emphasize again that as much as I see blockchain projects as a matter of the heart, we have to adhere to the valid legal regulations. Especially in the GDPR area, I see big problems here. The "right to be forgotten" does not seem to get along with blockchain, in my opinion. How do I enforce this consistently in a Blockchain network? Here, I am responsible as a manager, and in case of doubt, liable"	"It took ages until we got clearing from the responsible financial authority. Explaining the details to the authorities was indeed the most complex part."

		"The acceptance is not there yet. Companies have hundreds of people working in legal departs. Saying that blockchain is going to disrupt that. I don't know"	
Lack of trust to successfully execute blockchain	*Actors lost trust in the system, which led to decreased participation*	"We really struggled to convey the idea of transparent supply chain processes. I think we went to fast here for our partners (..) They don't want to publish everything they do to the whole world"	"Nevertheless, I believe that the two stories mentioned may have led to a loss of confidence in the system among the users - word gets around very quickly in these kind of cases"
		"The suppliers did not see value in this unprecedented transparency"	
	Inside trust established within Blockchain breaks in interfaces with outside world as "Outside trust"	"The movements of goods were visible on the blockchain, but not in the main system. In my opinion, it does not make sense to operate two systems in parallel"	"One single point of failure, of course, was the finance department which was managing the users' funds"
		"You have to be careful with the word 'trust' in front of a customer because if you promise trust, that could question if today there isn't there any trust? I remember a conversation where a business partner said, "We don't have a trust problem because we have a contract. And when we have a contract, we have a way of enforcing it."	
Lack of knowledge and skills	*Users (employees) lacked IT skills to fully grasp and use the blockchain system*	I was hoping for a user-friendly frontend, which I could use to map movements in a system on a simple system that maps the process. Simplicity was in the foreground for me. Whether our suppliers switch to another IT system or not was not really in the foreground for me."	"But I have to admit that the system was probably too complex for many users and they did not understand the trust aspects - and probably would have trusted the system in a non-decentralized way"
		"What is still not clear to me is to what extent blockchain smart contracts are now real contracts. I do not think they	It is frequently difficult for the user, particularly during the registration procedure. To log in,

		are, so we need contracts in any case. So, what's the point of the blockchain?”	the user's private key must be revoked. User management is also difficult for the user because the password cannot be reset by an administrator. There is no user management."
		"I expected that the blockchain would lead to processes becoming more cross-company. For example, who do I work with, what is my relationship with the business partner - even over a longer relationship - but at the same time, certain data can be exchanged more quickly.	"When it comes to designing Smart Contracts, Blockchain requires you to think at a very high level of abstraction in terms of business operations. You must grasp it and abstract - not everyone is capable of doing so"
		"In the end, however, we also have to say that our expectation of the blockchain was a simplification of the processes. Additional discussions about new responsibilities in the process, on top of the existing contracts, would rather complicate my work."	
	Potential Partners (suppliers) missed a fundamental understanding of what blockchain actually implied	" “The suppliers find the application, okay, but they are also very insistent that there is no additional work for them. Our suppliers are also corporate groups and do not have the direct option of changing their IT systems, which means we always generate extra work”	
		"We discovered that suppliers wanted to use blockchain technology to set themselves apart from the competition. Then there is the inconsistency. 'Yes, I want the benefit, but I don't want to participate, I don't want to invest, and I can't modify my procedures either,' we were told.	

4.5.1. Lack of Suitable Governance Structure

While the teams were able to implement technologically feasible and working solutions, they underestimated the governance complexity that blockchain projects brought along, which in turn resulted in the absence of a suitable governance structure to scale the project. Interviewees therefore stressed that, when onboarding actors, the main challenge was not situated at the technological level, but rather was related to the relational level:

> *"Blockchain turned out to be more of a relational conversation than a technical conversation" (Operations Manager)*

The governance complexities of blockchain projects manifested in several ways. First, we observed that governance issue hampered the necessary participation of stakeholders in the blockchain system. In the NewCert, for instance, external companies turned out to lack the motivation to act as active participants in the decentralized governance structure that the blockchain solution represented. Informants from our case company told us that, despite the potential advantages of decentralized blockchain governance, partners preferred to stay within their old processes and not invest becoming an active node in the blockchain system. Our data suggest that the stakeholders lacked the necessary commitment to embrace this alternative governance approach.

This lack of commitment was highly driven by skepticism about embracing a governance approach that would decide on important issues in an irreversible and preprogrammed manner. Especially NewCert informants noted that, for particular decision in supply chains, interpretation of the specific situation may be required and that exceptions to pre-specified agreements must be established. Complex business circumstances need situational awareness and context, which a Smart Contract cannot provide. Disputes are frequently resolved by a healthy shared personal relationship. Informants expressed that they would not want to rely on a Smart Contract, enforcing deterministic decisions at any price. Similar comments were made by informants from the NewProtect project, who expressed their dissatisfaction with the fact, that decisions made by the Smart Contract could not be ruled back post-hoc. In fact, the absence of a “Service Team” or “Complain Button” was seen as a fundamental flaw in the governance of the blockhain project.

Discussing the governance complexities associated with blockchain, interviewees expressed that privacy issues impeded the successful setup of the application as they stood in stark contrast to blockchain’s general radical transparency. Immutability and irreversibility are two critical attributes of blockchain technology. Once data or transactions are appended and accepted by a network, they can no longer be changed. Theoretically, data on a blockchain will stay there indefinitely. In blockchain applications, the data of registered users might no longer be under the company's control that contributed the data. Likewise, informants expressed that data audit trails may continue to exist on the blockchain even after dismantling the industrial facility. A third party, hosting the data on their node meant giving up control of the data. However, not being able to finally delete data from a system could mean a violation of European judicial privacy requirements. This illustrates that indirect stakeholders such as governmental bodies and regulatory authorities strongly influence the ability to adopt blockchain-based solutions. For example, it was not clear whether NewProtect had to be labeled as a finance product or not because the authorities could not form a conclusive opinion on the matter. If so, the product would have been subject to insurance tax and more complex insurance regulations. This would have significantly increased costs for users on the insurance rate. These uncertainties minimized applicability of the blockchain application for some users. Informants told us that, in case the blockchain environment fails to provide a clear legal framework, the willingness of actors to embrace blockchain-based solutions in an industrial setting is likely to remain low.

4.5.2. Lack of Trust to Successfully Execute Blockchain

This challenge revolves around the need for high levels of trust to successfully execute blockchain application. First, we highlight that trust in the blockchain system itself is needed to build and maintain engagement. Second, we observed how unanticipated problems with the blockchain applicaton can lead to a breakdown of trust.

Exemplary for the lost “inside trust” in the blockchain is a situation in NewProtect concerning a perceived (but actually correctly executed) malfunction of the application’s governance. As described earlier, within NewCert there was no way of questioning or even rolling back a decision, which led to lost trust in the application being able to responsibly manage its users claims. The rules by which the decisions are made within the network need to be known to all participants. If there is no mutual understanding of the rules and mechanism confusion, mistrust and disparticipation are the effect.

Another example of the complexities with trust in the blockchain system revolves around the core feature of radical transparency. A manager sketched an illustration of the network architecture he attempted to establish to illustrate the trust intricacy. The involved companies were identified as case company 'A,' supplier company 'B,' and logistics company 'C,'. All three organizations were required to be a part of the digitalization project of the incoming procurement and supply chain process. The manager elaborated on the challenge that case company 'A' was concerned that logistics company 'C' would discover that ‘A’ works with 'B..' As a result, he feared that 'C' would use that knowledge to approach employees of 'B' and directly offer their services, thereby sabotaging prices. He said that he had failed to eradicate this anxiety because he misjudged the difficulty of engaging partners to trust a system that builds fully on transparency. Not only felt it obscure to the partners to make their all data transparent, also the mechanics by which decisions would have been made felt opaque and not trustworthy..

> *"Still, the long-term benefit that there is a new transparency on both sides, and that penalties can be handled, or that the process could be improved and automated, the suppliers don't trust that" (Head of Procurement)*

The second manifestation of this challenge revolves complexities with blockchain trust and its interfaces with the outside world. In the NewCert project, certificates where shared by an external certifier. Informants told us that the trust the blockchain provides can only be as good as the certificate itself. A fraudulent certificate from the outside world on the blockchain, for instance, would hamper the actually provided "inside trust" in the system itself. In other words, the blockchain may not be able detect "corrupted" data as such. For instance, the proof of provenance for digital assets in NewCert could be verifiably tracked and traced in the blockchain network. In these scenarios, the reliability of the data is high, and blockchain can simplify collaboration and coordination. However, agents such as sensors, QR codes, cameras, and other input forms are needed to digitally capture what is happening in the physical world. Physical actions cannot be impacted by a blockchain-based setup. It was achievable in NewProtect to follow the flow of money on the settled and refused claims blockchain. It got more complex to understand in the NewCert project whether a procured part did arrive on-site in the right quality and quantity as claimed by the suppliers. There had to be an additional – physical – check to ensure that the information registered on the ledger was similar to the transferred physical assets. Furthermore, the blockchain application itself does not sit in a vacuum but is embedded in a greater environment with existing regulations and legislation in both the digital and physical realms. To maintain conformity with the embedding system's regulation, blockchain's interaction with the legacy processes could become a trust bottleneck with the power of potentially undermining blockchain's entire value proposition.

While the blockchain may be able to cater for data integrity, it is not suited well for improving data quality. Poor data quality on the system, however, destroys trust in the (blockchain-) system.

Here, we would like to mention another anecdote that was reported by an informant. The manager found that using blockchain might even jeopardize current trust. The manager said that companies must be cautious when offering increased 'trust' in front of customers, according to that manager. The implied message might be seen as a lack of trust in today's world. We were advised that implementing blockchain may call into question a long-standing partnership. (S)he said that when our example firm introduced blockchain to a supplier, this became a reality. According to the manager, the particular supplier was astonished that there had never been any conversation about trust.

4.5.3. Missing Know How

Missing competence has shown to be a major challenge for successfully implementing blockchain. Accordingly, we found two major manifestations of the challenge “Missing Know”. First, “Missing knowhow of users” and second “Missing know how of potential partners”.

Within NewProtect we observed that users lacked IT skills necessary to fully grasp and use the blockchain system. In turn, they struggled to understand the blockchain application’s outcomes and questioned decisions of the Smart Contract. The same held true for NewCert, where interviewed users had an exaggerated expectation that blockchain could automate cumbersome and manual error-prone tasks and simplify user interfaces. This expectation, however, differs from blockchain’s actual value proposition fundamentally. Blockchain was occasionally seen as a silver bullet to improve processes, rather than a tool that allows for decentralized, transparent decision making.

Other users anticipated that the blockchain application would serve as a streamlined, unified communication platform for internal and external stakeholders to share relevant data. What was given, however, was an application that was fundamentally different and reportedly more complicated than previous centralized programs. While consumers had few contacts with the blockchain application's underlying crypto mechanisms, there were several distinguishing specialties that caused confusion and eventually distrust. As a result, employees who had previously been trained in procurement, logistics, and other operations-related jobs suddenly needed to be knowledgeable in a variety of highly specialized IT sectors. For instance, users needed to comprehend how data is structured and how various digital databases could come to a consensus and resolve potential disputes.

> *"I must also say quite clearly that this exceeds the required knowledge of my employees in their normal working environment. The user always needs two core expertise: Blockchain expertise and subject matter expertise. You have to be an expert on both topics to be able to understand what happens, interpret and discuss" (Procurement Manager).*

The same held true for NewProtect, where the Smart Contracts' regulating behavior was not well understood. The blockchain application operated substantially differently from what the users expected. The Smart Contracts handled the mobile phone insurance claims in accordance with pre-defined community standards. Clearly, no management decision may override or modify the outcome of a decision. Some people, however, were taken aback by this. The fact that choices would be made automatically by a decentralized surprise and could not be modified later necessitated a lot of additional explanation from the project team. For instance, users still expected to have a focal point of contact that they could reach out to in case unexpected results occurred. However, the affected users reported remaining unclear about whom to reach out to in case of unexpected behavior triggered by the preprogrammed Smart Contract.

A lack of understanding applied not only to the users but also to the external partners who were expected to engage and participate much more. We observed partners' misunderstanding of the actual core advantages of blockchain as well as a lack of knowledge about what decentralization means, paired with an unwillingness to invest capacity and efforts in it. This led to a low rate of excitement as suppliers were not really sure of the advantages this innovation had brought. Reportedly, some suppliers did show a willingness to participate, but this was more related to the general openness for joint innovation projects and to use the engagement in the blockchain project as a marketing asset instead as an outcome of excitement to participate in decentralized, transparent governance system. Moreover, the suppliers very much insisted, that there cannot be additional efforts involved in engaging in the blockchain project, otherwise they would be forced to drop out.

4.6. DISCUSSION

In this section, we discuss the implications of our findings. We reflect whether blockchain truly establishes a trust architecture and if blockchain can be viewed as a new type of governance framework. Subsequently, we summarize implications for managers and highlight the core limitations of our research.

4.6.1. Is Blockchain an Architecture of Trust?

It is suggested in the literature that blockchain-designed projects may remove the need for trusted third parties and intermediaries (e.g. Meijer & Ubacht, 2018). Blockchain's trust architecture is said to potentially mitigate opportunistic behavior capabilities and even disrupt existing organizational collaboration capabilities (Lumineau, Wang, et al., 2020; Mahtab et al., 2019; Saberi et al., 2019; van Rijmenam et al., 2017).

We would want to provide some nuance to this conversation. In both our cases, the intention was present to shift from bilateral trust between actors to trust in a decentralized system, ensuring that all decisions are made by fair and pre-established business rules aligned with all actors. In theory, Iansiti and Karim (2017, p. 217) find that this decentralized trust architecture could potentially "protect assets and set organizational boundaries." However, we experienced that establishing a blockchain does not automatically establish confidence in the system. Participating actors, for example, must be certain that Smart Contracts invoke the unanimously agreed-upon rules. They must have faith in their implementation, which may be ensured by preceding conventional trust-providing governance methods like contracts between the actors, which also lay out the rules for situation of doubt and disagreement. Furthermore, the necessity for confidence in the actor, who establishes and drives the whole system, is sometimes overlooked, but it became evident in our examples. Dedicated trusted actors are required who push forward the blockchain system across actors, but who also act as a point of contact in case of uncertainty and doubt to avoid loss of trust in the blockchain system. Of course, these actors can also be the participating parties themselves, but full transparency on the rules and decision-making structure must always be ensured.

In this sense, we find that blockchain does not so much replace trust (Meijer & Ubacht, 2018), but rather relies on trustworthy actors that drives and secure the system. Once the system is established – and trusted – blockchain may further supplement and secure existing trust among actors by the means of its inherent technological features such as immutability and radical transparency. As a result, if traditional trust-providing institutions and trusted governance do not exist to build confidence in the blockchain system itself, the blockchain will likewise fail to develop trust.

Furthermore, it should be noted that Smart Contract-based, decentralized judgments may only be trusted in settings in which all parties involved have the technological and intellectual competence to participate. In other words, if actors do not operate their own nodes and get direct decision-making powers on the blockchain, they cannot confirm the correct execution of the Smart Contract and hence cannot trust it. Only by being part of the infrastructure, they can control what is written on the distributed ledger. Otherwise, individuals or groups with greater technological capabilities and more dominant hierarchical positions are likely to assert their authority or set rules and incentives to advance a specific objective or perspective (Brattström & Faems, 2020; Hambrick & Finkelstein, 1987). Ultimately, a one-sided or uneven blockchain implementation reverses the original idea of enabling trust and collaboration and gives the actors a strong power advantage to run their blockchain nodes over those who do not use these tools.

4.6.2. Can Blockchain be Seen as a Governmental Framework?

It is being discussed that blockchain-based collaboration may be distinct from traditional contractual and relational governance, for instance, because of blockchain's inherent "trust mechanism." (Lumineau, Schilke, et al., 2020); Lumineau, Wang, et al. (2020). However, we argue that blockchain does not install a governmental framework but actually relies on it. In other words, the blockchain as the technical implementation can only be the second step after a governance framework has been formulated that allows to actually use this technology.

To overcome environmental legal, regulatory, or other outside factors impeding the successful implementation of blockchain, it is crucial to have binding agreements in place before implementation that regulate how to deal with data on the blockchain. Another example is that, in contrast to a public blockchain network like bitcoin, where actors are mostly anonymous, data will be shared and distributed across all actors in private business blockchain networks. It is therefore required to rightly couple data entries and changes with the identity of each actor. In addition, new participants need to be vetted and approved by all stakeholders before they can access the network. This causes significant coordination efforts for which existing structural or relational governance mechanisms are needed as well.

Conventional contracts between stakeholders are needed to define the parties' rules and responsibilities where and when to enter data into the blockchain to enable the Smart Contracts to function as planned. Likewise, the terms by which the Smart Contract governs participants need to be agreed upon in a contract-like document. If this is not defined, the business partners may rely on a (pseudo) mutual understanding of how they are governed, which is dangerous and undermines a blockchain's project objective. In our case research, we have also shown that this framework needs to include statements, stipulating which decisions should and should not be taken over by the actors on the blockchain, and what happens in case of unplanned events and disputes. We therefore believe that blockchain governance needs a well-defined governance structure based on traditional contracts. If no governance framework is in place before adoption, the blockchain will be unable to establish one

4.6.3. Does Blockchain Lower Transactional Costs?

Blockchain is described as a technology that can reduce the costs of recording data on each potential transaction partner at the most fundamental level. At the same time, the identification of contingencies and pre-coded conditions in smart contracts allows for decreased negotiation because transactions can be automatically and autonomously carried out (Malhotra et al., 2021). One of the most prominent and commonly repeated reasons for blockchain's expected unique way of joint value creation and transaction cost reduction is its potential to strengthen existing and create new ecosystems while at the same time offering great potential for new business models (Holotiuk et al., 2017).

Management literature already recognized the necessity for companies to bring their specialized capabilities to a blockchain-based ecosystem (Felin & Lakhani, 2018). Our cases, however, demonstrate the complexity of blockchain-enabled value creation in developing ecosystems. There must be "in it" something for each ecosystem partner. Otherwise, the blockchain project may fail to scale and deliver value. It may be easy to convince external users to join a common innovation project without fully understanding the technology and its paradigms. However, the blockchain will not unfold its potential in these scenarios, mostly since the users will not expect different and potentially novel governance. To participate in the governance, actors will have to invest in technical know-how and infrastructure. In our cases, this was not always the case.

We also found cases with unclear regulatory settings that restricted blockchain's economic impact. Therefore, we strongly support Fenwick et al. (2018) who state that blockchain may only unfold its cost-cutting effect when regulators will be even more attentive to disruptive innovations and their economic and social effects within companies.

In our cases, blockchain reduced transaction costs for certain parties but not for others, and the initiatives were abandoned due to a lack of value for other key players. Subsequently, financial difficulties occurred as no viable business model that produced value for all members could be identified. As a result, although blockchain may reduce transaction costs for certain parties, it does not necessarily do so for others. This demonstrates the need for properly aligning incentives to guarantee that all participants share a similar interest in the blockchain's long-term viability and integrity.

Interestingly, previous research points out that blockchain's preprogrammed standardized and automated governance provides value to all participating actors (Schmeiss et al., 2019). However, we discovered that this is only true if there is a clear value proposition and demonstrable benefit for all involved parties. Implementing blockchain may provide one-sided commercial value through process automation or system integration as long as a clear value proposition for all participating parties is not identified. As a result, the expense of running a decentralized app for this purpose may exceed the value.

4.6.4. Managerial Implications

When it comes to blockchain for business, limited attention has been devoted to organizational constraints that have inhibited collaborative blockchain-based disruption and resulted in one-sided blockchain-enabled procedures in our examples. Our findings suggest that maintaining an experimental, yet also collaborative mentality is critical. Managers should continue to investigate blockchain-based technologies when new ones arise, however, only stringent training of stakeholders on the foundations of blockchain can address a lack of knowledge of decentralization and collaboration and hence ensure project success. Especially managers who sponsor and market the implementation of blockchain applications must have a thorough understanding of the technological impact on trust and decentralized governance. Users have to be extensively instructed and educated on the basics of blockchain in order to meaningfully interact with the program and its details. Employee training must include not just how to use the blockchain-based application correctly but also a broader perspective: what are the actual benefits of blockchain, how do the new concepts differ from what users are used to, and what does decentralize blockchain governance represent.

After examining several flaws in the use of blockchain technology in business, we recommend that practitioners, who want to reap the potential benefits of blockchain technology, to build on existing strategies to establish inter-organizational relationships via formal contract structures, which safeguard and coordinate the inter-organizational relationship (Faems, Janssens, Madhok, & Van Looy, 2008). To retain administrative control over multi-dimensional blockchain for commercial contexts, practitioners should not rely solely on blockchains. Blockchain may be incorporated as a technology implementation to supplement traditional governance procedures once a robust cooperation and governance framework has been established. We further advise managers at this point not to use blockchain technology in mission-critical business scenarios due to ambiguous regulatory environments and other external considerations. In the blockchain sector, successful business cases are uncommon and have yet to be discovered. As a result, businesses must work together to determine how decentralization might enhance their interfirm business relationships.

Starting blockchain initiatives without forming partnerships is difficult, and it has shown to be a barrier to the creation of new blockchain-enabled ecosystems in our situations. As a result, the trust advantage will continue to be one-sided. Using blockchain within a single company is unlikely to yield significant decentralization-enabled benefits. Thus, managers must establish a unified vision from all business partners with independently formulated value propositions in order to make the transition from a blockchain-enabled process to a collaborative blockchain disruption. The benefits of blockchain architectures are best realized in multi-party processes centered on verified data that can be shared across organizational lines. This is where blockchain excels in enabling confidence in certain areas, such as data integrity. In this context, our managerial advice is to change the focus from whether the technology will work to how its features can support existing collaboration and governance structures by enabling multi-party workflows around distributed, validated data, all while being supported by the right organizational framework.

4.6.5. Limitations and Future Research

There are a few shortcomings in our research that need to be addressed. First, we treat suppliers like monolithic blocks, which is a significant simplification. Second, it is demonstrated that the existence of multiple belief systems inside a single partner organization may result in a politically charged process in which different coalitions within that organization react to inter-organizational events in different ways (Brattström & Faems, 2020)

We only looked at permissioned blockchains in terms of the blockchain itself. With an endless number of possible participants on the network who may read and upload data to the system, doing business on public blockchains opens up a whole new chapter. In the area of public blockchains, more study is needed to confirm our results. Furthermore, both initiatives were started and completed by the same firm, so more study is needed to see if our findings are applicable to other businesses. Because of the nature of our initiatives, they were conducted in circumstances where just a few partners were involved in collaborative governance. As a result, an effective governance structure could not be built, and more study is needed. We further encourage scholars to look at incentive systems to entice firms to participate in activities in difficult use case scenarios. This is especially true when a large number of companies join together to collaborate and compete at the same time.

Blockchain is still an emerging but well-received topic in management literature. In contrast to existing assumptions that blockchain can be seen as an alternative governance structure that create trusts, we found that blockchain should rather be seen as a complementary governance structure that relies on preexisting trust. We hope that our findings can inspire other scholars to further explore the challenges of blockchain implementation in a wide variety of settings. At the same time, we hope that our insights can support managers in improving the success rate of applying blockchain application in business projects.

CHAPTER FIVE: FINAL CONCLUSION

5.1. RESEARCH CONTRIBUTIONS

This cumulative thesis makes several contributions to the domain of innovation ecosystems. The first paper explores novel NLP developments to enhance bibliometric analysis in management research. We introduce an approach combining BERTopic, GPT-based categorization, and the semi-supervised CorEx model, allowing for a semi-automated process that incorporates human judgment in the topic-modeling process. The use of these techniques on a corpus of 2567 journal paper abstracts from innovation ecosystems has led to the identification of 27 distinct themes, showcasing an enhanced ability to process and interpret large sets of textual data in the management domain. The outcomes of our bibliographic analysis provide multifaceted insights for the academic community. Our insight into 27 innovation ecosystem topic trends over time can act as a compass for researchers to understand the evolutionary trajectory within the domain. Thus, our paper demonstrates the utility of its novel topic modeling approach in conducting a bibliometric review of innovation ecosystems. This approach has enabled a detailed and scalable analysis of a large volume of text, providing a quantitative perspective on the various thematic structures within a broad body of scholarly work on innovation ecosystems. Ultimately, our application shows the approach's capability to handle extensive research domains efficiently and effectively.

My second paper enriches the literature on ecosystem governance through contribution in two specific ways. First, we point to the potential relevance of an alternative ecosystem leadership style, where instead of claiming a centralized leadership role, the potential orchestrator relinquishes their dominant position. While centralized governance might look attractive in terms of individual rent-generating opportunities, too much domination by one company can breed distrust, which hampers ecosystem growth. In this respect, democratic multi-body governance with shared decision-making and the cultivation of a give-first mentality seems viable alternatives to secure partner buy-in in sustainability-centric ecosystems.

Our findings also clearly highlight that ecosystem managers should remain flexible and adapt governance approaches over time to address emerging challenges. Imposing full centralization or decentralization risks ecosystem instability or rejection. Overall, an adaptive leadership style coupled with inclusive governance and voluntary contributions appears vital for ecosystem success.

The third paper explores how and why blockchain projects applications fail to scale in a corporate setting. We identify three organizational challenges, explaining why blockchain-based companies struggle to scale: (i) Lack of suitable governance structure (ii) Lack of trust to successfully execute blockchain, and (iii) Lack of knowledge and skills. Jointly, our findings add to the emerging management literature on blockchain in two important ways. First, whereas recent management literature has positioned blockchain as a substitute for existing interfirm governance structures and third parties (Smits & Hulstijn, 2020), our findings point to blockchain as a complementing governance framework. In particular, we argue that traditional collaborative governance institutions still need to be implemented before blockchain-based governance can fully realize its potential. Finally, as opposed to popular belief that blockchain can be seen as a trust-enabling silver bullet, we find that the successful implementation of blockchain actually relies on preexisting trust, provided by sound governing structures.

5.2. FINAL REMARKS

In this thesis, I have intensively explored, analyzed, and discussed thriving and struggling decentralized innovation ecosystems. Reflecting on my work throughout the journey, I would like to add some final remarks in the following.

Adaptive Governance as the fundamental driver. My dissertation, initially titled "The Role of Blockchain for Innovation Ecosystems", began with a focus on exploring innovation ecosystems from a rather technological perspective. This was driven by blockchain's potential for decentralization, transparency, and security, suggesting a technological answer for fostering innovation. However, my research evolved to demonstrate that the significance of blockchain, whether in centralized or decentralized innovation ecosystems, is actually secondary to the implementation of effective, adaptive governance frameworks. This realization led me to refocus my thesis on examining the underlying governance mechanisms within these innovation ecosystems, emphasizing their critical role. Particularly, I have shown that in the absence of centralized control, a robust governance structure that ensures a balance between autonomy and order within the innovation ecosystem plays the main role in ensuring innovation ecosystem success. In other words, in decentralized innovation ecosystems, governance frameworks serve a purpose far beyond providing support. They establish clear protocols for collaboration and value creation and set boundaries that safeguard the integrity and direction of the ecosystem. This is particularly important in decentralized environments often characterized by a diverse range of actors and interests, where governance frameworks provide the necessary coherence to align individual actions with the overarching goals of the ecosystem. However, I have also shown that in decentralized scenarios, these governance structures must be flexible and adaptive, capable of evolving with the changing dynamics of the ecosystem. This dynamic is driven by the need to balance efficiency, inclusivity, and capability to technological and organizational challenges. Thus, for ecosystem architects, the process of developing governance structures in decentralized innovation ecosystems should be seen as a constantly evolving, iterative process and include mechanisms for continuous improvement to adapt to new challenges and (technological, i.e., blockchain) opportunities as they arise.

Interplay of Technology with Governance. The pivotal role of governance in shaping innovation ecosystems led me to reconsider its role with the technological tools at hand, particularly blockchain. The examination of this interplay underscores a critical insight: Governance does not merely coexist with technology; it fundamentally shapes how technology is deployed and thus impacts innovation in such ecosystems The relationship between technologies like blockchain and governance models in decentralized ecosystems is complex: The way governance is structured can either enhance or hinder the effective utilization of these technologies in driving innovation. For instance, blockchain's attributes of transparency, security, and decentralization can be harnessed to streamline decision-making, build trust, and manage interactions between actors. On the other hand, governance models that are overly rigid or not well-aligned with the capabilities of blockchain can have a disadvantageous effect. Such misalignment may result in underutilization of the technology's potential, leading to inefficiencies and exclusion. A governance framework that does not adequately address blockchain's capacity for transparency might actually fail to build trust, thereby hindering collaboration and the creation of value for participants within the ecosystem. This scenario exemplifies the critical role of governance in ensuring technology serves its intended purpose within these ecosystems. In the same vein, particularly exciting to me was the exploration of areas where the synergy between technology and governance was rejected by ecosystem participants. This observation suggests that the acceptance and effectiveness of technology-governance interplays are contingent upon the ecosystem participants' perceptions and experiences. Effective governance thus emerges as a dynamic and responsive framework that accommodates technological capabilities and ecosystem participants' needs. I hope to inspire academics and practitioners alike to explore further how blockchain can be integrated into existing governance models, thereby supporting a more effective innovation and collaboration environment. Exploring the relationship between technology and governance in decentralized ecosystems, it became clear that the efficacy of these elements is deeply rooted in a foundational component: Trust.

Trust as a Prerequisite. Blockchain acts as a tool that, when effectively integrated within governance frameworks, can add to the maintenance of trust. The establishment of trust, however, is contingent upon governance practices that are consistent, transparent, and equitable. Throughout my exploration of blockchain-supported innovation ecosystems, I witnessed that trust is not a given. Also, blockchain does not automatically foster trust. Instead, it is earned through the consistent application of governance practices like clear and open communication channels, accountability, and sometimes even conventional trust-providing governance methods like contracts between the actors. This ensures that all participants are sufficiently informed and can contribute to decision-making processes. Inclusivity in decision-making not only fosters trust but also ensures diverse perspectives are considered, enhancing the ecosystem's overall problem-solving and innovation capabilities. Moreover, accountability mechanisms play a crucial role in sustaining trust. These mechanisms are essential to ensure governance practices are not just promises but are actually followed through, and actions and decisions within the ecosystem align with agreed-upon rules and standards. The presence of accountability mechanisms reassures participants that their contributions are valued and protected, which is essential for ongoing cooperation and collaboration. On the other hand, the absence of a strong trust foundation can lead to significant challenges in collaboration and cooperation, ultimately slowing down the innovation process. Yet, without trust, decentralized innovation ecosystems may face issues such as fragmentation, where participants are unwilling to share ideas or resources, or even outright rejection of new initiatives due to skepticism or fear of exploitation. Exciting to me was that trust, as a long-standing, traditional component in organizational dynamics, contrasts to some extent with the relatively recent emergence of technologies like blockchain. This emphasizes that while technology can support trust, the core principles of building trust remain grounded in human interactions - and governance practices. This contrast also underscores the need for a balanced approach in decentralized innovation ecosystems, where established principles should be integrated with new technological innovations.

Balancing Traditional and Innovative Approaches. Balancing traditional governance practices with innovative approaches has been shown to be critical in decentralized ecosystems. This balance ensures stability while embracing the benefits of new technologies. Traditional governance methods provide robustness and structure, serving as the foundation upon which the ecosystem's integrity is built. At the same time, innovative approaches, often enabled by new technologies, can bring more flexibility and transparency, which are crucial in rapidly evolving innovation environments. The core challenge in achieving this balance lies in developing a governance framework that is both robust and flexible. Achieving this equilibrium requires a dynamic integration of new technological advancements without losing sight of the ecosystem's core objectives and capabilities. Thus, an understanding of both traditional and rather new governance approaches is needed alongside the ability to apply them in a way that complements and strengthens each other. For instance, decision-making can involve combining hierarchical and consensus-based approaches to optimize decision equity and stakeholder engagement. Trust building necessitates conventional trust-providing governance approaches like contracts, potentially supported by “smart” digital additions on the blockchain that could enforce actions based on certain conditions. Strategy development can require merging traditional planning and goal setting with machine-automated transaction control from various stakeholders within the ecosystem. When it comes to technology implementation, this balance means leveraging cutting-edge technologies while ensuring they align with the existing technological infrastructure and ecosystem participants’ know-how. It's about creating an environment where technology serves as a facilitator for governance rather than as a disruptor.

Ultimately, my dissertation aims to build a bridge between academia and industry practice. For academia, my goal is to inspire further studies at the intersection of decentralization, innovation, and technology. As the dissertation provides several practical insights, I hope to guide practitioners not only in understanding the potential of blockchain technology but also in considering the wider contexts of governance and organizational structures for effective implementation. I am optimistic that my thesis provokes thoughts and can be a valuable resource for both academics and practitioners in these critical areas.

CHAPTER SIX: REFERENCES

Aarikka-Stenroos, L., Ritala, P., & Thomas, L. (2021). Circular economy ecosystems: A typology, definitions, and implications. In. https://doi.org/10.4337/9781789906035.00024

Adner, R. (2006). Match Your Innovation Strategy To Your Innovation Ecosystem. *Harvard Business Review*, *84*, 98-107; 148.

Adner, R. (2012). *The Wide Lens: A New Strategy for Innovation*. Portfolio Penguin. https://books.google.de/books?id=4iGCndhBnBgC

Adner, R. (2016). Ecosystem as Structure: An Actionable Construct for Strategy. *Journal of Management*, *43*(1), 39-58. https://doi.org/10.1177/0149206316678451

Adner, R. (2021). *Winning the Right Game: How to Disrupt, Defend, and Deliver in a Changing World*. https://doi.org/https://doi.org/10.7551/mitpress/12752.001.0001

Adner, R., & Kapoor, R. (2010). Value Creation in Innovation Ecosystems: How the structure of technological interdependence affects firm performance in new technology generations. *Strategic Management Journal*, *31*(3), 306-333. http://www.jstor.org/stable/40587479

Adner, R., & Kapoor, R. (2016). Right Tech, Wrong Time

How to make sure your ecosystem is ready for the newest technologies. *Harvard Business Review*(Nov 2016), pp.60–67. https://hbr.org/2016/11/right-tech-wrong-time

Alkhudary, R., Brusset, X., & Fenies, P. (2020). Blockchain in general management and economics: a systematic literature review. *European Business Review*, *ahead-of-print*. https://doi.org/10.1108/EBR-11-2019-0297

Andre Mediote de, S., & Karin, B. (2021). Pro/Anti-Vaxxers in Brazil: A Temporal Analysis of COVID Vaccination Stance in Twitter. https://doi.org/10.5753/kdmile.2021.17467

Autio, E. (2022). Orchestrating ecosystems: a multi-layered framework. *Innovation*, *24*(1), 96-109. https://doi.org/10.1080/14479338.2021.1919120

Baldwin, C. Y., Bogers, M. L. A. M., Kapoor, R., & West, J. (2024). Focusing the ecosystem lens on innovation studies. *Research Policy*, *53*(3), 104949. https://doi.org/https://doi.org/10.1016/j.respol.2023.104949

Bawden, D., & Robinson, L. (2009). The Dark Side of Information: Overload, Anxiety and Other Paradoxes and Pathologies. *J. Information Science*, *35*, 180-191. https://doi.org/10.1177/0165551508095781

Bennett, M. (2021). 2093 Is it time to stop talking about blockchain? IBM Think 2021, Virtual.

Blei, D. M., & Lafferty, J. D. (2006). Dynamic topic models. ICML 2006 - Proceedings of the 23rd International Conference on Machine Learning,

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation [Article]. *Journal of Machine Learning Research*, *3*(4-5), 993-1022. https://www.scopus.com/inward/record.uri?eid=2-s2.0-0141607824&partnerID=40&md5=505ce8839ae28d1cb56a7ff91bd0ad2d

Bocken, N. M. P., & Geradts, T. H. J. (2020). Barriers and drivers to sustainable business model innovation: Organization design and dynamic capabilities. *Long Range Planning*, *53*(4), 101950. https://doi.org/https://doi.org/10.1016/j.lrp.2019.101950

Brattström, A., & Faems, D. (2020). Inter-Organizational Relationships as Political Battlefields: How Fragmentation within Organizations Shapes Relational Dynamics between Organizations. *Academy of Management Journal*. https://doi.org/10.5465/amj.2018.0038

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., . . . Amodei, D. (2020). *Language models are few-shot learners* Proceedings of the 34th International Conference on Neural Information Processing Systems, Vancouver, BC, Canada.

Buterin, V. (2015). On public and private blockchains. *Ethereum blog*, *7*, 180.

Camacho-Collados, J., & Pilevar, M. T. (2017). On the Role of Text Preprocessing in Neural Network Architectures: An Evaluation Study on Text Categorization and Sentiment Analysis.

Casino, F., Dasaklis, T. K., & Patsakis, C. (2019). A systematic literature review of blockchain-based applications: Current status, classification and open issues. *Telematics and Informatics*, *36*, 55-81. https://doi.org/https://doi.org/10.1016/j.tele.2018.11.006

Chang, J., Boyd-Graber, J., Gerrish, S., Wang, C., & Blei, D. M. (2009). Reading tea leaves: How humans interpret topic models. Advances in Neural Information Processing Systems 22 - Proceedings of the 2009 Conference, Chen, Y., Pereira, I., & Patel, P. C. (2020). Decentralized Governance of Digital Platforms. *Journal of Management*, 0149206320916755. https://doi.org/10.1177/0149206320916755

Chen, Y., Richter, J. I., & Patel, P. C. (2020). Decentralized Governance of Digital Platforms. *Journal of Management*, *47*(5), 1305-1337. https://doi.org/10.1177/0149206320916755

Cozzolino, A., & Geiger, S. (2024). Ecosystem disruption and regulatory positioning: Entry strategies of digital health startup orchestrators and complementors. *Research Policy*, *53*(2), 104913. https://doi.org/https://doi.org/10.1016/j.respol.2023.104913

Damert, M., Morris, J., & Guenther, E. (2020). Carbon Footprints of Organizations and Products. In W. Leal Filho, A. M. Azul, L. Brandli, P. G. özuyar, & T. Wall (Eds.), *Responsible Consumption and Production* (pp. 59-72). Springer International Publishing. https://doi.org/10.1007/978-3-319-95726-5_12

Daymond, J., Knight, E., Rumyantseva, M., & Maguire, S. (2022). Managing ecosystem emergence and evolution: Strategies for ecosystem architects [https://doi.org/10.1002/smj.3449]. *Strategic Management Journal*, *n/a*(n/a). https://doi.org/https://doi.org/10.1002/smj.3449

De Filippi, P., Mannan, M., & Reijers, W. (2020). Blockchain as a confidence machine: The problem of trust & challenges of governance. *Technology in Society*, *62*, 101284. https://doi.org/https://doi.org/10.1016/j.techsoc.2020.101284

Deloitte. (2020). C-Suite Briefing 5 Blockchain Trends for 2020.

Derstappen. (2021). Evaluating and applying machine learning algorithms to analyze entrepreneurial ecosystems. *Master thesis WHU*.

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding* https://doi.org/10.18653/v1/n19-1423

Du, J., Nielsen, B. B., & Welch, C. (2023). From Buzzword to Biz World: R ealizing B lockchain's P otential in the I nternational B usiness C ontext. *California Management Review*, *66*(1), 124-148. https://doi.org/10.1177/00081256231202266

Edmunds, A., & Morris, A. (2000). The problem of information overload in business organisations: a review of the literature. *International Journal of Information Management*, *20*(1), 17-28. https://doi.org/https://doi.org/10.1016/S0268-4012(99)00051-1

Eisenhardt, K. M. (1989). Building Theories from Case Study Research. *Academy of Management Review*, *14*(4), 532-550. https://doi.org/10.5465/amr.1989.4308385

Erhard, J., Kubis, C., Krebs, J.-M., & Liedke, A. (2019). *Overcoming barriers for corporate scope 3 action in the supply chain*. W. Germany. https://mobil.wwf.de/fileadmin/fm-wwf/Publikationen-PDF/WWF-Overcoming-barriers-for-corporate-scope-3.pdf / Accessed on 03.05.2020

ESTAINIUM. (2023). Time to act - Insights from the ESTAINIUM
working groups.

Farsan, A., Chang, A., Kerkhof , A., Bence, N., Yan, C., Villasana, F. R., & Labutong, N. (2018). *Value Change in the Value Chain: Best Practices in Scope 3 Greenhouse Gas Management*. Climate-KIC.

Felin, T., & Lakhani, K. (2018). What problems will you solve with blockchain. *MIT Sloan Management Review*, *Fall 2018*. https://www.researchgate.net/profile/Teppo-Felin/publication/328598250_What_problems_will_you_solve_with_blockchain/links/5d1d26e1a6fdcc2462bdb287/What-problems-will-you-solve-with-blockchain.pdf

Fenwick, M., Kaal, W. A., & Vermeulen, E. P. M. (2018). Why 'Blockchain' Will Disrupt
Corporate Organizations. *ECGI Working Paper Series in Law*.

Fuso Nerini, F., Sovacool, B., Hughes, N., Cozzi, L., Cosgrave, E., Howells, M., Tavoni, M., Tomei, J., Zerriffi, H., & Milligan, B. (2019). Connecting climate action with other Sustainable Development Goals [Review]. *Nature Sustainability*, *2*(8), 674-680. https://doi.org/10.1038/s41893-019-0334-y

Gallagher, R., Reing, K., Kale, D., & Ver Steeg, G. (2016). Anchored Correlation Explanation: Topic Modeling with Minimal Domain Knowledge. *Transactions of the Association for Computational Linguistics*, *5*. https://doi.org/10.1162/tacl_a_00078

Gallagher, R. J., Reing, K., Kale, D., & Ver Steeg, G. (2017). Anchored Correlation Explanation: Topic Modeling with Minimal Domain Knowledge. *Transactions of the Association for Computational Linguistics*, *5*, 529-542.

Gan, J., Tsoukalas, G., & Netessine, S. (2023). Decentralized Platforms: Governance, Tokenomics, and ICO Design. *Management Science*, *69*(11), 6667-6683. https://doi.org/10.1287/mnsc.2021.02076

Ganne, E. (2018). Can Blockchain revolutionize international trade? *World Trade Organization*.

Gaur, V., & Gaiha, A. (2020). Building a Transparent Supply Chain - Blockchain can enhance trust, efficiency, and speed. *Harvard Business Review*.

Gioia, D. A., Corley, K. G., & Hamilton, A. L. (2013). Seeking Qualitative Rigor in Inductive Research: Notes on the Gioia Methodology. *Organizational Research Methods*, *16*(1), 15-31-31. https://doi.org/10.1177/1094428112452151

Goldsby, C. M., & Hanisch, M. (2023). Agency in the algorithmic age: The mechanisms and structures of blockchain-based organizing. *Journal of Business Research*, *168*, 114195. https://doi.org/https://doi.org/10.1016/j.jbusres.2023.114195

Gorkhali, A., Li, L., & Shrestha, A. (2020). Blockchain: a literature review. *Journal of Management Analytics*, *7*(3), 321-343. https://doi.org/10.1080/23270012.2020.1801529

Granstrand, O., & Holgersson, M. (2020). Innovation ecosystems: A conceptual review and a new definition. *Technovation*, *90-91*, 102098. https://doi.org/https://doi.org/10.1016/j.technovation.2019.102098

Greene, D., O'Callaghan, D., & Cunningham, P. (2014, 2014//). How Many Topics? Stability Analysis for Topic Models. Machine Learning and Knowledge Discovery in Databases, Berlin, Heidelberg.

Grootendorst, M. (2022a). BERTopic: Neural Topic Modeling With a Class-Based TF-IDF Procedure. https://doi.org/10.48550/arxiv.2203.05794

Grootendorst, M. (2022b). BERTopic: Neural topic modeling with a class-based TF-IDF procedure. *arXiv preprint arXiv:2203.05794*.

Hambrick, D., & Finkelstein, S. (1987). Managerial discretion: A bridge between polar views of organizational outcomes. *Research in Organizational Behavior*, *9*, 369-406.

Hannigan, T. R., Haans, R. F. J., Vakili, K., Tchalian, H., Glaser, V. L., Wang, M. S., Kaplan, S., & Jennings, P. D. (2019). Topic Modeling in Management Research: Rendering New Theory from Textual Data. *Academy of Management Annals*, *13*(2), 586-632. https://doi.org/10.5465/annals.2017.0099

Hastie, T., Tibshirani, R., & Friedman, J. (2009). The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition *Springer Series in Statistics*.

Hellwig, D., Karlic, G., & Huchzermeier, A. (2020). *Build your own blockchain. a practical guide to distributed ledger technology* [Non-fiction]. Springer. https://search.ebscohost.com/login.aspx?direct=true&db=cat07353a&AN=whu.54751&site=eds-live

http://doi.org/10.1007/978-3-030-40142-9

Hoffman, M. D., Blei, D. M., & Bach, F. (2010). Online learning for Latent Dirichlet Allocation. Advances in Neural Information Processing Systems 23: 24th Annual Conference on Neural Information Processing Systems 2010, NIPS 2010,

Holotiuk, F., Pisani, F., & Moormann, J. (2017). *The Impact of Blockchain Technology on Business Models in the Payments Industry*.

Hopkinson, P., De Angelis, R., & Zils, M. (2020). Systemic building blocks for creating and capturing value from circular economy. *Resources, Conservation and Recycling*, *155*, 104672. https://doi.org/https://doi.org/10.1016/j.resconrec.2019.104672

Iansiti, M., & Karim, L. (2017). The Truth About Blockchain. *Harvard Business Review.*, *Jab-Feb 2017*, pp.118–127. https://hbr.org/2017/01/the-truth-about-blockchain

Jacobides, M. G., Cennamo, C., & Gawer, A. (2018). Towards a theory of ecosystems [https://doi.org/10.1002/smj.2904]. *Strategic Management Journal*, *39*(8), 2255-2276. https://doi.org/https://doi.org/10.1002/smj.2904

Jurafsky, D., & Martin, J. (2023). Speech and Language Processing

An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition. *Stanford University*, *(3rd ed. draft)*. https://web.stanford.edu/~jurafsky/slp3/ed3book_jan72023.pdf

Korhonen, J., Nuur, C., Feldmann, A., & Birkie, S. E. (2018). Circular economy as an essentially contested concept. *Journal of Cleaner Production*, *175*, 544-552. https://doi.org/https://doi.org/10.1016/j.jclepro.2017.12.111

Labutong, N. (2018). *How can companies address their scope 3 greenhouse gas emissions?* C. Worldwide. https://www.cdp.net/en/articles/companies/how-can-companies-address-their-scope-3-greenhouse-gas-emissions / accessed on 08.05.2020

Lacity, M., & Van Hoek, R. (2021). What We've Learned So Far About Blockchain for Business. *MIT Sloan Management Review*, *62*(3), 48-54. https://search.proquest.com/scholarly-journals/what-weve-learned-so-far-about-blockchain/docview/2502933036/se-2?accountid=44866

Lacoste-Julien, S., Sha, F., & Jordan, M. I. (2009). DiscLDA: Discriminative learning for dimensionality reduction and classification. Advances in Neural Information Processing Systems 21 - Proceedings of the 2008 Conference,

Lakshmanan. (2019). *How, When, and Why Should You Normalize / Standardize / Rescale Your Data?* https://towardsai.net/p/data-science/how-when-and-why-should-you-normalize-standardize-rescale-your-data-3f083def38ff

Lei, C. F., & Ngai, E. W. T. (2023). Blockchain from the information systems perspective: Literature review, synthesis, and directions for future research. *Information & Management*, *60*(7), 103856. https://doi.org/https://doi.org/10.1016/j.im.2023.103856

Leonnardo Benjamin, H., & Derwin, S. (2022). Indonesian Hoax News Classification With Multilingual Transformer Model and BERTopic. *Informatica*. https://doi.org/10.31449/inf.v46i8.4336

Lin, S., Hilton, J., & Evans, O. (2022). *TruthfulQA: Measuring How Models Mimic Human Falsehoods*. https://doi.org/10.18653/v1/2022.acl-long.229

Liu, Y., Liu, Z., Chua, T.-S., & Sun, M. (2015). Topical Word Embeddings. *Proceedings of the AAAI Conference on Artificial Intelligence*, *29*(1). https://doi.org/10.1609/aaai.v29i1.9522

Lloyd, S., Hadziosmanovic, M., Rahimi, K., & Bhatia, P. (2022). Trends Show Companies Are Ready for Scope 3 Reporting with US Climate Disclosure Rule. *World Resource Institute, 2022*. https://www.wri.org/update/trends-show-companies-are-ready-scope-3-reporting-us-climate-disclosure-rule

Lu, J., Javeed, S. A., Latief, R., Jiang, T., & Ong, T. S. (2021). The Moderating Role of Corporate Social Responsibility in the Association of Internal Corporate Governance and Profitability; Evidence from Pakistan. *International Journal of Environmental Research and Public Health*, *18*(11).

Lumineau, F., Schilke, O., & Wang, W. (2020). *Organizational Trust in the Age of the Fourth Industrial Revolution*. https://doi.org/10.13140/RG.2.2.20789.50401

Lumineau, F., Wang, W., & Schilke, O. (2020). Blockchain Governance—A New Way of Organizing Collaborations? *Organization Science*.

Lumineau, F., Wang, W., Schilke, O., & Schneider, L. (2021). How Blockchain Can Simplify Partnerships. *Harvard Business Review*, *4/21*. https://hbr.org/2021/04/how-blockchain-can-simplify-partnerships

Mahtab, K., Joseph, S., & Qingyun, Z. (2019). At the Nexus of Blockchain Technology, the Circular Economy, and Product Deletion [article in journal/newspaper]. https://doi.org/10.3390/app9081712

Malhotra, A., O'Neill, H., & Stowell, P. (2021). Thinking strategically about blockchain adoption risks and risk mitigation. *Business Horizons*. https://doi.org/https://doi.org/10.1016/j.bushor.2021.02.033

Mallick, T., Bergerson, J. D., Verner, D. R., Hutchison, J. K., Levy, L.-A., & Balaprakash, P. (2023). Analyzing the Impact of Climate Change on Critical Infrastructure From the

Scientific Literature: A Weakly Supervised NLP Approach. https://doi.org/10.48550/arxiv.2302.01887

Masucci, M., Brusoni, S., & Cennamo, C. (2020). Removing bottlenecks in business ecosystems: The strategic role of outbound open innovation [Article]. *Research Policy*, *49*(1). https://doi.org/10.1016/j.respol.2019.103823

Meijer, D., & Ubacht, J. (2018). *The governance of blockchain systems from an institutional perspective, a matter of trust or control?* Proceedings of the 19th Annual International Conference on Digital Government Research: Governance in the Data Age, Delft, The Netherlands. https://doi.org/10.1145/3209281.3209321

Morabito, V. (2017). *Business Innovation Through Blockchain*. https://doi.org/10.1007/978-3-319-48478-5

Ni Ki, C., Hosseinian-Far, A., Daneshkhah, A., & Salari, N. (2022). Topic modelling in precision medicine with its applications in personalized diabetes management [Article]. *Expert Systems*, *39*(4), Article e12774. https://doi.org/10.1111/exsy.12774

Nowiński, W., & Kozma, M. (2017). How Can Blockchain Technology Disrupt the Existing Business Models? *Entrepreneurial Business and Economics Review*, *5*(3), 173-188. https://EconPapers.repec.org/RePEc:krk:eberjl:v:5:y:2017:i:3:p:173-188

OpenAI. (2023). *Documentation / Models*. https://platform.openai.com/docs/models

Papalexopoulos, D., & Klein, C. (2022). Towards an EU Action Plan for a Digitally Enabled Green Transition. *European Round Table for Industry*.

Peltola, T., Aarikka-Stenroos, L., Viana, E., & Mäkinen, S. (2016). Value capture in business ecosystems for municipal solid waste management: Comparison between two local environments. *Journal of Cleaner Production*, *137*, 1270-1279. https://doi.org/https://doi.org/10.1016/j.jclepro.2016.07.168

Pelz, B. (2021). Research Methods for the Social Sciences. *Herkimer College / SUNY* https://courses.lumenlearning.com/suny-hccc-research-methods/

Pidun, U., Reeves, M., & Schüssler, M. (2020). Why Do Most Business Ecosystems Fail? *CORPORATE FINANCE AND STRATEGY*, *June 22, 2020*. https://www.bcg.com/publications/2020/why-do-most-business-ecosystems-fail

Pujadas, R., Valderrama, E., & Venters, W. (2024). The value and structuring role of web APIs in digital innovation ecosystems: The case of the online travel ecosystem. *Research Policy*, *53*(2), 104931. https://doi.org/https://doi.org/10.1016/j.respol.2023.104931

PwC. (2021). *PwC's Global Blockchain Survey*

. Retrieved 14.02.2021 from www.pwccn.com/global-blockchain-survey-2018

Radford, A., & Narasimhan, K. (2018). Improving Language Understanding by Generative Pre-Training.

Rahman, M., Frame, J., Lin, J., & Nearing, G. (2020). *Hidden Stories: Topic Modeling in Hydrology Literature*. https://doi.org/10.31223/osf.io/2sy7a

Rajala, R., Hakanen, E., Mattila, J., Seppälä, T., & Westerlund, M. (2018). How Do Intelligent Goods Shape Closed-Loop Systems? [Article]. *California Management Review*, *60*(3), 20-44. https://doi.org/10.1177/0008125618759685

Reeves, M., Lotan, H., Legrand, K., & Jacobides, M. (2019). How Business Ecosystems Rise (and Often Fall). *MIT Sloan Management Review*, *July 30, 2019*. https://sloanreview.mit.edu/article/how-business-ecosystems-rise-and-often-fall/

Řehůřek, R., & Sojka, P. (2010). *Software Framework for Topic Modelling with Large Corpora*. https://doi.org/10.13140/2.1.2393.1847

Ricketts, J., Barry, D., Guo, W., & Pelham, J. (2023). A Scoping Literature Review of Natural Language Processing Application to Safety Occurrence Reports. *Safety*, *9*(2).

Romero, A. (2021). *Understanding GPT-3 In 5 Minutes*. https://towardsdatascience.com/understanding-gpt-3-in-5-minutes-7fe35c3a1e52

Saberi, S., Kouhizadeh, M., Sarkis, J., & Shen, L. (2019). Blockchain technology and its relationships to sustainable supply chain management. *International Journal of Production Research*, *57*(7), 2117-2135. https://doi.org/10.1080/00207543.2018.1533261

Schmeiss, J., Hoelzle, K., & Tech, R. P. G. (2019). Designing Governance Mechanisms in Platform Ecosystems: Addressing the Paradox of Openness through Blockchain Technology [Article]. *California Management Review*, *62*(1), 121-143. https://doi.org/10.1177/0008125619883618

Smits, M., & Hulstijn, J. (2020). Blockchain Applications and Institutional Trust [Original Research]. *Frontiers in Blockchain*, *3*(5). https://doi.org/10.3389/fbloc.2020.00005

Snihur, Y., & Bocken, N. (2022). A call for action: The impact of business model innovation on business ecosystems, society and planet. *Long Range Planning*, *55*(6), 102182. https://doi.org/https://doi.org/10.1016/j.lrp.2022.102182

Steyvers, M., Smyth, P., Rosen-Zvi, M., & Griffiths, T. (2004). Probabilistic author-topic models for information discovery. KDD-2004 - Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Stonig, J., & Müller-Stewens, G. (2019). Navigating the Challenges of Ecosystem Emergence: A Multi-Level Review of Leader and Complementor Strategies. *Die Unternehmung*, *73*, 288-307. https://doi.org/10.5771/0042-059X-2019-4-288

Tapscott, D., & Tapscott, A. (2016). *Blockchain Revolution: How the Technology Behind Bitcoin Is Changing Money, Business, and the World*. Portfolio.

Teerikangas, S., Onkila, T., Koistinen, K., Mäkelä, M., Aarikka-Stenroos, L., Ritala, P., & D. W. Thomas, L. (2021). *Research Handbook of Sustainability Agency*. Edward Elgar Publishing. https://doi.org/10.4337/9781789906035

10.4337/9781789906035.00024

Tettamanzi, P., Venturini, G., & Murgolo, M. (2022). Sustainability and Financial Accounting: a Critical Review on the ESG Dynamics. *Environmental Science and Pollution Research*, *29*(11), 16758-16761. https://doi.org/10.1007/s11356-022-18596-2

Thomas, L., & Autio, E. (2019). Innovation ecosystems. *Available at SSRN 3476925*.

Thomas, L., & Autio, E. (2020). Innovation ecosystems in management: An organizing typology. In. https://doi.org/10.1093/acrefore/9780190224851.013.203

Thomas, L., & Ritala, P. (2022). Ecosystem Legitimacy Emergence: A Collective Action View. *Journal of Management*, *48*(3), 515-541. https://doi.org/10.1177/0149206320986617

Treiblmaier, H. (2019). Toward More Rigorous Blockchain Research: Recommendations for Writing Blockchain Case Studies [10.3389/fbloc.2019.00003]. *Frontiers in Blockchain*, *2*, 3. https://www.frontiersin.org/article/10.3389/fbloc.2019.00003

van Kessel, P. (2019, 29.12.2023). Interpreting and validating topic models. *Pew Research Center: Decoded*. https://medium.com/pew-research-center-decoded/interpreting-and-validating-topic-models-ff8f67e07a32

van Rijmenam, M., Schweitzer, J., & Williams, M.-A. (2017). A Distributed Future:How Blockchain Affects Strategic Management, Organisation Design & Governance. *Academy of Management Proceedings*, *2017*, 14807. https://doi.org/10.5465/AMBPP.2017.14807abstract

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention Is All You Need. *CoRR*, *abs/1706.03762*. http://arxiv.org/abs/1706.03762

Vincent, R. (2021). Exploring Topic Modelling using Semi-Supervised Learning (Correlation Explanation). *medium.com*. https://medium.com/version-1/exploring-topic-modelling-using-semi-supervised-learning-correlation-explanation-b81d2603c9a2

Vogel, J., Hagen, S., & Thomas, O. (2019). Discovering Blockchain for Sustainable Product-Service Systems to enhance the Circular Economy [text]. https://search.ebscohost.com/login.aspx?direct=true&db=edsbas&AN=edsbas.F27CFBBA&site=eds-live

WEF. (2023). The "No-Excuse" Framework to Accelerate the Path to Net-Zero Manufacturing and Value Chains. *Industry Net Zero Accelerator Initiative, Jan 2023*.

Weinhold, M., & Faems, D. (2022). *Understanding Challenges and Intricacies surrounding the Scalability of Blockchain Applications within Large Industrial Corporations* EURAM Conference 2022 Winterthur

Weinhold, M., & Faems, D. (2024). Developing an Integrated Topic Modeling Approach for Knowledge Accumulation. *Academy of Management Conference 2024*.

Weinhold, M., Faems, D., & Beitinger, G. (2024). Ecosystem governance dynamics: Navigating the tension between centralization and decentralization in sustainability-centered ecosystems. *Long Rangle Planning*.

Werbach, K. (2018). The Blockchain and the New Architecture of Trust. *The MIT Press*.

Yasin Musa, A. (2023). Towards Election Forecasting Using Sentiment Analysis: The Zambia General Elections 2021. *Zambia Ict Journal*. https://doi.org/10.33260/zictjournal.v7i1.148

Zachariadis, M., Hileman, G., & Scott, S. V. (2019). Governance and control in distributed ledgers: Understanding the challenges facing blockchain technology in financial services. *Information and Organization*, *29*(2), 105-117. https://doi.org/https://doi.org/10.1016/j.infoandorg.2019.03.001

Zwitter, A., & Hazenberg, J. (2020). Decentralized Network Governance: Blockchain Technology and the Future of Regulation [Hypothesis and Theory]. *Frontiers in Blockchain*, *3*. https://doi.org/10.3389/fbloc.2020.00012