Mikael Sarja

Valleys of Death and Business Configuration

Case B2G Service Pilot

School of Marketing and Communication
Master's Thesis
Marketing Management

Vaasa 2026

UNIVERSITY OF VAASA
School of Marketing and Communication
Author: Mikael Sarja
Title of the thesis: Valleys of Death and Business Configuration: Case B2G Service Pilot
Degree: Master of Science in Marketing
Degree Programme: Marketing Management
Supervisor: Hannu Makkonen
Year: 2026 **Pages:** 99

ABSTRACT:

Julkisen sektorin digitalisaatio ja palvelujen uudistaminen ovat lisänneet tarvetta uusille digitaalisille ratkaisuille, mutta lupaavien innovaatioiden eteneminen varhaisesta kokeilu- ja kehitysvaiheesta vakiintuneeseen käyttöön on edelleen haastavaa. Tämä tutkielma tarkastelee ilmiötä startup-lähtöisesti business-to-government (B2G) -kontekstissa. Tavoitteena on edistää ymmärrystä siitä, miten digitaalisia palveluinnovaatioita voidaan skaalata ja miksi lupaavat ratkaisut jäävät pilotointivaiheeseen. Kuolemanlaaksoa (Valley of Death, VoD) ja pilottiloukkua (pilotitis) tarkastellaan liiketoimintamallin rakenteen, ekosysteemin koordinoinnin ja institutionaalisen kontekstin dynamiikkojen kautta. Tutkimus toteutettiin kirjallisuuskatsauksena ja laadullisena tapaustutkimuksena. Empiirisenä kohteena oli anonyymin GovTech-startupin pilottiprojekti suomalaisella hyvinvointialueella. Aineisto koostui pilotin kannalta tärkeimpien osapuolten kanssa toteutetuista viidestä puolistrukturoidusta haastattelusta sekä dokumenttiaineistosta. Haastatteluaineisto analysoitiin laadullisen sisällönanalyysin avulla. Analyysiä ohjasi kirjallisuuskatsauksen pohjalta rakennettu teoreettinen viitekehys. Tulokset osoittavat, että kuolemanlaakso näyttäytyy tässä kontekstissa kahden erilaisen konfiguraation välisenä ongelmana. Pilotti muodostaa paikallisesti toimivan, mutta väliaikaisen järjestelyn, mutta pilotin jälkeinen eteneminen edellyttää toisenlaista, toistettavampaa ja paremmin arjen toimintaan upotettua konfiguraatiota. Analyysin perusteella tunnistettiin kolme skaalaamisen kannalta huomioitavaa keskeistä kokonaisuutta: pilotin jälkeinen jatkopolku, toimitustavan muokkaaminen rutiiniin sopivaksi sekä laajemman ekosysteemin muodostuminen laajempaa etenemistä varten. Tutkielma kokoaa yhteen hajanaista kirjallisuutta kuolemanlaaksosta, pilottiloukusta ja digitaalisten palveluinnovaatioiden skaalaamisesta ja tuo nämä ilmiöt vahvemmin julkisen sektorin B2G-kontekstiin. Lisäksi se tuo kirjallisuuteen startup-lähtöisen näkökulman ja osoittaa, että pilotin jälkeinen eteneminen tai skaalaaminen riippuu liiketoimintamallin, ekosysteemin koordinoinnin ja institutionaalisen toimintaympäristön konfiguraatiosta ja yhteensopivuudesta.

KEYWORDS: startup-yritykset, julkinen sektori, julkiset hankinnat, pilotointi, käyttöönotto, liiketoimintamallit, liiketoimintaekosysteemit

Contents

1 Introduction 6
1.1 Background 6
1.2 Research gap 9
1.3 Research questions and objectives 11
1.4 Research approach and structure 12
2 Literature review 15
2.1 The challenges in scaling digital service innovations 15
2.1.1 The Valleys of Death (VoD & DVoD) 15
2.1.2 Pilotitis 18
2.1.3 From diffusion to embedding for scaling 20
2.2 Business model lens 22
2.2.1 Defining business models 23
2.2.2 Business model reconfiguration across phases 24
2.2.3 Scaling-relevant business model dimensions 25
2.3 Ecosystem lens 27
2.3.1 Defining ecosystems 27
2.3.2 Scaling-relevant ecosystem dimensions 29
2.4 Institutional context 30
2.4.1 Defining the institutional context as a B2G business environment 30
2.4.2 Deepening and contextualizing scaling constraints in B2G 32
2.4.3 Scaling-relevant institutional dimensions 34
2.5 Synthesis and theoretical framework 36
3 Methodology and research method 40
3.1 Research design and approach 40
3.2 Data collection and description of the data 41
3.3 Data analysis method 43
3.4 Validity and trustworthiness of the study 44
3.5 Case: Company X 47

4 Findings 48
4.1 Business model design in the pilot arrangement 48
4.1.1 The pilot as a mechanism for learning and benefit validation 48
4.1.2 The value logic pursued through the pilot 50
4.1.3 Cost structure and pricing logic under pilot conditions 52
4.2 Ecosystem coordination in the pilot configuration 53
4.2.1 The actor structure and coordination around the pilot 54
4.2.2 Dependencies, bottlenecks, and challenges in the ecosystem 55
4.3 Institutional context in the pilot configuration 57
4.3.1 Interaction constraints in the pilot 58
4.3.2 Legitimacy tensions in continuation decisions 60
4.3.3 Resource asymmetry in broader contracting 63
4.4 Summary of empirical findings 65
5 Discussion 68
5.1 Scaling-relevant focus areas 69
5.1.1 Building a continuation pathway 70
5.1.2 Reconfiguring delivery for embedding 71
5.1.3 Complementary actors and the ecosystem 73
5.1.4 Synthesis of the scaling-relevant focus areas 75
5.2 Answers to research questions & study contributions 76
5.3 Practical implications 77
5.4 Limitations and future research 82
6 Conclusion 86
References 88
Appendices 96
Appendix 1. Interview structure & interviewee-specific tailoring 96
Appendix 2. Analytical coding by lens and dimension 98

Figures

Figure 1. The Double Valley of Death, DVoD (European Commission, 2025, p. 2) 16
Figure 2. Theoretical framework: Lenses and dimensions for examining scaling 36
Figure 3. Data Analysis Process Guided by the Analytical Lenses 43
Figure 4. Empirically Enriched Framework: VoD as a Configurational Problem 68

Tables

Table 1. Summary of Conducted Interviews 42
Table 2. Future Research Directions & Questions 84

1 Introduction

1.1 Background

In order to modernize public services, the public sector is increasingly relying on *GovTech*, which is defined as digital solutions developed by private sector startups to modernize public service delivery (Bharosa, 2022; Hoekstra et al., 2022). In Finland, digitalization has been framed as a main enabler of public sector renewal, and the country has been ranked one of the EU's leading performers in public sector digitalization (Karjalainen, 2020). This change which promises to automate complicated tasks, improve decision-making, and radically alter citizen-government interactions, is today centered on artificial intelligence (AI) and other digital service innovations (Engin & Treleaven, 2019). This *Business to Government* (B2G) collaboration operates within an institutional setting with heavy procedures, such as those regarding procurement (Mergel et al., 2022).

Public procurement is a central feature of the B2G environment. A critical structural element is the procurement threshold that acts as a boundary for below-threshold procurement, which is often deemed more accessible to SMEs and startups (Hoekstra et al., 2022). According to Tukiainen et al. (2023), wellbeing services counties in Finland spend an estimated 8 billion euros annually on external purchases and about 2.5 billion euros of this is procured as above-threshold competitive tenders. At the same time, competition in these is generally weak, which raises concerns about the efficiency and the ability of public procurement to drive forward innovation and support new solutions. Business Finland's empirical material (2022) relatedly shows that buyers and suppliers describe competitive tendering as a challenge for the spread of new solutions, because the process is slow and the outcome does not always match the need (Merisalo et al, 2022). This creates a tension, because while the extensive below-threshold procurement resulting from these issues may lower entry barriers, the contract sizes could remain too small to support the investments needed for startups to *scale.*

Despite the potential of these publically procured GovTech solutions, a persistent risk known as *the Valley of Death* (VoD), originally applied to entrepreneurial context by Stephan Markham (2002), threatens the application of their advantages. This concept represents the early stages of innovation, specifically the gap between technical invention or market recognition of an idea and the efforts to commercialize it, where development is frequently obstructed by among many other reasons a lack of ecosystem support or overall funding (Ellwood et al., 2022). Merisalo et al. (2022) note that in public procurement this transition is supported when the procurement and the steps after are considered and structured already during the initial stages.

There are different forms of validation in the early stages such as Proof of Concept (PoC) and pilot projects, where innovative startups can demonstrate the technical feasibility of a technology that may still only exist on a small scale (McIntyre, 2014) and generate evidence regarding the efficacy and implementation of an innovation (Scarbrough et al., 2024). In the public sector, pilot projects are a common step toward crossing the gap towards implementation, as piloting allows the testing of products and services early on and to hedge risks before committing extensive resources (Mergel et al., 2022).

Although pilots seem like a good step towards commercialization, repeated pilots or being unable to move past the pilot phase is, in health and medical literature, labelled as the phenomenon of *pilotitis*. This describes a condition where digital innovations demonstrate technical success in controlled environments but fail to transition into full-scale production or achieve widespread adoption (Bhatia et al., 2020; Egermark et al., 2022; Gerhardt et al., 2015; Huang et al., 2017; Scarbrough et al., 2024, Schiffelers et al., 2025), effectively stalling in the VoD between funded experimentation and more permanent production (Egermark et al., 2022). Entry through relationships and a credible early stage offering may be enabled by pilot projects, even when later continuation depends on tendering. This can produce pilotitis by supporting temporary experimentation without creating a viable route into routine service, while also utilizing public resources.

The persistence of problems like pilotitis points that digital innovations may achieve technical feasibility yet still fail when they are not aligned with their environment (Schiffelers et al., 2025). Thus, when developing the value logic for early stage technology, it is not sufficient enough to focus on the technology in isolation. Instead, the commercial and systemic logic required to sustain it should be examined (Chesbrough & Rosenbloom, 2002). In the context of public service, the challenge is amplified because moving from a pilot to a scaled service depends on having a usable economic logic and also on the *embedding* of the innovation into a complex and regulatory heavy environment (Scarbrough & Kyratsis, 2022). This suggests that to understand scaling beyond the pilot stage, attention must be paid to the architecture of the offering and the wider configuration which value is created and maintained through.

This thesis adopts a three-lens theoretical framework. The first lens is *business model design*, which is used to examine how value is created, delivered, and captured through the offering (Teece, 2010; Wirtz et al., 2016). The second lens is *ecosystem coordination*. It focuses on the alignment, orchestration, and intermediation required between actors dependent on each other for the offering to work and move beyond the pilot stage (Adner, 2017; Kowalkowski et al., 2024). The third lens is *institutional context*. It is used to capture how the broader B2G environment affects what kinds of configurations are feasible, legitimate, and supportable in practice. These three lenses are used to examine scaling as a configurational issue emerging from their interaction. The further reasoning for combining these lenses regarding this challenge, and the research gap the framework addresses, is discussed in section 1.2.

To examine the challenge of moving digital service innovations beyond the pilot phase and across VoD in a public sector setting through the outlined theoretical framework, this thesis uses the case of Company X. At the time of writing, Company X is piloting an AI-assisted preliminary information collection service for social welfare clients, supporting professional-led service needs assessments in a Finnish wellbeing services

county. The name Company X is used to ensure anonymity, as the study aims to analyse sensitive material related to the company's business structure.

1.2 Research gap

Although VoD is recognized as a bottleneck in innovation, existing research on why digital services fail to progress past the initial stage is conceptually fragmented and spread across multiple domains. Management and technology commercialization studies discuss VoD as the transition between invention and commercialization, emphasizing that this phase remains difficult despite combat efforts (Ellwood et al., 2022). Drug development literature describes a similar pattern under the label translational gap, which happens between R&D and commercialization (Calza et al., 2021), whereas health innovation literature refers to the same underlying pattern as pilotitis, describing digital interventions that show promise in pilots but fail to become embedded in routine practice (Huang et al., 2017; Egermark et al., 2022; Bhatia et al., 2020; Schiffelers et al., 2025). Streams on university technology transfer and science commercialization examine early phases of moving or translating projects into ventures and they identify critical junctures such as opportunity recognition, entrepreneurial commitment, credibility and sustainability (Vohora et al., 2004), yet they do not explicitly use terminology like VoD. *These strands show that most innovations stall between proof of concept and larger-scale adoption, but they employ different labels, focus on different parts of the process, and rarely speak directly to each other.*

This fragmentation means different perspectives dominate different literatures. Much of the science commercialization literature focuses on the early stages and concentrates on funding gaps and proof of concept (Ellwood et al., 2022). Health and implementation research really focuses on diffusion in organizations and the embedding of innovations into routines (Huang et al., 2017; Scarbrough & Kyratsis, 2022). Vohora et al. (2004) show that ventures must repeatedly reconfigure their capabilities to move from research toward sustainable business, but they provide limited detail on how this is reflected in

the structure of the venture and its surroundings. *The perspective of the startup or early-stage venture as an actor that must navigate multiple institutional logics and design a viable commercial path through this phase remains underdeveloped.*

This need is very prominent in GovTech and B2G contexts. Studies in the field of GovTech show that startups developing solutions for the public sector operate within heavy regulation, procurement rules, and accountability requirements that make experimentation and scaling especially challenging (Bharosa, 2022; Hoekstra et al., 2022; Mergel et al., 2022). *This literature identifies obstacles for scaling in public organizations but it rarely analyzes how those constraints shape the configuration of a startup's business logic and surrounding requirements.*

Two well-developed perspectives of business model design and ecosystem coordination are especially relevant for analyzing this problem. Ecosystem literature challenges the assumption that a strong knowledge or innovation ecosystem will automatically evolve into a viable business ecosystem that can deliver value to end users, because the two operate under fundamentally different logics and drivers (Clarysse et al., 2014). The VoD can be understood as a manifestation of this disconnect, as it describes the gap between the research economy, driven by discovery and public investment, and the commercial economy, driven by profit and market logic (Jucevicius et al., 2016). In this transition, the business model is critical because it acts as a mediating construct. It translates inputs from the innovation ecosystem into an economic logic that can support a functioning business ecosystem (Chesbrough & Rosenbloom, 2002). This is why the business model helps specify how the venture is intending to create, deliver, and capture value, and helps distinguish technically promising directions from the ones that lack viable commercial potential (Chesbrough & Rosenbloom, 2002; Teece, 2010). *Recent work still finds that the business models of GovTech remain poorly understood which is limiting the integration of startups into public sector ecosystems (Bharosa, 2022).*

In B2G settings, business model design and ecosystem coordination cannot be understood properly without also accounting for institutional context. Public-sector rules, procurement practices, legitimacy requirements, and resourcing conditions shape what kinds of business models and ecosystem configurations are feasible, legitimate, and supportable in practice, yet this contextual shaping role receives limited attention in existing scaling research. *What remains unclear is how business model design and ecosystem coordination can be configured in B2G settings, and how institutional context shapes that configuration in practice.*

There is a clear gap at the intersection of these research streams, precisely where the most important institutional and structural problems lie. *We lack integrated, startup-oriented research that examines how digital service innovations in B2G settings are configured across business model design, ecosystem coordination, and institutional context, and how this configuration structures what is feasible beyond the pilot stage.* This thesis addresses that gap by focusing on the joint configuration of all three elements in a concrete B2G digital service pilot in the Finnish wellbeing services county context.

1.3 Research questions and objectives

The purpose of this thesis is to advance the understanding of how digital service innovations can be scaled by analyzing the configuration of business model design, ecosystem coordination, and institutional context. The study treats scaling as an outcome of the configuration between them. Using a GovTech pilot as the empirical setting, the thesis explains the VoD and the risk of pilotitis by analyzing how this configuration structures what is feasible beyond pilots.

Because the case company is in the pilot phase, scaling is examined analytically rather than as a measured outcome. The focus is on how the pilot is currently configured across the venture's business model design and surrounding ecosystem coordination, and how the institutional context shapes these. In this thesis, pilotitis is treated as a specific

manifestation of the broader concept of VoD in contexts where promising innovations struggle to move beyond piloting. The pilot therefore functions as the observable point where the transition challenges associated with VoD and pilotitis become visible in practice, enabling examination of how the surrounding configuration supports or limits movement beyond the pilot.

On this basis, the thesis aims to answer the following research questions **(RQ):**

- **RQ1:** What are the key dynamics between business model design, ecosystem coordination, and institutional context in the scaling of digital service innovations?
- **RQ2:** How are the business model design and ecosystem coordination configured in a digital service innovation pilot within a B2G institutional context?

The objective of RQ1 is to *develop a clear, literature-based understanding of these dynamics and clarify how business model, ecosystem, and institutional perspectives can be combined to analyze scaling challenges as configurational problems.* This objective is addressed through a structured review and synthesis of research on scaling, business models, ecosystems, and institutional context, resulting in a conceptual basis for interpreting the empirical case. The objective of RQ2 is to *examine how these configurational dynamics appear in a B2G pilot and how the institutional context shapes what is feasible.* This objective is addressed through a case study of Company X, using interviews and document analysis to identify how the case's business model design and ecosystem coordination are configured during the pilot and how the institutional B2G context shapes that configuration in ways that matter for moving beyond piloting.

1.4 Research approach and structure

This thesis uses a qualitative approach to examine the phenomenon from the perspective of the participants and see their interpretations. The study is taking a pragmatist stance and treats theory as an analytical tool and not as a set of fixed criteria.

The research is conducted as a single-case study of Company X in a B2G context. RQ1 is approached through a structured literature review that develops the conceptual basis around business model design, ecosystem coordination, and institutional context. Its final synthesis is completed later by bringing this framework into dialogue with the empirical findings. RQ2 is addressed through an empirical case analysis in which the pilot's current configuration is examined across these elements and their relationships.

The empirical data consists of semi-structured interviews with the case company and relevant stakeholders, complemented by document analysis. The data is analyzed using qualitative content analysis. The concepts developed in the literature review guide what is examined, while the interpretation remains open to case-specific patterns in how these elements are configured and interact. The aim is to describe the configuration clearly and interpret what it implies for feasibility beyond initial experimentation.

The scope of the study is deliberately delimited to the pilot-phase configuration of a GovTech startup operating in a Finnish wellbeing services county B2G context. The analysis focuses on the commercial and systemic setup around the pilot, how business model design and ecosystem coordination are arranged and how the institutional context shapes what is feasible. The study does not evaluate any of the technical performance of the solution, any user outcomes, or service effectiveness. Instead, this is an in-depth examination of how a pilot stage B2G innovation is configured in practice and how this configuration helps explain the Valley of Death and the risk of pilotitis in practice.

This thesis consists of six sections. The first section introduces the research problem, key concepts, RQs, and the positioning of the study. The second section the literature review and develops the framework used in the empirical analysis. It covers the challenges of VoD and pilotitis, embedding and the theoretical perspectives of business model design, ecosystem coordination, and institutional context. The third section explains the research design and methods, describes the data and analysis approach, and introduces

the case. The fourth section presents the empirical findings by describing how the pilot is configured across the framework elements and how they interact. The fifth section synthesizes the findings and presents implications, limitations, and directions for future research. The final section concludes the study.

2 Literature review

2.1 The challenges in scaling digital service innovations

In this section, the aim is to explain why digital service innovations often stall between early validation and wider use. Drawing from literature on the Valleys of Death, pilotitis, and innovation embedding, the section presents the scaling gap as a transition from experimentation toward routine service delivery. The section clarifies difficulties that prior research has associated with this transition and motivates examining scaling as a business configuration problem. This provides background for the core focus of the thesis: analyzing how business model design, ecosystem coordination, and the institutional context interact, and what scaling-relevant success areas are implied by that dynamic.

2.1.1 The Valleys of Death (VoD & DVoD)

The Valley of Death (VoD) has gathered multiple different definitions from different disciplines and is still not a single, universally defined concept. A systematic literature review by Gbadegeshin et al. (2022) shows that the term is used alongside several substitutes (e.g., chasm, funding gap, innovation drought) and is defined through multiple lenses, most commonly as a financial gap but also as a broader period or process marked by limited structure, resources, and expertise. Building on this, VoD is used in innovation management literature to describe the difficult transition from a technically working invention to a viable service offering that achieves sustained uptake in practice (Ellwood et al., 2022; Markham, 2002). Importantly, this framing already points beyond a mere lack of funding by treating VoD as a transition problem where the requirements for progress change as the innovation moves forward.

Although VoD is often described as a period where early funding and support ends before development is done, both Jucevicius et al. (2016) and Markham et al. (2010) also frame it as a deeper structural mismatch between research and commercialization. The

research phase prioritizes discovery and new knowledge, whereas routine delivery depends on the venture's ability to implement those innovations to the marketplace (Jucevicius et al., 2016). Ellwood et al. (2022) emphasize that these phases operate under different institutional logics, so the transition must be actively managed or those differences can become a barrier to transferring the innovation forward. Gbadegeshin et al. (2022) argue that many startups fall into the VoD because they lack the skills, competences, personnel, and infrastructure needed. This meaning that progress towards commercialization requires technical development and capability-building.

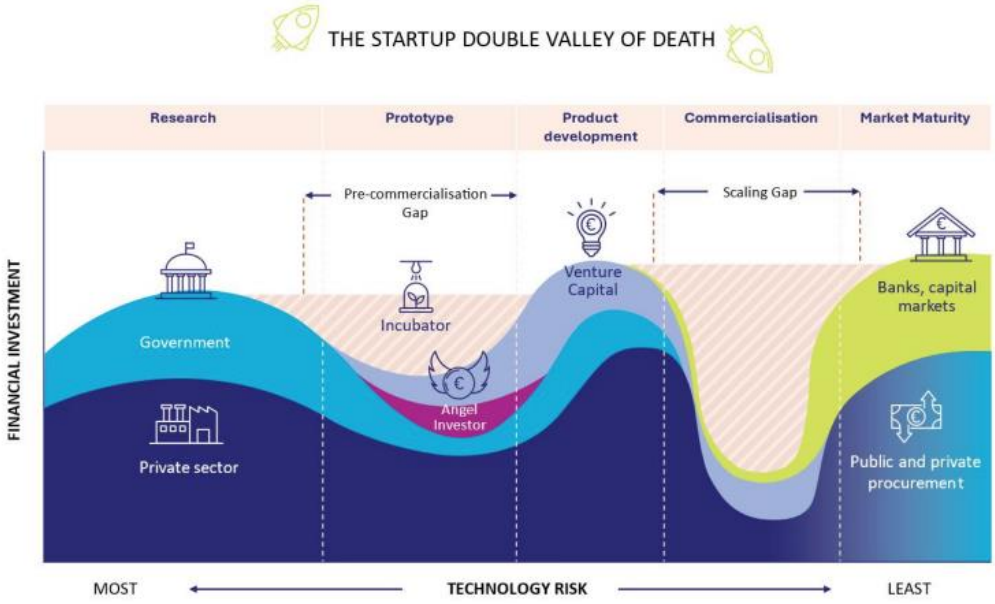

Figure 1. The Double Valley of Death, DVoD (European Commission, 2025, p. 2)

The European Commission's (2025, p. 2) policy analysis suggests that startup scaling in Europe often has two distinct gaps, shown in the idea of Double Valley of Death (DVoD) (see Figure 1). In their proposed framing, the first gap occurs when innovations fail to move from research to product, and the second when companies struggle to scale after early market entry. The Commission also links this scaling challenge to structural constraints in the environment which are e.g., fragmented markets, limited risk appetite, slow uptake of innovation, and underutilized public procurement (European Commission, 2025). Limited risk and slow uptake are also echoed by Popp et al. (2017), who note that

because new technology overall is seen as high-risk, industry and commercial investors often delay uptake until regulatory approvals are secured and trials are successful. This leaves ventures to carry financial burdens even longer.

Evidence from sectors that are capital-intensive supports a similar pattern with two gaps. Hartley et al. (2013, p. 9) report Zindler and Locklin's finding that financial shortfalls often rise in two places. First, when a technology is ready exit the lab and later when much larger funding is needed to prove its viability at commercial scale. They also note that the second gap is harder to cross because financial actors are not well positioned to fund the larger-scale deployment of new technologies, and because venture capital often requires short or medium term returns (Hartley et al., 2013). These perspectives suggest that getting technology to work in an isolated setting is not the same as getting it to actual scale, because later stage progress depends on different requirements than early stage validation (European Commission, 2025; Hartley et al., 2013).

Across the literature, VoD is repeatedly associated with recurring barriers that surface when ventures attempt to progress from technical validation toward routine service delivery. These include challenges related to financial continuity and fit-for-purpose resourcing (Hartley et al., 2013; Lefebvre et al., 2022), legitimacy under uncertainty and the ability to demonstrate reliability without a long track record (Ellwood et al., 2022; Lefebvre et al., 2022), capability shifts from development work to standardized delivery (Gbadegeshin et al., 2022), regulatory alignment under compliance uncertainty (Bonnin Roca & O'Sullivan, 2020), and ecosystem integration across fragmented actors and dependencies (Jucevicius et al., 2016; Klitsie et al., 2019). To summarize, VoD and DVoD research says that scaling is not a smooth continuation after technical success but multiple difficult transitions across gaps in resources, overall logic and integration.

2.1.2 Pilotitis

In digital service innovations the transition from an abstract idea to something tangible is typically handled through testing mechanisms such as PoC and pilot projects. These generally represent an early financing and development stage where funding and facilities are required to demonstrate the technical feasibility of a technology that may essentially only yet exist on a small scale (McIntyre, 2014). Scarbrough et al. (2024, p. 1) define pilots as “formally designated, discrete events aimed at generating evidence” regarding the efficacy and implementation of an innovation. These types of projects allow stakeholders to refine designs, prove commercial feasibility, and hedge risks by testing products with limited resources before committing to full operationalization (McIntyre, 2014; Mergel et al., 2022) and ultimately the more resource-intensive phase of scaling (Hartley et al., 2013).

Despite the intended bridging function of pilots, the digital health and public sectors frequently encounter the systemic failure of pilotitis. The problem is characterized by a cycle where innovations are repeatedly tested again in pilots in different places but rarely make it beyond this phase into full scale routine implementation (Bhatia et al., 2020; Egermark et al., 2022; Gerhardt et al., 2015; Huang et al., 2017; Scarbrough et al., 2024, Schiffelers et al., 2025). These small scale digital interventions might demonstrate promise in the pilot phase but fail to scale, effectively “dying” when the original funding runs out (Chamberlain et al., 2021, p. 1). The innovation landscape becomes fragmented with scattered projects that lack the critical mass required for significant evidence generation or sustainable adoption (Egermark et al., 2022). Projects often expire after the initial phase, failing to expand beyond their test sites or integrate into national systems (Gerhardt et al., 2015; Huang et al., 2017).

Chamberlain et al. (2021) argue that pilots often demonstrate a promising intervention, but do not test their sustainability pathway early enough. This increases the likelihood that innovations remain successful pilots but do not convert into routine services. Thus, it can be argued that pilotitis persists because pilot conditions often do not match the

conditions of routine delivery. Scarbrough et al. (2024, p. 2) note that pilots can be "exceptionally well resourced and managed in very controlled and supportive local environments." Chamberlain et al. (2021) make the same point in a practical form, exemplifying that in pilots funded by donors it can be possible to design and run an ideal solution because the project can pay for intensive support, tailored content, and operational work. These are all difficult to sustain at scale. When scaling begins, trade-offs are often required between ideal solution design and affordability, and solutions need to be designed within the limits of sustainability (Chamberlain et al., 2021). Another related issue is integration, and Schiffelers et al. (2025) show that pilots often function as add-ons to existing processes that are additionally supported by temporary adjustments and extra resources. When this is removed, the innovation struggles to survive.

Pilotitis is also closely aligned with the scaling gap that sits after initial validation in DVoD. While the first gap involves the transition from research to a working prototype (McIntyre, 2014), the second valley sits between validation and sustainable, widespread adoption (Egermark et al., 2022). This second gap exists because the requirements for scaling can be said to differ fundamentally from those of piloting. The failure to bridge this gap results in a cycle of re-piloting and starting again and again from zero. Innovators try to prove the solution works in new settings without ever achieving the momentum required for routine (Scarbrough et al., 2024). Thus, moving across requires shifting from a focus on technical feasibility or effectiveness to addressing broader organizational, financial, and regulatory constraints (Scarbrough et al., 2024).

In the context of the public sector, Chamberlain et al. (2021) argue that many pilots are not designed with government procurement policies, interoperability, and operational feasibility at scale in mind. They emphasize that the expectation that a pilot will be directly replicable at scale is unrealistic because of public sector issues such as procurement delays, funding stream alignment, and different components not rolling out in sync (Chamberlain et al., 2021). Pilotitis seems to mostly be a problem in regulated

or investment-heavy environments like the public sector. This is despite the problem recognized in EU funding programs like Horizon Europe and the EIC Accelerator that provide PoC and scaling support (European Commission, 2025; European Innovation Council, 2025).

Across literature regarding pilotitis, the core implication is that piloting supports scaling when the pilot is designed around routinization requirements and when it is used to test and build the integration, resourcing, governance, and evidence needed after the pilot setting ends. The innovation must move from being a stand-alone pilot add-on to being integrated into existing infrastructures and everyday workflows to convert this temporary situation into a routine (Schiffelers et al., 2025). Additionally, there must be a transition from temporary project-based funding to structural financing arrangements that support long-term operations in the limits of affordability and public sector constraints (Schiffelers et al., 2025; Egermark et al., 2022; Chamberlain et al., 2021).

The evidence base of the project must be translated from localized proof from the pilot into decision-relevant validation that can travel across contexts and support later stage implementation and adoption decisions (Scarbrough et al., 2024; Bhatia et al., 2020). Because scaling depends on broader uptake responsibility for sustaining the innovation needs to extend across the actors involved in routine delivery so that the work continues when the pilot setting ends (Currie et al., 2025). Finally, the service must align with user needs and implementation realities, so it works under constraints also and not of only under the pilot conditions (Huang et al., 2017; Schiffelers et al., 2025).

2.1.3 From diffusion to embedding for scaling

DVoD highlights the structural gaps between invention, piloting, and mainstream use. Crossing over these gaps requires a shift from diffusion to scaling. Traditional innovation research often describes the spread of technology as diffusion. It is described by Breaugh et al. (2021) as a passive process where innovations spread over time as people learn

about them. This description does not include the later gap in the second valley of DVoD where innovations can have proven value but still fail to progress (Hartley et al., 2013). To avoid clashing and to make sure of consistent concept use, this thesis uses Breaugh et al.'s (2021) definition of innovation scaling as the process of proactively and deliberately spreading an innovation to increase its impact with replication. This is important here because the transition to routine service does not happen passively because it requires intentional consideration.

Diffusion and adoption should be differentiated from embedding and implementation to clarify what the active scaling would then mean. Scarbrough and Kyratsis (2022) argue that diffusion and implementation should be seen as part of the single process of innovation embedding. It is required for sustaining innovation beyond early adoption. On its own, the mostly passive diffusion (Breaugh et al., 2021) can result in low or partial adoption where an innovation is technically present but not effectively utilized or sustained which in turn leading to weak improvements (Scarbrough & Kyratsis, 2022). It therefore focuses on the depth of innovation implementation within an organization. Damschroder et al. (2009, p. 3) describe it as "the critical gateway between an organizational decision to adopt an intervention and the routine use of that intervention". In short, embedding refers to the institutionalization of the innovation in its intended context, so it becomes a part of organisational routines, practices, and infrastructure (Scarbrough & Kyratsis, 2022).

When innovations are adopted by new actors they often need to be adapted (Breaugh et al., 2021), because a solution that works in one site rarely fits other sites as is (Greenhalgh & Papoutsi, 2019). Denis et al. (Scarbrough & Kyratsis, 2022, p. 239) describe this as keeping the core idea of the innovation stable and at the same time allowing a more flexible "periphery" to be adjusted to local conditions. Pollock and Hyysalo (Scarbrough & Kyratsis, 2022, p. 239) make a similar point with generalization, where solutions are designed around flexible standards so that lessons and adaptations from different implementation sites can be incorporated into the product. Adaptation is

not the opposite of scaling when seen this way but actually one of the mechanisms that makes embedding possible. The focus moves away from passive diffusion and replicating the exact form of the innovation toward replicating more of its impact across contexts by preserving core elements, as Denis et al. (Scarbrough & Kyratsis, 2022, p. 239) describe, while enabling tailoring by context (Greenhalgh & Papoutsi, 2019). This is also consistent with Breaugh et al.'s (2021, p. 15) definition of innovation scaling.

This is important when innovations move from testing into mainstream use, because what works in a protected pilot often needs to be reshaped for real settings with different constraints and expectations (Scarbrough et al., 2024). Martin et al. (2013) say that once a service starts moving toward routine use, sustainability is an ongoing process that requires continuous effort and adjustment. The work does not end after a successful pilot. The service must keep demonstrating and communicating its value to the actors who make funding and continuation decisions. Especially in environments where priorities and expectations can change quickly. They also show that sustaining a service depends on building and maintaining support with relevant stakeholders, aligning the service with existing organisational arrangements, and staying flexible in response to changes (Martin et al., 2013).

2.2 Business model lens

This section develops the business model as a theoretical lens for understanding how ventures structure a usable offering and sustain delivery over time. The business model is defined with value proposition, value creation and delivery, and value capture. It then explains why these elements often require reconfiguration beyond early validation and identifies the scaling-relevant dimensions within the business model.

2.2.1 Defining business models

The business model is a useful unit of analysis for explaining how a venture turns an innovation into economic value. As Teece (2010, p. 172) describes it, a business model specifies the design or architecture of how value is created, delivered, and captured. It includes how the firm delivers value to customers, induces payment, and converts revenues into profit. In this sense, the business model is a conceptual model for a business (Teece, 2010). According to some of the founding work on what a business model is (Chesbrough & Rosenbloom, 2002), it is sitting between two domains. One is technical and includes inputs, performance, and feasibility, and the other is the economic domain that has outputs, customer value, and profit. In this view, an innovation's value just remains latent unless it is commercialized with a working business model. Similarly, Amit and Zott (2001) similarly argue that a business model describes how a venture structures its transactions and activities to create and capture value, reinforcing the point that technical functionality alone does not determine whether an innovation becomes economically viable. This can be seen as especially relevant in VoD, as even a working solution can stall if the venture cannot specify a revenue architecture that fits market realities (Chesbrough & Rosenbloom, 2002).

This relevance becomes clearer in how VoD is described in commercialization research. In VoD, technical validation can be achieved with the commercialization problem still remaining unresolved. McIntyre (2014) characterizes this situation as a firm having a working prototype that is not yet developed enough to earn money through commercial sales. This is creating a continued need for resources until the offering becomes independent. Similarly, Ellwood et al. (2022) argue that progress through VoD becomes visible when the venture can articulate a credible commercial proposition that attracts the next major investment. These views reinforce that it is not enough that the solution works technically but the venture must also have a sustainable business model.

To operationalize the concept in this thesis, the business model is decomposed into three core components, drawing on Teece (2010, p. 172) and Wirtz et al. (2016, p. 38): *the*

value proposition, which specifies what problem the offering solves and why it is valuable to the customer (Wirtz et al., 2016, p. 38), *value creation and delivery*, which describes how the organization mobilizes resources and activities to produce and deliver that value (Teece, 2010, p. 172), and *value capture*, which specifies how revenues are generated relative to the cost structure so that the venture can remain financially viable (Teece, 2010, p. 172; Wirtz et al., 2016, p. 38).

2.2.2 Business model reconfiguration across phases

Research suggests that moving from a pilot to routine service delivery usually requires reconfiguration, because the first business model is rarely the final one. Chesbrough and Rosenbloom (2002) describe the early business model as more of a prototype strategy that changes with adaptation when the venture learns about the offering and the economic environment. Teece (2010) describes the business model as a conceptual architecture that is built on assumptions about customers and about how revenues and costs will behave and argues that these assumptions must be tested and adjusted through trial and error. This dynamic view is also reflected in Wirtz et al. (2016), who argue that a business model should be understood as something that may require evolution or innovation over time to respond to internal and external changes. All of these perspectives seem to imply that business model configuration needs to be dynamic and adjustable over different phases and institutional settings. Core design choices should be set before the pilot but they are refined as uncertainty declines and any constraints become clear.

A simple way to describe what changes is Baden-Fuller and Haefliger's (2013) distinction between project-based and scale-based business model configurations. Project based "taxi" models create value through tailored work for specific clients, while scale based "bus" models create value through more standardized offerings and repeatable routines (Baden-Fuller & Haefliger, 2013, p. 421). This also fits that the business model is an organizational and financial architecture (Teece, 2010) and redesigning that architecture

so value can be created repeatedly rather than through individual arrangements is required for scaling (Amit & Zott, 2001).

Reconfiguration is also driven by value capture limits that affect what kind of business model is possible. Teece (2010) explains that weak ability to capture returns and value spillovers can create market failures in innovation, meaning private business models may not create enough incentives for investment unless these innovators package together complementary elements or rely on public funding or other mechanisms. McIntyre (2014) makes a closely related point on that even when the prototype works the venture still needs a value capture logic that can keep delivery through sales or other stable income, otherwise it remains stuck in a technically validated but financially fragile configuration. Both reinforce Chesbrough and Rosenbloom's (2002) point on that the business model gives the logic that connects what a solution can do with how it generates economic value. They add that without this logic value remains unrealized regardless of the quality of the technology in the offering.

2.2.3 Scaling-relevant business model dimensions

Building on the definitions above, this section clarifies what becomes important in business model design as ventures attempt to scale. The three components of the business model provide a structured way to explain why a working solution does not automatically translate into sustained service delivery at scale.

From a value proposition perspective, scaling draws attention to how the offering is framed. Specifically, how it's framed in terms of customer benefit and economic utility. Also, if that framing can be carried past the early validation. In early developmental phases, the proposition may stay broad or provisional. It's often framed around novelty and feasibility to attract early support (Chesbrough & Rosenbloom, 2002). Moving toward wider adoption typically requires refining the proposition so that it speaks to a defined customer segment with demonstrated economic utility (Teece, 2010). Baden-

Fuller and Haefliger (2013) also point that a business model must identify the customer and engage with customer needs, which supports the idea that scaling requires the value proposition to move from a broadly promising concept toward a clearer and more transferable articulation. Thus, a central design issue is how far the value proposition can be standardized without losing the core benefit validated in early use. This matters for scaling. A value proposition that depends on extensive tailoring may work in a pilot setting but it struggles to support repeatable uptake in different contexts.

In value creation and delivery, the key issue is how delivery is organized. Does it depend on flexible ad hoc effort or can it be executed through repeatable routines with lower marginal effort. In early validation ventures often rely on flexible processes to solve problems as they arise. There the priory is effectiveness over cost-efficiency (Baden-Fuller & Haefliger, 2013). For scaling, attention shifts. It is toward repeatable processes and governance structures that reduce friction and lower the marginal cost of delivery (Amit & Zott, 2001). Another design issue concerns boundary choices, meaning which activities remain internal and which delivery depends on outside the venture, because delivering at scale may require access to assets the venture does not control directly (Teece, 2010). This also matters for scaling as a delivery architecture optimized for flexible problem-solving may lack the structural stability and process efficiency needed to support routine, higher-volume operations.

Value capture is where the VoD often becomes tangible: a working solution exists, but the logic for revenue and cost for sustained operation is still underdefined. Ventures can have a working prototype but miss a more developed way to create commercial sales, relying instead on external financing or development funding (McIntyre, 2014). Provisional arrangements like that can hide the full cost of delivery and delay the formation of a revenue model that is based on the market (Chesbrough & Rosenbloom, 2002). As Ellwood et al. (2022) point out, getting growth funding is a competitive selection problem where investors compare ventures against many alternative projects, making a convincing value capture logic central to progressing through the VoD.

Therefore to progress, the business model must specify who pays, how they pay, and how the venture appropriates a portion of the value created against the costs of operation (Teece, 2010). This often has navigating the compromise between value creation and capture. Because a venture might generate significant utility for users but fail to capture revenue if the design is missing the mechanisms to enforce payment or protect returns from imitation (Teece, 2010). For scaling once again, technical validation does not guarantee economic sustainability if the business model is missing a revenue logic that is capable of covering operating costs longer term.

2.3 Ecosystem lens

This shifts attention from firm-level performance to the production of a system-level outcome. There actors hold different roles, depend on each other, and must be aligned for the value proposition to materialize. Additionally, scaling often depends on whether the surrounding actor set has opportunities for value capture. The main key questions become coordination questions: who needs to do what, in what order, and under what alignment structure for the system to function.

2.3.1 Defining ecosystems

In this thesis, an ecosystem is defined as the “alignment structure of the multilateral set of partners that need to interact in order for a focal value proposition to materialize” (Adner, 2017, p. 40). This means an ecosystem is a specific set of activities and partners that must fit together for the value proposition to exist. In systemic innovations, the main value often is in the linkages between subsystems, which means actors within the influence domain must take action to adjust to changes (Lavikka et al., 2021). Ecosystem boundaries can be narrow or broad depending on the value proposition: the ecosystem is defined by relevance to the outcome in the system level and not by size (Adner, 2017). This distinguishes ecosystems from networks. Networks focus on patterns of connectivity and actor ties, but ecosystems are defined by the specific alignment

structure and the interdependence of partners required to create a coherent broader level outcome that no single participant could generate alone (Adner, 2017; Autio & Thomas, 2022). In addition to this, unlike supply chains that usually rely on formal supplier contracts, ecosystems can coordinate diverse participants with methods and mechanisms like role definitions and modular architectures (Autio & Thomas, 2022).

A related distinction is between knowledge ecosystems and business ecosystems. Knowledge ecosystems focus on generating new knowledge and often are around universities and research organizations, business ecosystems focus on delivering customer value through a network of companies (Autio & Thomas, 2022; Clarysse et al., 2014). While knowledge ecosystems have upstream value creation, business ecosystems have value creation that is not linear where interdependent companies deliver integrated solutions to end users (Clarysse et al., 2014). The formation of a business ecosystem does not automatically follow a knowledge ecosystem, because they operate under different logics. These are e.g., the different norms of science and business, and tend to involve different anchor tenants (Ellwood et al., 2020; Jucevicius et al., 2016).

Coordination in ecosystems also differs from traditional ones. Ecosystems coordinate multiple stakeholders to deliver a broader system output without relying on hierarchical control or direct supplier contracts (Autio & Thomas, 2022). Coordination relies on alignment structures, roles, and complementarities (Adner, 2017). When multiple actors are involved coordination becomes more complex because actors may have different motivations that need to be managed through innovation processes (Ellwood et al., 2022). There is also interdependence, meaning that if one partner fails to align the value proposition can weaken for all (Adner, 2017). For this reason, ecosystem leaders often orchestrate which means intentional actions to spark and manage innovation processes and align incentives with actors (Linde et al., 2021).

2.3.2 Scaling-relevant ecosystem dimensions

The *alignment* of partners regarding their roles and the flow of activities in the ecosystem is needed for scaling (Adner, 2017). Adner (2017) mentions this as the amount of shared agreement in these roles and positions, so that the actors are willing to do the activities required for the value proposition to become real. Value capture from these kinds of broader innovations is not always however evenly shared out because it can create barriers if the required changes give switching costs on any specific partners (Lavikka et al., 2021). This kind of alignment needs coordinating for the value creation, e.g., new resource building or defining standards that help communicate the value proposition to relevant actors (Stehn & Erikshammar, 2025).

For ecosystem leaders, scaling involves *orchestration*. It means intended work to start and manage innovation processes (Linde et al., 2021). This can include building partnerships identifying necessary partners and getting rid of unproductive relationships (Linde et al., 2021). Developing digital solutions across organizational boundaries can create coordination tensions that complicate collaboration and slow scaling progress (Garcia Martin et al., 2024). Coordination that keeps partners aligned and incentivized for the common good while properly structuring value creation interactions supports scaling (Linde et al., 2021). Adner (2017, p. 42) also argues that successful alignment must have "compatible incentives and motives" alongside with how activities are actually structured. To achieve this coordination often needs defining roles and finding a balance between stability and change in the overall ecosystem structure (Autio & Thomas, 2022).

Finally, scaling can benefit from *intermediation*. Stehn and Erikshammar (2025, p. 8) describe a capability of a “digital broker” that supports knowledge sharing and manages tensions in projects and technology development. Intermediaries can also reduce tensions in digital, like limits in data sharing and cybersecurity concerns that affect transparency (Garcia Martin et al., 2024). Effective intermediation can include phases ranging from establishing foundations and creating early projects to orchestrating

collaboration and supporting broader implementation (Garcia Martin et al., 2024). This brokering role also directly addresses the need identified by Clarysse et al. (2014, p. 34) to develop "boundary spanners" that can bridge the gap between ecosystem elements to bring larger actors into the ecosystem. This also could directly support the development of a shared value proposition and learning over projects in ways that allow the needed value capture for scaling (Stehn & Erikshammar, 2025).

2.4 Institutional context

This section introduces an institutional lens to explain why the B2G environment has a specifically constraining character and how that character shapes which business model and ecosystem configurations are feasible in practice. To understand why digital service innovations so often stall between the pilot phase and routine service delivery, the section examines the environmental pressures, structural constraints, and institutional forces that shape change attempts within public organizations. Earlier sections established the scaling gap in terms of venture configuration. That framing cannot alone explain why technically viable and resourced innovations still fail to progress. The section proceeds by defining the B2G environment as institutionally complex, contextualizing the specific scaling constraints this creates, and then deriving the institutional dimensions used in the empirical analysis.

2.4.1 Defining the institutional context as a B2G business environment

For new discoveries to be commercialized successfully, ventures need to navigate organizations with differing missions, incentives, and operating assumptions (Earle et al., 2019). Institutional theory can be used to get the broader view needed in order to describe how organizational action is shaped by rules, routines, and socially embedded expectations (Aksom & Vakulenko, 2024). This is relevant in settings like the B2G where public entities are restricted with internal routines and by a wider regulatory and governance environment. So, the adoption and continuation of new solutions cannot be

understood just in terms of productive efficiency (Aksom & Vakulenko, 2024; Obwegeser & Müller, 2018). From this perspective, public procurement regulations and control mechanisms can narrow buyers' room for action (Obwegeser & Müller, 2018) and importantly also complicate the progression of startups and innovations across the VoD.

In this broader institutional lens, institutional logics help explain how actors make sense of organizational situations. These institutional logics refer to socially constructed patterns of practices, values, beliefs, and rules that provide meaning and change action (Bitektine & Song, 2023). For B2G contexts what is appropriate is not determined by just technical promise. Judgments are instead filtered through the logics of the field (Wu et al., 2023; Vickers et al., 2017). For startups that want to avoid pilotitis and start scaling, this means that scaling is also an institutional issue.

Institutional complexity also defines the B2G environment well. It arises when organizations confront multiple logics at the same time and these logics may order different or incompatible courses of action (Greenwood et al., 2011). For example in the public sector public service delivery may combine state, market, and civil society logics and is not being guided by one clear and coherent logic alone (Vickers et al., 2017). Under conditions like that, organizations can face tensions and coordination difficulties when they try to respond to multiple expectations at once (Greenwood et al., 2011; Wu et al., 2023).

All these complexities are visible in public service delivery which often happens in hybrid institutional environments. Vickers et al. (2017) show that state, market, and civil society logics may coexist and interact in public service settings. More conventional private markets are for this reason institutionally different when comparing to B2G environments. Also, public organizations usually pursue broader objectives than private firms (Arlbjørn & Freytag, 2012). Additionally, studies from the supplier-side bring up more practical barriers e.g., limited interaction with procurers, over-specified tenders, low procurer competence, and weak risk management (Uyarra et al., 2014). B2G is an

environment with rules that change by market exchange, governance structures and objectives, and multiple institutional logics.

2.4.2 Deepening and contextualizing scaling constraints in B2G

Suchman (1995, p. 574) defines legitimacy as a assumption that actions are desirable, proper, or appropriate within a socially constructed system of norms, values, beliefs, and definitions. When this definition is applied to digital service innovations, it suggests that progression also depends on audiences that are relevant seeing continuation as appropriate. Organizational action evaluations from an institutional logics perspective are influenced by the logic that is important in a particular situation, which is because different logics make the different norms and criteria more available for judgment (Bitektine & Song, 2023). This is important in hybrid public service settings, which are where state, market, and civil society logics can coexist and interact (Vickers et al., 2017). This is why legitimacy may be harder to secure consistently across decision contexts.

Different institutional logics draw attention to different values and criteria which makes some considerations more important than others in the assessment process (Bitektine & Song, 2023). When there is institutional complexity organizations may also face multiple and partly incompatible expectations about what is appropriate action (Greenwood et al., 2011). Since legitimacy depends on if actions are perceived as appropriate within prevailing norms and beliefs (Suchman, 1995), scaling may become more difficult when pilot outcomes are not framed in a way that aligns with the criteria used in evaluating by those deciding on broader continuation (Bitektine & Song, 2023).

One other main hinderance in this environment moreover is that public procurement is organized with formal rules and procedures that can limit interaction between buyers and suppliers (Melander & Pazirandeh Arvidsson, 2020). For innovation long-term interaction and collaboration are often important for developing and refining new solutions, yet public procurement regulations may just restrict exactly these kinds of

relationships (Melander & Pazirandeh Arvidsson, 2020). Formalized tendering procedures may also make mutual and co-learning more difficult (Arlbjørn & Freytag, 2012). From the supplier perspective these difficulties can be seen in practice because lack of interaction with procuring organizations and reliance on tenders with overly specific criteria are identified as major barriers to innovation in public procurement (Uyarra et al., 2014). Chamberlain et al. (2021) reinforce this directly by pointing out that few digital development pilots are designed with government procurement policies and overall operational feasibility at scale in mind. This disconnect makes the transition from piloting to routine delivery structurally difficult for digital services that may depend on continued adaptation between the solution and the organization.

Another important tension in the literature is that public procurement is often shaped by competing policy priorities instead of one clear and stable objective. In the UK case Pickernell et al. (2011) examine, they show that cost efficiency and value for money compete with objectives that position procurement as a tool for innovation policy and wider economic development. Similarly, Akenroye et al. (2024) in their study identify competing priorities in policies, cautious culture, and an attitude great for large suppliers as influential barriers to SME participation in public procurement. This is relevant because smaller innovative companies may contribute to public innovation procurement, but their limited resources and the large size of some tenders still can make participation difficult (Saastamoinen et al., 2018). A pilot can as a result enter a procurement environment where the broader continuation they want depends also on how decision-making actors are able to balance e.g., the risk in innovation and supplier fit with efficiency.

Smaller firms may also face structural difficulties when moving from pilot projects to larger contracts because public procurement often involves administrative burdens and resource demands that are harder for SMEs than for larger suppliers. Akenroye et al. (2024) identify bureaucracy as one of the influential barriers to SME participation and note that government structures can create bureaucracy and hierarchical barriers.

Saastamoinen et al. (2018) likewise note that public procurement of innovations is often characterized by large contracts and that small firms may lack the resources to compete for tenders such as that. This problem from the supplier perspective too is consistent with wider evidence that contract size and tender requirements can operate as barriers to innovation-oriented public procurement participation (Uyarra et al., 2014). These conditions together may suggest that even when a pilot shows promise, progression to broader contracting may remain difficult for smaller suppliers. At the same time, these constraints should not be understood only as fixed structural conditions. Breaugh et al. (2021) note that the legislative environment is a changeable modal factor that exists in a dynamic relationship with scaling efforts. They say that it is not an antecedent although frequently identified as an external barrier in scaling literature. Therefore, the institutional environment is treated in this study as multiple pressures that the venture's configuration is actively engaging with.

2.4.3 Scaling-relevant institutional dimensions

Synthesizing and building on the constraints identified above, this section derives three analytical dimensions that help explain how the institutional B2G environment shapes the feasibility of scaling digital service innovations beyond the pilot stage.

The first lever, *interaction conditions*, focuses how formalized procurement rules shape the scope and quality of buyer–supplier interaction through which new solutions can be refined, adjusted, and aligned with organizational needs. This lever is relevant because the literature suggests that innovation in public procurement is constrained when regulations restrict interaction and long-term collaboration between buyers and suppliers (Melander & Pazirandeh Arvidsson, 2020). Relatedly, formalized tendering procedures may hinder the possibility to mutually learn and gain value (Arlbjørn & Freytag, 2012), while suppliers identify lack of interaction with procuring organizations and reliance on over-specified tenders as major barriers to innovation in public procurement (Uyarra et al., 2014). In the transition from pilot to scaled operation this is

also relevant, as Chamberlain et al. (2021) note that few digital development pilots are designed in mind with government procurement policies and operations at scale.

The second lever, *legitimacy conditions*, focuses on continuation depending on if the pilot and its outcomes are judged to be appropriate and credible, also worthwhile under the evaluative criteria of the institutional environment. According to Suchman (1995), legitimacy depends on if actions are seen as desirable, proper, or appropriate in a socially constructed system of norms and beliefs. From an institutional logics perspective, these are built by the logic that is activated in the situation, because different logics make different norms and criteria more available in evaluation (Bitektine & Song, 2023). Organizations may therefore face multiple and partly incompatible expectations about appropriate action in settings that are known for their institutional complexity (Greenwood et al., 2011). In B2G contexts, this means that scaling may become difficult if pilot outcomes are not framed in a way that fits the criteria emphasized by those deciding on broader continuation.

The third lever, *resourcing conditions*, concerns whether smaller ventures have the financial and organizational capacity to move from pilot stage to broader contracts in the B2G environment. Akenroye et al. (2024) identify bureaucracy, competing policy priorities, cautious culture, and a pro-large supplier attitude as influential barriers to SME participation in public procurement. Saastamoinen et al. (2018) likewise state that public procurement of innovations often has large contracts and that small firms may lack the resources to compete for such tenders. Contract size and demanding tender requirements can thus make broader procurement participation more difficult even where a smaller firm has demonstrated innovation potential (Uyarra et al., 2014). These constraints at the same time should not be treated as fixed, as Breaugh et al. (2021) suggest that these external conditions should be understood as changing factors that are in a more dynamic relationship with efforts in scaling. The institutional environment is treated here as pressures that ventures must engage with.

2.5 Synthesis and theoretical framework

The previous sections have developed three theoretical lenses for examining scaling digital service innovations beyond the pilot stage. Scaling is treated here as an outcome of the configuration between business model design, ecosystem coordination, and institutional context. Business model design explains how the venture structures its offering and the economic logic that must sustain it. Ecosystem coordination explains how the surrounding actor set must be aligned for value delivery to become possible in practice. Institutional context explains how the broader environment shapes what configurations are feasible and supportable. Using a GovTech pilot as the empirical setting the thesis addresses the Valley of Death and the risk of pilotitis by analyzing how this configuration structures what is feasible beyond pilots. The three lenses form a configurational framework together where scaling is not as the outcome of any single factor but as the result of how all of these elements interact (see Figure 2).

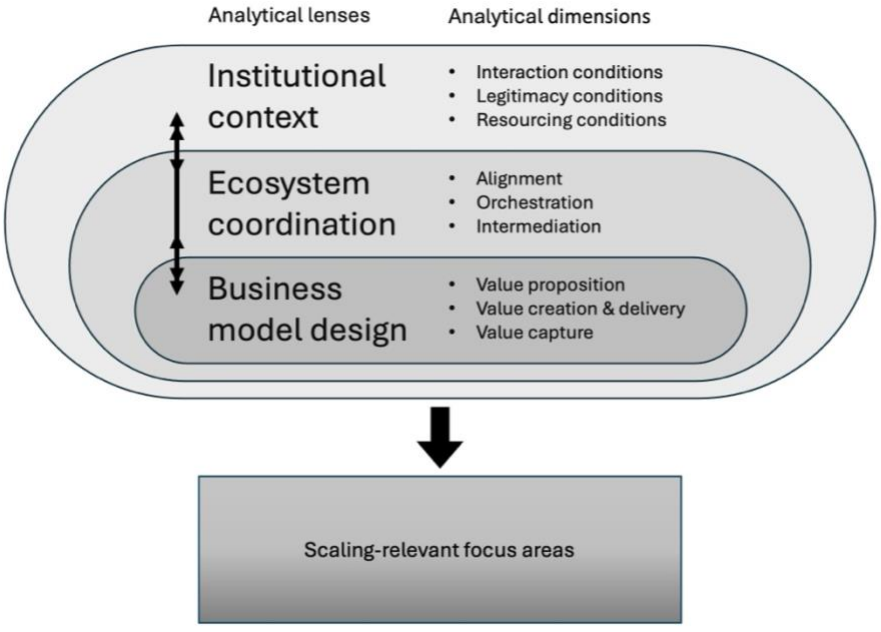

Figure 2. Theoretical framework: Lenses and dimensions for examining scaling

As shown in Figure 2, the relationship between the lenses is nested and thus provides micro, meso, and macro levels of the configuration. The micro level is formed by business model design, because this is the nearest level to the venture and concerns how the focal venture itself structures value. The meso level is formed by ecosystem coordination, because delivering that value depends on the roles, contributions, and alignment of other actors that the venture cannot control alone. The macro level is formed by institutional context, because public-sector rules, legitimacy pressures, and resourcing conditions define the wider space within which both the venture and the ecosystem must operate. The double-headed arrows between the layers are there to indicate that these relationships are mutually affecting each other. The venture's value proposition influences what kinds of actors are needed around it, the available ecosystem influences what kinds of business model configurations are realistic, and the institutional environment shapes which combinations of both can be sustained in practice. The framework therefore treats scaling as a layered configuration in which micro-, meso-, and macro-level conditions continuously interact.

The business model lens contributes to the framework at the micro level by providing a structured view of how the venture creates, delivers, and captures value. The three dimensions of value proposition, value creation and delivery, and value capture identify the internal design choices that scaling requires. Moving beyond a pilot setting typically involves refining the value proposition so that it can travel across contexts, redesigning delivery so that it can be executed through repeatable routines and not ad hoc effort and developing a value capture logic that can support sustained operation rather than depend on pilot-specific funding. In this sense, the business model lens explains why a solution that works in a protected pilot does not automatically become a viable service at greater scale. It captures the internal logic of the venture, but it also points to the limits of looking only at the venture itself.

That limit leads directly to the meso level. Even a well-configured venture cannot scale if the broader actor set required for delivery is not organized in a workable way. The

ecosystem lens therefore shifts attention from firm-level design to the system-level conditions required for that design to function. Its three dimensions of alignment, orchestration, and intermediation capture how the surrounding actor network must be organized. Alignment concerns whether the necessary partners agree on their roles and are willing to perform them. Orchestration concerns the deliberate work required to maintain that agreement and resolve coordination tensions as conditions change. Intermediation concerns whether actors capable of brokering knowledge, managing dependencies, and connecting different parts of the ecosystem are present. In the B2G context, where multiple actors across organizational and sectoral boundaries must collaborate, these meso-level coordination conditions become especially visible when a pilot attempts to move toward routine delivery.

However, the meso level does not operate in a vacuum either. The way ecosystems can be built and maintained is shaped by the wider public-sector environment in which they are embedded. This brings the analysis to the macro level, where the institutional context lens explains why the B2G environment has the specific constraining character it does. The three dimensions of interaction conditions, legitimacy conditions, and resourcing conditions capture the environmental pressures that shape what the venture and its ecosystem can actually do. Interaction conditions concern how procurement rules affect the scope and continuity of buyer-supplier collaboration. Legitimacy conditions concern how continuation depends on being judged appropriate, credible, and worthwhile under the evaluative criteria that are in the institutional environment. Resourcing conditions concern whether smaller ventures have the financial and organizational capacity to participate in broader contracts after the pilot ends. These macro-level conditions explain why institutional pressures make a dimension for the scaling problem and also show why they are not reducible to funding gaps or technical barriers alone.

The three levels taken together form one analytical whole. The micro level captures the venture's internal design logic, the meso level captures the relational and coordination

conditions required for delivery, and the macro level captures the environmental pressures that limit what is feasible and supportable. This means that the framework is not additive. It does not assume that business model design, ecosystem coordination, and institutional context independently produce scaling, because none of these perspectives is sufficient on its own. A venture can have a well-designed business model that cannot be delivered without the right actors in place. Its fragmented ecosystem can prevent an otherwise promising value proposition from working. Restrictive procurement rules or demanding legitimacy requirements may block business model development and ecosystem stabilization. Thus, the main task is to examine how the three levels interact in combination.

The nine analytical dimensions across the three lenses are used here as guides that are theory-informed for structuring the empirical section. They are not hypotheses or predetermined findings. They direct attention toward the aspects of the pilot's configuration that the literature suggests are most consequential for moving past the early validation while still remaining open to the specific ways those appear in practice. The downward arrow shown in Figure 2 is the output logic of the analysis. The purpose of applying the framework is not to evaluate the dimensions themselves. The purpose is to identify with reading of all the levels combined which aspects of the current pilot configuration influence how digital service innovations can be scaled past the pilot stage. That is why the scaling-relevant focus areas are not defined in advance. They emerge from how the three lenses interact when applied to the empirical material. The framework serves as a bridge between theory and empirics, structuring the analysis while leaving the final synthesis open to what the case reveals.

3 Methodology and research method

This chapter presents the methodological choices of the thesis. It explains the research design and overall approach, describes how the empirical data were collected and from whom, and outlines the procedure used to analyze the material. The chapter also addresses how the trustworthiness of the study was ensured and introduces the empirical case of Company X.

3.1 Research design and approach

This thesis employs a qualitative research design, which is appropriate when the aim is to understand phenomena through the perspectives and meanings of the actors involved (Creswell & Poth, 2018, p. 5). The study is a single-case study of Company X which is a GovTech startup currently piloting an AI-assisted service in a Finnish wellbeing services county. Case studies are well suited for developing an in-depth understanding of complex issues in their real settings (Crowe et al., 2011). The case provides an example of the transition challenges associated with the Valley of Death and pilotitis in a currently live B2G pilot context. This makes it possible to examine how the configuration of business model design, ecosystem coordination, and institutional context interact.

The study takes a pragmatist stance. It is treating theory as an analytical tool and not as a set of fixed criteria to be tested. This is reflected in the use of an abductive research logic where the theoretical framework developed in the literature review guides what is examined in the empirical material and in what order while the empirical material still builds how findings are interpreted (Dubois & Gadde, 2002). The analytical dimensions extracted from the three theoretical lenses provide a clear and structured starting point for reading the case. The analysis still remains open to patterns that emerge from the data and does not force the case into predetermined categories or topics.

RQ1 was approached primarily through a structured literature review and conceptual synthesis that drawed on prior academic research and peer-reviewed journal articles.

Literature was identified using software Publish or Perish and the Google Scholar search engine. Searches were performed using keywords related to the topics of Valley of Death, pilotitis, business models, ecosystems, scaling, GovTech, B2G, public procurement, and digital service innovation, including relevant combinations of these terms. The search process was iterative. It was also complemented by backward and forward searching to capture influential contributions across the literature streams relevant to the thesis. Sources were selected based on their relevance to the research questions and their conceptual or empirical contribution. The literature was synthesized to clarify most important concepts, build the framework, and derive analytical dimensions guiding the empirical research. RQ2 was addressed through an empirical single-case study described in the following sections. The final synthesis of RQ1 was completed later by combining the literature-based framework with the empirical findings.

3.2 Data collection and description of the data

The primary data consisted of five semi-structured interviews conducted between February and March 2026. This format was chosen because it allows for open-ended questions that leave room for additional questions to emerge through dialogue which makes it well suited to exploring experiences and perspectives in depth (DiCicco-Bloom & Crabtree, 2006). The interviews were conducted online via Microsoft Teams and were all held in Finnish, as this was the native language of all participants and allowed for more natural and precise expression on topics especially those involving procurement law, organizational practice, and other sector specific terminology.

Participants were selected through purposive sampling which means they were chosen based on their relevance to the case and their capacity to provide information that is rich on the phenomenon instead of for statistical representativeness (DiCicco-Bloom & Crabtree, 2006). The main aim here was to capture the perspectives of actors positioned at different points around the pilot. This including the case company itself, the procuring organization, and external actors with relevant expertise. Access to participants was

obtained through mutual connections. Four of the five participants are directly involved in or connected to the case company or pilot. P4 is an external expert who is part of the case company's broader professional network but is not directly involved. This interview was included to provide an informed perspective on public sector dynamics and typical scaling pathways, and how these dynamics tend to operate in practice. This supported the study by strengthening the contextual understanding needed to interpret the case configuration within the wider institutional environment. Table 1. describes the five participants, their roles, interview dates, and interview lengths.

Participant	Role	Date	Length	Mode
P1	Case company founder	11.2.2026	42 min	Teams video call
P2	Public sector procurement lead	20.2.2026	47 min	Teams video call
P3	Public sector service area lead	23.2.2026	50 min	Teams video call
P4	External expert	18.2.2026	48 min	Teams video call
P5	Case company board member	15.3.2026	43 min	Teams video call

Table 1. Summary of Conducted Interviews

Interview themes and structures were informed by the theoretical framework, which ensured alignment with the thesis focus on configuration and feasibility beyond pilots, while leaving room for respondents to introduce case-specific issues, constraints, and dependencies most important in their views. While the interviews were based around the same structure, some of each interview's questions were tailored to the participant's role (see Appendix 1). In each interview, questions were organized around our three analytical lenses and adapted to reflect the specific perspective and knowledge that each actor was expected to bring to the conversation. To complement the interview material and strengthen accuracy especially around terminology, documents such as a pitch deck provided by the case company were analyzed as documentary sources. Documentary material can provide background, contextual and verifying evidence that helps support the credibility of qualitative findings (Shenton, 2004). Together, the interviews and documents provided the empirical basis for describing how the pilot is configured across

business model design and ecosystem coordination, how the institutional context shapes that configuration, and what this implies for feasibility beyond the pilot stage.

3.3 Data analysis method

The interview data was analyzed using qualitative content analysis. Qualitative content analysis is a systematic method for identifying, coding, and categorizing patterns in textual data (Elo & Kyngäs, 2008), and it was appropriate for this study because the aim was to develop a structured understanding of how the pilot's configuration appears across multiple actor perspectives.

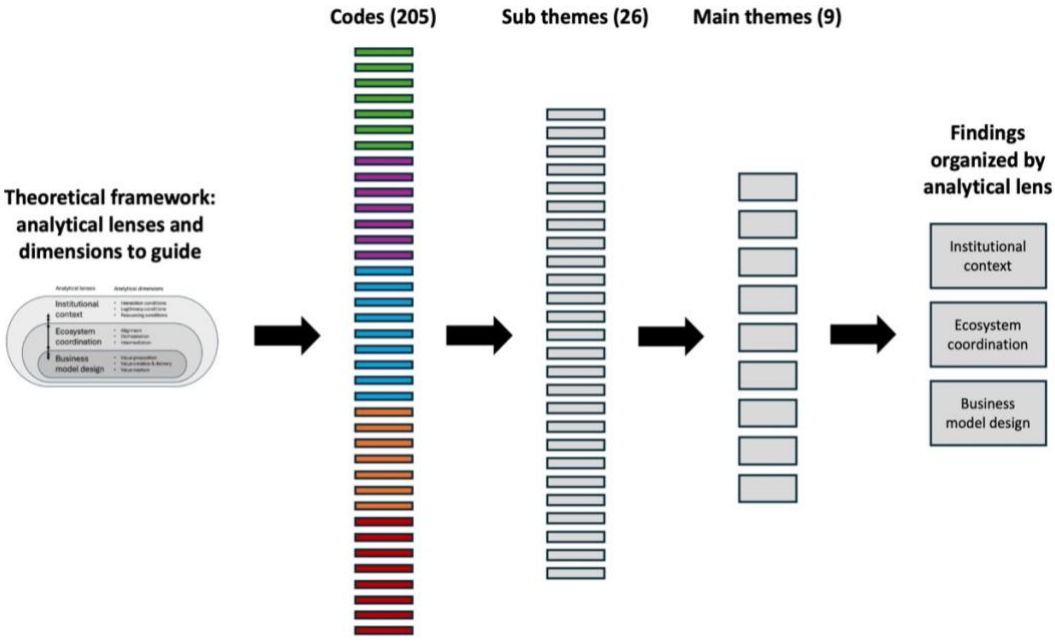

Figure 3. Data Analysis Process Guided by the Analytical Lenses

The analysis proceeded in three stages, as illustrated in Figure 3. In the first stage, all interview recordings were transcribed, fully anonymized, and openly coded in Finnish. The initial coding remained close to the interviewees' own wording, so that each code captured the substantive meaning of a passage as expressed in the material. This stage

produced 205 codes across the five interviews. To maintain traceability throughout the analysis the codes from each interview were visually distinguished by colors.

In the second stage, the codes were compared, then grouped, and clustered into higher categories with examining similarities within the material. The analysis followed an abductive logic at this stage: the theoretical framework guided what kinds of issues were looked for in the material, while the coding and theme formation remained open to patterns emerging from the interview data itself. The three analytical lenses and their nine analytical dimensions were used as concepts that directed attention to scaling-relevant aspects of the case. The codes were organized into 26 subthemes and 9 main themes with this process. Theme formation was therefore empirically grounded while remaining still mostly theory-guided, because the framework helped indicate which kinds of patterns were analytically important while the actual content of the themes was derived from the interview material.

In the third stage, the main themes were reviewed in relation to the analytical framework and used to structure the findings under the three analytical lenses. Where a theme touched on more than one lens, it was placed according to its main emphasis to avoid unnecessary overlap. This thematic structure guided the organization of Chapter 4 under the three analytical lenses. To make the coding process more transparent, Appendix 2 presents an analytical coding summary by lens and dimension, showcasing the main themes together with selected examples of subthemes and codes to exemplify how the empirical material was structured and what the process looked like.

3.4 Validity and trustworthiness of the study

The trustworthiness of this study is assessed through four criteria: credibility, transferability, dependability, and confirmability (Shenton, 2004). Transparency is the most important principle that overarches in assessing qualitative trustworthiness, and in qualitative research the researcher is often the main research instrument (Adler, 2022).

This framing fits well here because the thesis examines a current phenomenon in its real context through a qualitative case study design (Halinen & Törnroos, 2005).

Credibility concerns whether the study presents a believable and accurate picture of the phenomenon or issue being studied (Shenton, 2004). Shenton (2004) states that triangulation through different informants and supporting documents can strengthen credibility. Credibility was supported in this study with interviewing actors in different positions around the same pilot. These participants were selected because they were well placed to provide informative perspectives on the pilot and the environment it is operating in. Together, the interview material, supported by the documentary sources described in Section 3.2, provided a sufficiently broad basis for interpreting the case from multiple relevant positions. Pratt (2009) argues that qualitative claims become more convincing when researchers show data and maintain a clear link between data and interpretation. The findings chapter includes therefore interview quotations and aims to clearly show how the themes were formed from the interview material.

Transferability is the extent to where findings may be applied to other situations. In qualitative research, this depends on providing sufficient contextual information so that the reader can make that judgment (Shenton, 2004). Halinen and Törnroos (2005) also point out that network studies are closely tied to their context and cannot be understood separately from it. This thesis does not claim statistical generalizability for that reason. Instead of that it provides sufficient detail about the Finnish wellbeing services county setting, the pilot, the actor formation, and the continuation conditions so that readers can assess whether the findings are relevant to similar contexts.

Dependability examines if the research process is described in detail for the reader to follow how the study was conducted and also repeat the procedure in practice (Shenton, 2004). Detailed reporting of the research design and data gathering is required to do this (Shenton, 2004; Adler, 2022). Pratt (2009) adds that studies where empirical data is based in interviews should include the interview questions as an appendix.

Dependability is strengthened through systematic description of the case design, participant roles, a common interview structure for them, data collection process, and the detailed description of the data analysis procedure. The appendix of the interview structure, the data analysis figure, and the analytical coding appendix make the research process more traceable and transparent to the reader. Also, this level of methodological detail allows the reader to examine how the presented interpretations were reached and if a similar reading of the material would be plausible.

Confirmability concerns whether the findings arise from the data and not from the researcher's personal assumptions or opinions (Shenton, 2004). Confirmability is strengthened when the study uses more than one source of evidence and when the research process is described clearly enough for the reader to follow it step by step (Shenton, 2004). Qualitative research must be transparent in both theory and method (Adler, 2022). Confirmability was strengthened by making the theoretical framework explicit, by showing how the empirical material was coded and grouped into themes, and by documenting how the findings were organized under the three analytical lenses. Pratt (2009) argues that qualitative work should make its chain of evidence clear. Interview quotations, coding summaries, appendices, and analytical figures keep the link between empirical material and interpretation visible, which shows that the findings made were grounded in the material.

Overall, the choices made here were made to strengthen the trustworthiness, but they do not remove the limitations of the single-case design of the study. Eisenhardt (1989) describes case study research overall as a strategy for understanding the dynamics that are present within single settings. Studies of networks must take context and complexity into account (Halinen & Törnroos, 2005). This means in this study that the connected actor network around the pilot is treated as part of the phenomenon being studied. The strength and trustworthiness of this thesis lies in contextual depth, analytical transparency, and a visible link between data, coding, and interpretation.

3.5 Case: Company X

This study examines Company X, a startup operating in the Finnish public sector context. Company X is developing a digital service intended for professional use in public sector service processes and operates in a business-to-government setting where its activities take place in collaboration with a public buyer. Company X is referred to anonymously because the analysis addresses the company's current pilot, business configuration, and operating environment. Company X was selected as the empirical case because it provides access to an information-rich setting for examining how a digital service innovation is configured during the pilot phase and what kind of conditions the Finnish public sector provides for innovative startups. It also represents a relevant case because in this context innovation and development and possible continuation are closely intertwined with public sector practices and its organizational processes.

The digital service examined in this study is an AI-assisted solution designed to support the collection of preliminary information from social welfare clients for a professional service needs assessment. The solution is positioned in the early phase of the assessment process and is intended to support professional work. The service forms part of an ongoing effort to explore how digital tools can be used in public sector service processes. At the time of the study, the service was being piloted in collaboration with a Finnish wellbeing services county scheduled to continue a year. The pilot is treated as the empirical setting through which the broader research problem can be examined. The case is well suited for the purposes of this study because continuation past the pilot phase remains open at the time of writing, making it possible to study the current configuration without focusing on final outcomes. Company X therefore provides a relevant empirical setting for answering the research questions of this thesis.

4 Findings

This chapter presents the empirical findings of the study, describing how the digital service innovation pilot is currently configured across its business model design and surrounding ecosystem coordination, and how the institutional B2G context shapes this configuration. The views of the different actors involved are used to identify the configurational dynamics that define current state of the pilot project and picture what is feasible for the case company within the current public sector environment. The findings are organized using the theoretical framework introduced in Chapter 2 (Figure 2). The chapter is structured by the three analytical lenses of business model design, ecosystem coordination, and institutional context, and the analysis is guided by the corresponding analytical dimensions under each lens.

4.1 Business model design in the pilot arrangement

This section presents the empirical findings related to the business model design lens (see Figure 2). It describes how the pilot was talked about in the interviews in relation to the offering's purpose, the benefits, the way it is currently delivered, and the commercial logic attached to the pilot project and possible continuation. The section is organized into three subsections: the pilot as a mechanism for learning and benefit validation, the value logic pursued through the pilot, and cost and pricing logic in pilot conditions.

4.1.1 The pilot as a mechanism for learning and benefit validation

Pilots have been defined as formally designated, discrete events aimed at generating evidence regarding efficacy and implementation (Scarbrough et al., 2024). Across the interviews, the case company's pilot project was described as a setting where the practical use of the technology is explored and clarified. The case company founder (P1) in particular described the technology as something new and presented the pilot as a way to learn what can be achieved. The same also appeared on the buyer side, where piloting was discussed as relevant especially when something is being tried in a form that

has not previously been used in the same way. That point was brought up by the public sector procurement lead (P2) who framed piloting as the appropriate approach when the practical implications are not yet fully known beforehand. Additionally, the interview with the public sector service area lead (P3) placed the pilot in a broader organizational setting with referring to it as part of a wider learning process. The board member (P5) pointed out that validation from the pilot is a strategic requirement for building market legitimacy, which points towards a successful pilot serving as the necessary proof to position the startup as a credible actor.

Repeated references to validation also appeared in the data. For example in the P1 interview, the pilot project was described to be connected to benefit hypotheses that had been identified in advance and then converted into measurable indicators on the buyer side. The interviews also described piloting more generally as something that should produce evidence about whether a solution is worthwhile. P2 described piloting through criteria that are used to assess pilot solutions in practice, such as in terms of productivity, quality, usability, and reliability. Furthermore, P5 identified a successful pilot as the primary determinant for the case company's future growth, because measured and verified results could create the pull needed for other buyers to consider the solution.

In the case company's pilot, continuation was not described as automatic even if the pilot works in practice. P1 described the current agreement as ending at the pilot stage. More generally, the possibility that a pilot project may not automatically lead to continuation, especially in collaboration with the company doing the pilot, was also present in the interviews. This appeared in the P3 interview as a direct statement saying that a pilot project can end with the conclusion that the solution does not work. P4 noted that if a short-term change in operations is not made permanent, a pilot project can even be seen as a hindrance as it forces the organization to go back to old ways of working after a period of change. The P2 interview described the pilot as bounded under the procurement threshold and discussed the situation after the pilot through the starting

point of tendering. Overall, the case pilot is described to be within a wider institutional context where pilots are currently structured as temporary experiments.

4.1.2 The value logic pursued through the pilot

The business model lens structures the data reporting around value proposition, value creation and delivery, and value capture (Teece, 2010; Wirtz et al., 2016). This subsection describes how these value dimensions were discussed in the context of the case pilot.

The value logic of the case offer is built on an AI-assisted service that collects preliminary information from social welfare clients for a service needs assessment conducted by a professional. Based on the basic information provided in the case company documents, the value proposition that is desired would revolve around the improvement of the feeling of inclusion and engagement of the end customer when they are in the beginning stages of engaging with the service process. The material describes a shift in the service experience where the interaction is prioritized and ensures greater feelings of participation and security of the end customer. This automated data collection is presented to enable the professional's focus to move from the current administrative tasks more toward professional assessment. For the buyer organization, they promise improved process quality and speed by identifying requirements earlier.

In the case company documentation, value creation and delivery are structured through a secure and accessible system designed to be flexible and suit technical environments of the clients'. The service is described as a system that uses automated dialogues to standardize client information collection. For value capture the company's documentation brings out two main advantages, which are the scalability of the operating model and the standardization of documentation practices. The material states that value capture is achieved using a standardized approach where consistency in quality is guaranteed no matter of the performance of an individual. The documentation also mentions the possibility of expanding this same operating model to

multiple service points with moderate costs once the initial infrastructure is established more.

The interviews provide further detail on how these intended value dimensions are perceived and how they have become apparent in the pilot environment. The pilot was talked about through a value logic that repeatedly returned to work time use and the service process itself. P1 described how the service produces information and how that information would move into professionals' daily processes. Key point they emphasized was that for the value logic to be fully realized in a scaled version the solution must eventually flow seamlessly into other systems and not require manual data entry.

The same topic appeared on the buyer side, where P3 discussed locating where the benefits would show up in practice. Value was described in language that connected the pilot to staffing capacity constraints and resource pressure. The pilot was described as an initiative for quality improvement and in relation to the idea that the benefits gained from it should appear as a more efficient use of professional labour. P3 stated that in a staff-oriented organization, the benefit must ultimately materialize as a reduced need for personnel to perform the same tasks:

> "*Elikkä suoraan sanottuna me tarvitaan vähemmän henkilöstöä tekemään niitä samoja asioita mitä tällä hetkellä*" (P3, personal communication, 23.2.2026).

P5 described this requirement as a critical step for business and said that potential benefits such as time savings need to be converted into monetary values and staffing equivalents so that the value proposition is clear for decision makers. P5 also noted that what counts as value differs across functions because some more frontline actors tend to emphasize workload reduction while financial decision makers focus on budget impact and in monetary savings.

Value was discussed through the client perspective too, for example including attention to how clients experience the piloted solution. P1 identified that the value proposition

currently depends partly on practical conditions that are outside the service itself. They explained that in the target group basic requirements for clients using a digital solution are not always in place and that this has only become visible when the piloting has started. P1 also described that their value proposition is currently affected by the technology itself not responding with the same level of reliability in every interaction and that variation appearing as the system sometimes failing to catch what the user is trying to communicate. P5 expanded on this with that value is at the end experienced through the client. With this view in mind, the value logic includes reducing the burden of repetition where the end client no longer needs to explain the same situation multiple times because the data is recorded in writing.

P2 described the concept of value in piloting more generally through dimensions of practical assessment. Piloting was discussed in terms of checking whether there are productivity gains, if service quality or efficiency is improved, and does the system actually work in everyday use from the angle of usability and reliability. P2 said that in this specific context, the value of the AI cannot be autonomous decision-making, and that the AI's role is preliminary interpretation that a human must verify. Alongside these, this procurement interview also focused on information security and data protection being a tightly connected part of what gets assessed in technology pilots once client data is involved, e.g., where data is stored and who can process it.

4.1.3 Cost structure and pricing logic under pilot conditions

In VoD, value capture is central because a venture can have a working solution but still lack a defined mechanism to generate sustained revenue and cover ongoing costs (McIntyre, 2014; Teece, 2010). In the founder interview, the pilot’s cost base was described mainly through project work. The costs of the pilot were broken down by P1 as coming from specification work, technical development, and the technologies used. A variable cost element was also described. They mentioned technology and AI-related costs through usage time and stated that the current price contains assumptions about

how long each use of the client end solution will take. Longer usage times were described as changing the cost side for the provider and this was discussed as an uncertainty within the pilot by P1.

The pilot was described as fixed price for the buyer side in the same interview, but pricing beyond the pilot for a more scaled solution was discussed as something that could take different forms. P1 stated that in this operational context the buyer ultimately decides the format of pricing and gave examples such as a monthly price for unlimited use. The interview also included a description of a preferred structure from the provider side as a combination of a fixed element and a usage-based element. P5 supported this more hybrid approach and proposed a model where a base fee covers the technical environment and support, while a variable transaction fee scales with the volume of cases. P5 also emphasized the importance of maintaining a high initial price point to preserve credibility, as underpricing can be interpreted as a sign of low quality.

Moreover, the pilot budget was described as a constraint on what can be done during the pilot. The case company founder described the pilot being procured under the 60,000€ threshold and said that this limits the amount of work that can be placed into the project. The threshold was discussed across all interviews as defining what can be done within the pilot without moving into tendering.

4.2 Ecosystem coordination in the pilot configuration

This section presents the empirical findings related to the ecosystem coordination lens (see Figure 2). It describes how the pilot is coordinated across the actors involved, how decision-making and responsibilities are described in the interviews, and what kinds of interdependencies and bottlenecks are attached to delivery and embedding in practice. The section is organized into two subsections: the actor structure and coordination around the pilot and dependencies, bottlenecks, and challenges in the ecosystem.

4.2.1 The actor structure and coordination around the pilot

Ecosystems can be defined as the alignment structure of the multilateral set of partners required for a focal value proposition to materialize (Adner, 2017). This section describes the actor set and role division identified in the interview material. The pilot is organized around a small group of central actors where technical and administrative responsibilities are divided between the company and an external partner. On the company side, P1 described the pilot-stage setup as being run with a very small internal team, where the same people end up carrying multiple responsibilities and roles:

> *"Nythän me ollaan täällä [case-yrityksessä] niin tosi pienellä porukalla 'me myself and I' henkisesti tehty näitä kaikkia rooleja" (P1, personal communication, 11.2.2026).*

P1 further described Company X as being responsible for and carrying roles related to domain knowledge, customer ownership, and project management, while the technical development, delivery, and maintenance are provided by a separate technology partner.

On the buyer side, the actor structure consists of a network of leadership and administrative duties. The public sector service area lead P3 provides general leadership for the large service area where the pilot takes place. This role is described as a general manager who looks for practical improvements and oversees the environment, with decision-making strongly guided by financial considerations because of the environment being described as constrained in resources. In the P3 interview, top management is identified as a distinct actor whose support provides the necessary weight and priority for the pilot and new initiatives overall to move forward within the organization.

Procurement acts as a strategic coordinator within the buyer's structure, focusing on oversight, guidance, and legal alignment. The public sector procurement lead P2 manages strategic coordination and works with service areas when substantive needs and rules are weighed against each other. The preparation of actual documentation and the overall procurement process is done by specialists and a procurement manager that

handle all of the pilot's operational procurement work. This role is described as distinct from hands-on implementation, and an actor that e.g., provides information on the legality and feasibility of procurement proposals.

The interviews also positioned professional end-users as a relevant actor group in the pilot configuration. In the P1 interview, professionals were described as being involved early in work sessions where the pilot and its use were worked on together. In the P3 interview, professionals were similarly discussed as the users of the piloted solution in the service process.

The environment also includes system providers who own and operate the core client information systems used by wellbeing services counties. These actors were described as part of the wider configuration around the pilot because they own and operate the systems that hold and manage client information. In addition, P4 referred to the presence of multiple parallel systems alongside the main client information systems.

Overall, the pilot stage ecosystem actor coordination was framed as a combination of boundaries defined in the contract and regular decisions made along the way. The case company founder described that the contract sets the baseline of what is being done, but practical decisions keep happening during delivery especially when content and functionality are adjusted. There the wellbeing services county is the actor that defines what is wanted on the content and functional side, and the case company evaluates whether additions are implemented and then communicates the final implementation choices to the technology partner for execution.

4.2.2 Dependencies, bottlenecks, and challenges in the ecosystem

Ecosystem coordination involves managing interdependencies that can constrain value delivery if required contributions do not align (Adner, 2017). A bottleneck kept coming up in the data that concerned how the information produced through the pilot becomes

part of routine professional work and existing systems. Information transfer in the pilot may involve manual workarounds when content is moved into professional documentation, according to the case company founder. They described that the intended goal is for the output to move smoothly into existing workflows and systems instead. Outside of the case pilot a similar description appeared in the P3 interview, where the environment was discussed through the need to enter information again because it does not transfer automatically into the core system which becomes an additional step to the overall process. P5 suggested a way to manage this dependency in ideal conditions the venture should coordinate the ecosystem by providing its own standardized interface for others to connect to. Not customize the solution for multiple various systems.

Additionally, the P1 interview described that a single pilot delivery cannot simply be replicated from one environment to another. P1 stated that the solution can accept unique content. Where the main differences appear is when connecting to the customer environment. The founder stated that each wellbeing services county has its own systems and that the way information is received and recorded in one environment may not be the preferred way in the next environment.

Embedding was described through the practical difficulty of changing routines inside service units. From the company perspective, P1 described an environment with a many systems that determine whether new tools become seamless or just add layers, this way connecting to the embedding challenge. On the buyer side, coordination was described as extending into operative adoption and everyday embedding work within the service units. P3 described that new ways of working do not spread simply because leadership wants them to, and that spreading and embedding requires peer examples:

> *“Nämä ei lähde niin, että johto käskee, että nyt otetaan tämän käyttöön vaan se vaatii toisilta oppimista ja sitä kautta leviämistä”* (P3, personal communication, 23.2.2026).

P3 later emphasized that the support and time needed by personnel is often underestimated. They described that adoption depends on engagement inside the units, and mentioned operative level managers as the actors who provide the local support and peer examples that enable a new way of working to be embedded and spread to other environments. P5 described that progress is supported in organizations when all functions are engaged in the same discussion so the initiative does not get stuck in any internal gatekeeping or issues over e.g., budgets and decision rights.

Feasibility for broader delivery came up on the company side as reliance on contributions outside the company's direct control. P1 described the technology partner as critical for delivery and described that replacing that partner would create uncertainty. In the same interview it was also described that if several projects surface the team would need to grow, and further development and keeping the technology up to date was described as ongoing work. P5 similarly framed scaling feasibility as requiring organizational capability building beyond the pilot team, including dedicated functions all operations as volume and market scope grow. P5 made an additional point that scaling wider even within the domestic market means a bigger organization than what they had at the pilot stage, and international would need additional resources for getting into markets and building relationships.

4.3 Institutional context in the pilot configuration

This section presents the empirical findings related to the institutional context lens (see Figure 2). It describes how the B2G environment, its procurement rules, regulatory layers, and public sector logics were discussed across the interviews. The section is organized around the three analytical dimensions derived in the theoretical framework: regulatory interaction constraints, legitimacy tensions, and resource asymmetry. Each subsection reports how these institutional pressures appeared in the interview material and in the conditions surrounding the pilot.

4.3.1 Interaction constraints in the pilot

Public procurement regulation can restrict interaction between buyers and suppliers and thus constrain things like collaborative learning during innovation procurement (Melander & Pazirandeh Arvidsson, 2020; Arlbjørn & Freytag, 2012). The interviews show that the pilot was influenced by three connected boundary conditions: the procurement threshold, the rules governing post-pilot continuation, and the fragmented regulatory environment across counties. The 60 000 € below-threshold limit emerged as the primary structural condition shaping the arrangement. P1 identified it as the defining feature and noted that it both enabled the pilot and constrained its scope. P2 confirmed this and described the threshold as a financial ceiling and the boundary that determines which procurement procedure applies and how much development work can be placed inside the arrangement.

Another complexity has to do with how the procurement category of a digital solution is determined. P2 described an ambiguity about whether a digital tool should be classified as an ICT service procurement which falls under the 60 000 € threshold or as a social services procurement which has a much higher procurement threshold. P2 was clear that digital tools operating within the social welfare process cannot straightforwardly be reclassified as social services procurements, and that the classification has major practical consequences for how a pilot or any continuation can be structured.

Once the pilot ends, the institutional default is competitive tendering. P1 stated that the current pilot agreement does not include a clear continuation path or route. P2 likewise described tendering after the pilot as the standard expectation. They stated even when a pilot is viewed positively by the operational side continuation cannot be treated as a direct extension of the pilot arrangement. P3 similarly acknowledged that a pilot can end with the finding that the solution does not work and that even a respected pilot does not automatically translate into a continued relationship. The interviews also included descriptions of what continuation moving to tendering can mean from the case company's perspective. Both P3 and P4 described the possibility that other actors could

offer their solutions and win the tender after the pilot and used this as an example of why continuation is uncertain from the pilot company's perspective.

Another boundary comes from the prohibition on contract splitting. According to P2, the legal rule against contract splitting works as an active compliance boundary. They also gave an example of a direct procurement request related to an existing pilot that was declined because adding it would have pushed the overall cumulative value of purchases related to the same project over the threshold. P3 said that supplementary purchases connected to ongoing pilots can easily enter grey areas where the same concern arises. This points out a broader challenge in how the legal rule interacts with the practical needs to support and adjust projects during their run. The direct procurement process involves internal oversight. Procurement reviews and endorses the legality and appropriateness of the decision before it is approved. P2 noted that if a service area still decides to proceed without that endorsement, the responsibility for the decision rests with the service area and not with procurement.

The overall procurement process itself was described as placing strict limits on the interaction that can take place between buyer and supplier during and after tendering. P2 observed that fast and agile development together and continuous collaborative adjustment are structurally difficult for a public procurement unit. It is because once a service is competitively tendered the contract content and scope are tightly fixed, which ends up leaving very limited room for modification. P4 described a related dynamic and noted the procurement unit and the operational service area functioning as separate, with the first focusing more on legal compliance and the latter on the underlying need and implementation. This kind of separation limits the depth of interaction that influences what gets procured and how. P2 and P4 both connected this separation to the difficulty of ongoing adaptation between the solution and the organization. P1 described longer-term co-development as something that would be more sustainable from the supplier perspective than the current pilot arrangements but additionally mentioned that it does not fit within a threshold pilot budget like it is currently being used.

These interaction constraints are further complicated by the fragmented regulatory environment across wellbeing services counties. P1 described each county as operating its own procurement guidelines, its own interpretation of applicable rules, its own procedures for data protection, data security, and AI-related assessments. All of these additionally further influenced by their own customer information systems, noting that these layers have accumulated over decades. In this context, fragmentation means that each new county requires repeating the same processes from the start. This is limiting how much a provider can build on prior gained experience.

Innovative procurement procedures exist in law as a potential alternative but were described as rarely used in practice. P2 mentioned the innovation partnership as being available under the procurement act but untested in real context. P4 described the same gap. While innovative procurement procedures could in principle be used, doing so would require the procurement unit to initiate and lead the process from the beginning, with procurement preparation alone taking months due to administrative processes. P4 framed this as a more meaningful organizational leap and not a simple and straightforward alternative. P4 also noted an alternative where the case company could sell development work rather than a product like in the current case but described that as a different type of commercial arrangement altogether.

4.3.2 Legitimacy tensions in continuation decisions

Legitimacy concerns whether actions are perceived as desirable, proper, or appropriate within a socially constructed system of norms and beliefs (Suchman, 1995), and evaluations of this legitimacy can differ under institutional complexity (Greenwood et al., 2011; Bitektine & Song, 2023). In the data cost savings and efficiency were repeatedly described as the dominant criteria which new solutions are assessed through in wellbeing services counties. P3 indicated that the benefit from any new solution must ultimately materialize as a reduced need for personnel to perform the same tasks and

that this calculation is what pushes decision-making at their organizational level. P5 also mentioned this and discussed financial impact as the indicator that carries the most weight with those making decisions. They mentioned that this cost-cutting orientation may also discourage engagement with suppliers that are perceived as riskier even when those suppliers may also promise most efficiency.

Legitimacy requirements specific to AI-assisted solutions appeared in the interview data. Data protection and data security were said by P2 to be an inseparable part of how technology pilots are looked at and assessed once client data is involved. This includes questions about where data is stored, under what conditions it can be processed, and whether it crosses any administrative boundaries. P2 also identified AI's role as restricted to preliminary interpretation that a human must verify as a regulatory in the social welfare context. P1 confirmed that these requirements had been present from the beginning of the pilot setup. The buyer had required modifications to technical definitions before the contract was signed and AI transparency obligations had been worked through during the pilot process.

The requirement for compliance validation was described as not transferable between wellbeing services counties. P4 described the data protection impact assessment process as something every county must conduct independently, meaning a provider who has completed it with one buyer cannot rely on that outcome when approaching another. P4 stated that the idea of national clearance had begun to surface in discussions but that a bigger structural change of this kind was distant. P1 framed the same issue with practical progression and stating that completing the auditing process in one environment gives no guarantee of how it will proceed elsewhere:

> "*Jos oot jostain päässyt niinku auditoinnin läpi niin se ei todellakaan takaa että se menee samalla tavalla jossain muualla*" (P1, personal communication, 11.2.2026).

The risk logic of public sector actors came up in a way that reflected institutional terms. P4 described that in the role of an official in this context inaction is often seen as safer than action. This drives the cautious stance toward new suppliers and new technology. They mentioned that when the regulatory environment around AI is unclear and still developing, this cautious tendency is amplified as actors interpret the available boundaries as tightly as possible to limit exposure to risk.

This risk logic was reflected in how startups and large incumbents are perceived in continuation decisions, too. P2 observed that choosing a large established supplier creates a different accountability situation than choosing a startup. Negative outcomes connected to small provider choices carry visible reputational and legal consequences in a way that comparable outcomes connected to large incumbent choices do not:

> "*Meillä on paljon helpompi käydä jonkun ison toimijan kanssa se keskustelu. Iso toimija voi mokata ihan niin paljon kun haluaa. Ei kukaan siinä enää reagoi*" (P2, personal communication, 20.2.2026).

A related pattern from the venture side was described by P5, who said that a successful pilot at a credible public sector buyer is their primary mechanism for establishing the credibility required to possibly enter other market contexts. P5 also noted that demonstrating the solution works in the highly constrained public sector environment could reduce perceived risk for other potential adopters of the company's offering:

> "*Kaikki tietää kuinka vaikeata se on toimia näiden julkkareiden kanssa, ja sitten jos sä sen pystyt todentamaan, niin se toimii hyvänä ponnahduslautana ja referenssinä*" (P5, personal communication, 15.3.2026).

The interviews also suggested that the institutional environment is not experienced only restrictive, because experienced buyers may still find room to shape how a later tender is specified. P3 mentioned that in practice, tender specifications can be formulated in a way that they align procurement more closely with a preferred solution’s direction. They added that this can be true especially when the buyer has gotten a clearer sense of the

need and the market through the pilot. They positioned this as a realistic mechanism available to experienced buyers and said that a provider who has demonstrated clear benefits during a pilot is in a stronger position when tendering is later prepared around that demonstrated performance. P5 noted that the data gathered during the pilot is also described as essential for predicting future cost structures and justifying the business case in the larger following tendering processes.

The interviews described some competing agendas in public organizations that can influence whether these new solutions advance forward. P5 described situations where individual actors act as gatekeepers, and that having someone willing to pursue the initiative internally becomes essential and necessary for progress. P3 similarly pointed out that top management support carries real weight in determining whether these new initiatives receive the priority and resources needed to move forward.

4.3.3 Resource asymmetry in broader contracting

Innovation procurement can disadvantage SMEs because meeting participation requirements can be administratively heavy and resource intensive (Saastamoinen et al., 2018; Akenroye et al., 2024). The participation conditions attached to competitive tendering were described across the interviews as presenting barriers for smaller suppliers. P2 described conditions such as turnover requirements, credit rating thresholds, and reference requirements as features that vary between wellbeing services counties and that can effectively completely exclude startups from participation in the tender. They illustrated how a reference requirement specifying prior deliveries to public sector clients would in practice exclude any firm that had not already competed in comparable procurements. Startups may not have developed the capabilities yet, meaning ones that go beyond technical expertise. They also acknowledged that their county had worked to minimize unnecessary participation barriers but noted that the administrative and compliance still demands require those broader capabilities.

P4 described the same pattern at a broader level, and that procurement conditions are typically built around solutions with demonstrated histories. They also said that this effectively completely filters out innovative entrants regardless of their technical capability. P4 also described startups' short-term revenue pressures in this context by saying that pilots may be accepted even when continuation is uncertain because the immediate need for revenue outweighs the structural uncertainty around what follows.

The contract size structure of public procurement was described as a distinct capacity barrier. P4 described the scale of procurement contracts in the wellbeing services sector as running into the tens of millions, noting that organizations too small to compete for or sustain contracts at that scale are disadvantaged from the point of entry regardless of the quality of their offering. P5 noted that the below-threshold pilot arrangement, while accessible, produces contract values insufficient to sustain the investment required for broader development. The layered compliance burden was described alongside these disadvantages. P2 described the multiple regulatory frameworks that must be satisfied simultaneously in any procurement. This includes sector-specific legislation, data protection regulation, and AI-related obligations. P1 described the procurement law and associated requirements as having expanded dramatically from simpler earlier versions.

Legislative unpredictability was described as creating another burden to add to the complexity of current requirements. P2 described the regulatory landscape as having tightened continuously over a twenty-year period with new obligations coming up from both EU directives and national implementation on an ongoing basis. P2 noted that a multiple year contract period creates the risk that the technology becomes not compliant or completely obsolete before the contract ends, which is making investment decisions harder to justify. P2 continued that while EU-level discussions around regulatory simplification could eventually ease conditions Finland's own implementation patterns have historically added more layers.

For suppliers, the fragmentation of regulatory environments also acts as a direct multiplier of costs. In the pilot arrangement what looks like repeated procedures and processes becomes repeated compliance demands across counties. P5 suggested that the state could instruct counties to harmonize their operating models but has not done so and framed this as a gap that leaves suppliers facing different standards and system environments in each county despite a shared national legislative framework.

An in-house company structure was described as a potential workaround to some of these capacity barriers but its viability was presented as uncertain. P1 mentioned it as a route through which continuation without competitive tendering might be possible but emphasized uncertainty about its current legal standing. A more detailed explanation came from P4, who specified more that an in-house actor could procure the solution and wellbeing services counties could then purchase through that actor without running their own procurement process. P4 also however explained that a procurement law reform introduces an ownership limit for these in-house companies which will likely be resulting in disruption. They pictured the workaround as strategically possible but noted that the legal situation was actively changing and that the likelihood of successfully using this mechanism had decreased. The interviews also pointed towards B2G ventures needing to explore alternative scaling paths outside of the public sector because of these structural constraints. P5 described that because the continuation path after piloting is so uncertain the venture considers building parallel scaling paths in other places to reduce dependency on any single route to market.

4.4 Summary of empirical findings

Across the three analytical lenses, the findings describe how the pilot is currently configured and what conditions surround it. At the business model level the pilot functions as validation where for future decisions evidence is collected and premade benefit hypotheses are tested. The pilot is also the proof point that builds external credibility and supports later scaling. The value proposition is focused at its core on more

efficient professional work and improved client experience. The interviews held however revealed that the proposition currently depends on conditions that are not fully within the venture's control. This most notably includes e.g., how consistently the technology performs and integration with existing systems. The cost structure is shaped by the direct procurement threshold of 60 000 €. The pilot operates under a fixed-price arrangement with a variable cost element tied to usage, and the revenue logic for a scaled operation has not yet been fully specified. Reliance on a single continuation route is fragile, increasing the importance of parallel scaling options beyond one buyer path.

At the ecosystem level, the pilot is organized around a small actor set in which responsibilities are divided between the case company, a technology partner, and the buyer organization. The buyer side involves multiple roles e.g., service area leadership, procurement specialists, and professional end-users. They each have their distinct roles in how the pilot is managed and assessed. The findings describe recurring dependencies around system integration, manual workarounds in information transfer, and the challenge of embedding new routines into established workflows. Replication across counties is described as requiring significant repetition of setup work due to differences in systems and operating environments between counties. Scaling is also an ecosystem coordination problem that depends on arrangements that remain stable across multiple units and partners instead of being rebuilt as a one-off specific pilot setup each time. The interviews also imply that scaling and market access may depend on higher-level coordinating actors and networks not just individual units.

At the institutional level, the procurement threshold defines the boundaries of the pilot arrangement and determines what development work can be included within it. The post-pilot default is competitive tendering, and continuation with the same provider is not structurally guaranteed. Legitimacy requirements that are specific to AI solutions in the social welfare context are present from the beginning. These include e.g., data protection, AI transparency, and human verification obligations, and compliance validation for most of these must be repeated independently in each county. The risk

logic of public sector actors, the credibility gap between startups and large incumbents, and the participation conditions most times associated with competitive tendering together describe an environment that places significant demands on smaller suppliers that are seeking to move beyond their initial pilot project.

5 Discussion

The empirical analysis shows that Company X's pilot is functioning under the conditions that currently support it. The solution is delivering value in the pilot setting, but the central question of this thesis is not whether the pilot works. It is whether what works in the pilot can survive the transition to routine delivery. That question points to a more fundamental issue. A pilot is a temporary configuration, often enabled by conditions that do not persist after the pilot ends. Scaling is therefore not a continuation of what the pilot set up but requires reconfiguration of that setup under different conditions. This chapter synthesizes the empirical findings from that perspective.

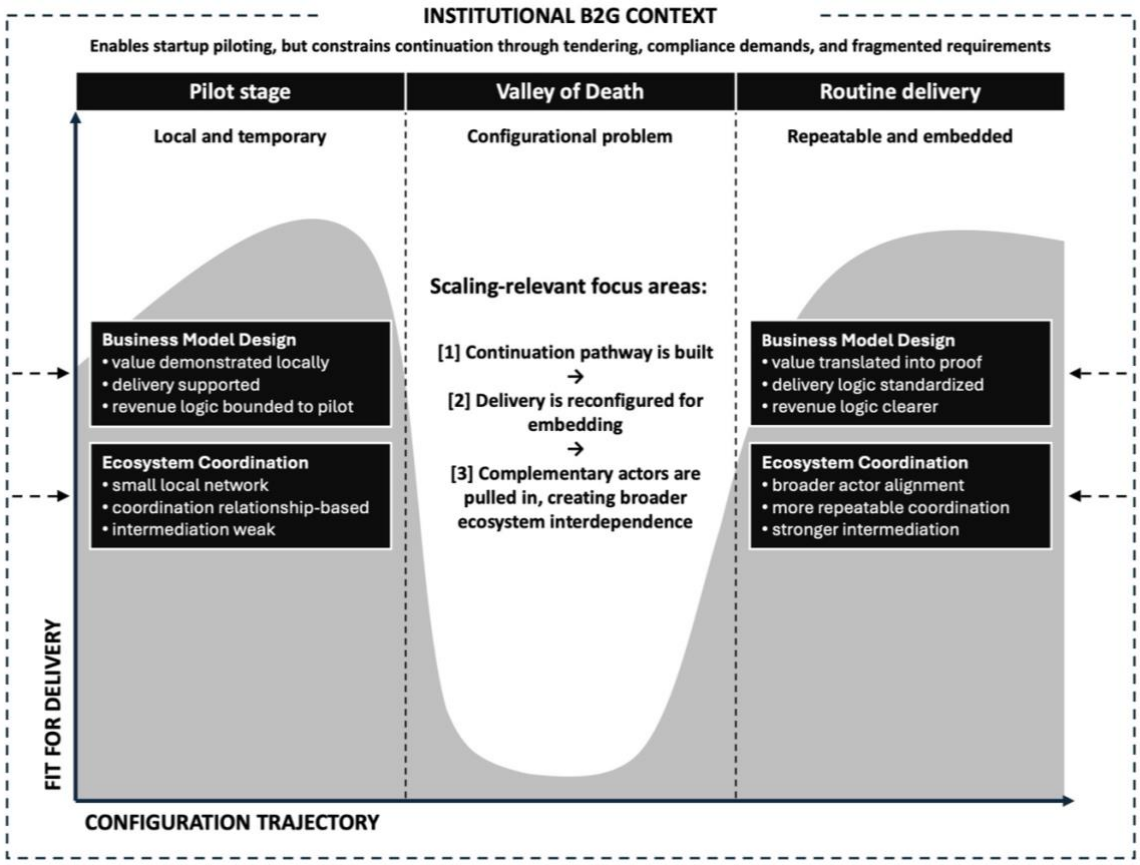

Figure 4. Empirically Enriched Framework: VoD as a Configurational Problem

Figure 4 presents the empirically enriched framework developed with the case. It shows the Valley of Death in this B2G setting as a configurational problem. The pilot project

represents one locally working and temporary configuration, while routine delivery requires a different configuration that is more repeatable and embedded. The vertical axis, fit for delivery, refers to how well the current configuration matches the delivery conditions of that stage. In the pilot stage, fit is relatively high because the arrangement is supported by pilot-specific conditions like temporary arrangements, local actor-level coordination, and otherwise supported delivery. In routine delivery, fit must be achieved again under different conditions, where continuation, embedded use, and broader coordination must hold without the same temporary support. The horizontal axis, configuration trajectory, refers to the movement from pilot stage configuration toward routine delivery configuration. It describes the path of reconfiguration where the venture attempts to move from one type of arrangement to another.

Between these two states is a gap where the pilot stage configuration no longer carries over automatically but the routine delivery configuration has not yet been achieved, this framed here as Valley of Death. The curve illustrates a loss of fit for delivery during the transition stage. This gap highlights three scaling-relevant focus areas that emerge from applying the theoretical framework to the case. Together, they point to the key configuration shifts that require attention to move beyond the pilot. The following section examines these three focus areas in greater depth.

5.1 Scaling-relevant focus areas

The three scaling-relevant focus areas are derived from the interaction between business model design, ecosystem coordination, and institutional context. Together they explain what should change for the configuration to move from validation in the pilot stage to actual routine delivery. Both focus areas [1] and [2] concern business model reconfiguration under institutional B2G conditions. Focus area [3] concerns the broader ecosystem coordination that may begin to form once that fit has become more stable and durable. It can be argued that this sequence matters because without a business model that is compatible with continuation and more embedded, the conditions that

would attract complementary actors and enable broader ecosystem formation do not yet materialize.

5.1.1 Building a continuation pathway

The first focus area is the pathway from pilot project to continuation. The main claim is that scaling in a B2G context depends on two connected conditions. First, whether pilot outputs can be converted into proof for procurement and second, if there is a viable continuation route that maintains the momentum to avoid resetting the process each time the pilot ends.

This is where problems like the VoD and pilotitis show as structural. Scarbrough et al. (2024) argue that pilot evidence must travel from the local setting in which it was generated to other contexts and decision points to support scaling. The problem is that the evidence a pilot produces is often in forms that give an idea of local usefulness but do not yet match the logic of continuation in public procurement. In the institutional environment of Finnish wellbeing services counties the main pathway after a pilot project is competitive tendering. Also, continuation decisions are shaped by cost, compliance, and broader supportability instead of pilot performance alone. This requires the venture to convert what it has learned and demonstrated in the pilot into forms of proof that can support continuation.

The case of Company X displays this dynamic clearly. The pilot is seen more as a mechanism for validation and credibility-building and treated as an opportunity to gather the kind of proof needed to move forward. Yet the after the pilot the default is to return to open competitive tendering. This means that continuation with the same buyer is not guaranteed and continuation with new buyers requires the process to be restarted in each new county. Two features of the case enhance the challenge. The evidence from the project should be shown of service quality and client experience and especially in forms that translate directly into argumentation for procurement. These could be staffing effects, reductions in workload, and other economically measurable outcomes.

The second is that even where proof like that exists the route forward remains fragile because compliance and validation requirements are not transferable across counties. What works in one setting may still need to be revalidated in the next.

This means that the continuation pathway is a question of whether the process can move forward without losing the movement created in the pilot. In the case company's situation, the risk is that the pilot produces a local proof point but the institutional route after it still returns the venture to a new procurement cycle, repeated work and another moment of uncertainty for continuation. In that situation, the pilot may validate the solution without reducing the structural distance to broader delivery.

The focus area therefore concerns business model design reworked under institutional B2G conditions. At the levels of value proposition, this means expressing value in forms that matter for continuation decisions. For value capture it means moving beyond a pilot-specific commercial arrangement toward a clearer continuation logic after the pilot. Without both, the venture remains in a configuration where local piloting is possible, but broader continuation is repeatedly reset. The continuation pathway is therefore the first condition for crossing to begin because without it the business model may work in the pilot but still fail to move past it.

5.1.2 Reconfiguring delivery for embedding

The second scaling-relevant focus area is the reconfiguration of delivery for embedding. The central claim is that continuation is a delivery architecture problem, not a procurement problem only. More specifically, the constraint in the case is not the diffusion of the innovation, meaning the spread of awareness and adoption across potential buyers. It is more the embedding, which means the process of integrating the service into existing systems and routines so that it becomes a normal part of operating practice rather than a parallel project that requires separate management.

This argument is well grounded in the literature on embedding and routinization. Scarbrough and Kyratsis (2022) show that scaling depends on whether the innovation becomes embedded in organizational routines and practices. The transition from pilot to embedded routine service involves what Damschroder et al. (2009) describe as the institutionalization of the innovation within its intended context. Pilot project conditions include extra attention, temporary workarounds, project-level support, and manual processes that help the service function in the short term but do not represent the conditions of routine delivery, which is why it matters. When all of those supports are removed, the service may no longer fit the actual conditions of everyday use.

This is clearly visible in Company X's case because the value proposition seems to depend heavily on integration into routine workflows, information systems and the broader overall service process. The findings show manual information transfer steps, system frictions, and repeated county-specific setup work. These are not minor operational details and actually show that the current delivery logic is still partly supported by pilot conditions and has not yet reached the embedded form required for routine delivery. In Baden-Fuller and Haefliger's (2013) terms, the configuration now is closer to a project-based "taxi" logic than to a more repeatable "bus" logic.

This challenge is also part of the business model. In the case for Company X, value is sustained by whether the service can be delivered as part of ordinary work. If information must be transferred manually, integration continues to be incomplete, or if the service adds extra steps to a process, the value demonstrated in the pilot becomes harder to maintain beyond it. Therefore, the main issue is if the service can be delivered in a form that fits routine use. In Company X's case, this means making information move more directly into existing systems and avoiding repeated setup work when the service is introduced in a new county. Without such changes, the service may continue to function in individual pilots, but delivery remains too site-specific and resource-intensive for broader continuation.

This is why the second focus area concerns business model design at the level of value creation and delivery. The first two focus areas describe the business model reconfiguration for crossing to begin. The first concerns whether the offering can move forward under procurement conditions and the second concerns if it can be delivered in a more repeatable and embedded form.

5.1.3 Complementary actors and the ecosystem

The third scaling-relevant focus area concerns the broader ecosystem that may begin to form around a configuration that is more embedded and compatible with continuation. The claim here is that scaling in the case is not mainly a question of finding more buyers. It is a question of whether the solution can begin to connect to a broader ecosystem past similar one-off pilot projects. To be more specific, once the business model has been reworked and has a clearer continuation pathway and a more embedded delivery logic, the offered solution becomes more reasonable for complementary actors outside the current case. As that happens, a broader ecosystem may begin to form which could make market access and repeated delivery more feasible.

This is where the ecosystem lens can be connected and becomes important. Adner (2017) emphasizes that an ecosystem is defined by the alignment structure required for a value proposition to materialize. Clarysse et al. (2014) likewise show that early-stage innovation arrangements do not automatically become viable business ecosystems. Therefore in the case, the relevant ecosystem issue is not simply the current pilot actor set. It is the actual absence of a broader external layer of intermediation and coordination that could help the solution travel beyond one case.

The case suggests that such a broader layer is currently weak or missing. The existing setup is local, relationship-based, and held together through direct case-specific coordination. At the same time the interviews imply that broader progression may depend on actors and mechanisms outside the current case itself, something such as

intermediating or structures that span boundaries that help the solution gain access to the fragmented B2G field. In this sense, the ecosystem is not the current network simply made bigger. It is the broader layer through which a more stable configuration might begin to connect to repeated market access.

This is supported by the ecosystem literature used in the thesis. Stehn and Erikshammar's (2025) digital broker, Garcia Martin et al.'s (2024) intermediary roles, and Clarysse et al.'s (2014) boundary spanners all point to the importance of actors who reduce friction between projects, organizations, and contexts. In the case, the absence of intermediation like that helps explain why scaling remains so dependent on local relationship work. These broader structures are unlikely to form around an arrangement that remains purely pilot project stage in its character. They become more plausible when the solution is begins to look more compatible with continuation and more supportable beyond one local supported setting.

The legitimacy implications reinforce this point. Bitektine and Song (2023) show that organizational legitimacy is evaluated differently by different evaluators applying different institutional logics. In Company X's case, the solution must be credible to actors who prioritize operational usefulness, procurement compliance, and cost control. These evaluations do not necessarily reinforce each other and this makes single-actor relationship management insufficient as a scaling strategy. Thus, a broader set of structures that can help the solution travel across these different evaluation stages becomes important.

The implication for scaling feasibility is the possible emergence of a broader ecosystem. In practical terms, this could mean the intermediating and cross-boundary structures that can lower entry friction with reducing repeated work across counties and that connect the solution to a more repeatable route to market. Their absence shows a big missing layer that is currently being compensated for with project-level relationship

work. That may be enough for a single local case, but it is unlikely to support broader progression when volume and geographic scope increase.

5.1.4 Synthesis of the scaling-relevant focus areas

The three focus areas describe the Valley of Death in this B2G context as a configuration break between two different states. The pilot represents one working configuration: local validation, temporary support, and narrow relationship-based coordination. Routine delivery represents another: business model design that is more compatible with continuation, delivery that is more embedded and repeatable, and a broader ecosystem meso-layer larger than a one local and supported case. The valley between them appears because the first configuration does not automatically carry over, while at the same time the second has not yet been assembled.

The transition over the gap operates in a pattern that begins when business model design is reworked under institutional B2G conditions. This is building a continuation pathway that translates pilot outputs into procurement-relevant proof and connects them to a route forward. The second step consists of restructuring delivery in such a way that the product is less reliant on project-specific support and more capable of embedded and repeatable use. With both of these beginning to hold, they create a more coherent and supportable basis of progression past the protected pilot setting. Around that, complementary actors may begin to see a role for themselves, and a broader ecosystem layer may start to form.

This means that the ecosystem layer is not separate from the dialogue of business model and institutional context, but actually grows around it. Therefore, to move past the pilot is not a purely internal business model problem or an external ecosystem problem. It is a configurational transition where business model design must first become more compatible with continuation and more embedded in delivery, and only then a broader ecosystem can begin to form around that configuration.

5.2 Answers to research questions & study contributions

This section answers the two research questions and clarifies the main contributions of the study. **RQ1** is addressed first, then **RQ2**. The section then explains what the study contributes conceptually and empirically when the two answers are read together.

RQ1 asked what the key dynamics are between business model design, ecosystem coordination, and institutional context in the scaling of digital service innovations. Chapter 2 addressed this question at a conceptual level by developing the three-lens framework. The literature review showed that the three lenses interact in ways that shape what is feasible in scaling. Chapter 5 refines that answer through the case. The empirical synthesis shows that the key dynamic begins with the relationship between business model design and institutional context. In the B2G context, business model design should not be designed without procurement rules, evaluation criteria, and participation barriers in mind because these shape which configurations are compatible with continuation and what are limited to the pilot stage. The first two scaling-relevant focus areas capture this directly because moving beyond the pilot requires the business model to be configured for both continuation and for routine delivery. There should be a continuation pathway that makes progression institutionally possible and delivery should be configured so that the service able to function under routine conditions. The ecosystem lens then adds another dynamic by showing that a broader ecosystem can begin to form only when this more compatible configuration creates visible opportunities for complementary actors.

RQ2 asked how business model design and ecosystem coordination are configured in a digital service innovation pilot within a B2G institutional context. The answer is that in the pilot business model design is configured as a locally demonstrated and temporarily supported arrangement. The value proposition is proven in local setting, its delivery still relies partly on pilot conditions, and its revenue logic remains tied to the pilot. Ecosystem coordination is also a narrow and local arrangement with relationship-based coordination and weak intermediation. Both are feasible for local piloting because the

institutional context enables temporary experimentation more easily than broader continuation. Both remain constrained in movement beyond the pilot because of elements like tendering, compliance demands, and fragmented requirements by county.

The study makes three main contributions. First, it shows that in this GovTech pilot context the Valley of Death is best understood as a configurational problem between a pilot stage arrangement and a routine delivery arrangement. This shifts attention away from treating the problem primarily as a funding gap, adoption gap, or technical validation problem like it currently mostly has been treated as. Second, it shows that institutional context is the surrounding condition that determines what is possible for business model design and ecosystem coordination. This only becomes visible when the three lenses are examined together and it explains why a configuration used and works in piloting could still fail in continuation. Third, it refines the role of the ecosystem in analysis of pilot stage scaling. In this case the ecosystem is not a mature surrounding structure already in place. Instead, it reveals the absence of broader intermediation that enables the market and helps explain why broader ecosystem development could begin after business model design has been reconfigured for continuation and embedded delivery.

These contributions support a business configurational view of scaling beyond a pilot in institutionally complex environments. This movement is not explained by any single factor. It depends on if business model design, ecosystem coordination, and institutional conditions align in a way that allows the venture to move from a local and temporary arrangement toward one that is more embedded and repeatable across contexts.

5.3 Practical implications

The findings suggest that moving a digital service innovation past and beyond the pilot stage in a B2G context requires more than just a positive result from the pilot itself. The practical issue is whether the conditions for continuation, embedded delivery, and

broader market access are built around the solution. The implications therefore concern the case company, other startups operating in similar contexts, public sector buyers, and the wider structures that control how solutions move forward.

For Company X and other GovTech startups the main implication is that the pilot cannot be treated only as validation. It should also function as continuation-building. In practice this means using the pilot to to produce the evidence that matter in continuation decisions. In this case, the evidence would be expressed as the effects on staffing, the reduction of workload, and the measurable operational impact. The company should also consider the architecture of delivery as a strategic matter. If integration is incomplete or each new county requires repeated setup work, maintaining the value from the pilot becomes a difficult task. So, the practical implications are clear: building a continuation pathway and making delivery less dependent on the conditions of the pilot. More broadly, startups in similar contexts should design pilots from the beginning to test whether continuation and routine use are institutionally and operationally realistic. Otherwise, the result is a model that might function in a single pilot but remains too fragile for benefits for the company or the buyer in a broader sense.

Additionally, startups must develop deep domain understanding of the sector they are in. Successful solutions come from founders who understand the technology and the operational realities, regulatory constraints, and workflow patterns of public sector organizations. Technical excellence alone does not translate to procurement success when compliance requirements, professional practices, and approval processes remain unclear to the supplier.

For public sector buyers and procurement professionals, the implications are similar to those for startups, that pilots should be structured to test whether the solution could continue under routine conditions. A pilot may generate useful learning while still leaving the continuation route unclear. Buyers can address this by defining expected evidence more clearly from the start, by paying attention to how the service would fit

existing systems, routines, and operational practice if it were to continue and by thinking through the continuation route already before the pilot is initiated. The pilot should not be treated as a separate experiment detached from later decisions. It should be more an early assessment of whether broader use is feasible.

The findings also suggest that the move from pilot to competitive tendering often breaks continuity and causes issues like VoD and more specifically pilotitis. If continuation always returns the process to a full reset, the value of piloting as a pathway to innovation remains limited even when it succeeds. In that situation, the system becomes expensive for everyone involved. The startup spends resources producing proof that does not proceed forward, while the public sector spends time and money that does not lead to sustained solutions, which in turn causes back and forth for professionals using the solution during the pilot. What is needed is a more clarified bridge from pilot to routine contracting. This could include innovative procurement methods such as innovation partnerships that enable longer development relationships, or framework agreements that reduce restart friction when pilots prove successful.

Moreover, procurement professionals should recognize the tension between risk avoidance and innovation. Participation requirements, such as minimum turnover, credit ratings, or reference requirements, designed to minimize procurement risk often systematically exclude the very startups most capable of bringing novel solutions to persistent problems. Balancing institutional security with market access for innovation requires conscious adjustment of procurement criteria and potentially different approaches for established and emerging solutions.

For policymakers and other actors involved in building the public sector environment, the main implication is that this transition should be recognized as a distinct problem requiring real and fast intervention. The findings suggest that the Finnish public sector context enables local piloting relatively well, but broader continuation remains structurally difficult. Supporting more pilots alone is unlikely to solve the scaling gap and

the issues the pilot was procured to solve in the first place. What matters is reducing unnecessary restart between contexts and improving the conditions for validated solutions to move forward.

In practice, this points toward several specific improvements. First, clearer continuation pathways are needed that reduce the gap between pilot and routine. Second, compliance and assessment processes across counties need harmonization. The current situation where each county conducts its own data protection impact assessments, security audits, and AI risk assessments creates major friction and forces the same work to be repeated. A national-level certification or approval mechanism that transfers across counties would reduce market entry barriers while maintaining needed supervision. Third, stronger mechanisms for sharing what has already been validated would prevent duplicated effort and allow proven solutions move forward efficiently.

A critical issue identified in the interviews is the disconnect between investment and outcomes. Although Finland ranks highly in public sector digitalisation and digitalisation is framed as a key enabler of public sector renewal, publicly visible evidence on how these efforts translate into measurable performance improvements remains limited. This suggests a systemic problem with how digital investments are evaluated, procured, and implemented. The procurement process design itself plays a decisive role in what can scale and what cannot. Without changes such as these, successful pilots may demonstrate value without progression, which drains the resources of both the startup and the county without creating lasting benefit for either.

Furthermore, digital leadership remains underrepresented in public sector decision-making structures. Very few healthcare regions have representation in that level in management, meaning ICT is most times positioned more as a support. This affects individual procurement decisions and the broader organizational capacity to understand and derive value from digital innovation. Taking on the scaling challenge therefore requires process changes and developing capabilities in organizations.

Another implication is the broader ecosystem that remains largely absent in the case. The study suggests that scaling depends on direct buyer-supplier relationships, but also on intermediating actors and structures that help solutions travel across contexts. These include actors who reduce entry friction, connect validated solutions to new settings, and lower the need to rebuild relationships and processes from the beginning in each county. The practical challenge is therefore to make market access less dependent on local relationships. The interviews highlighted the potential role of intermediary organizations, such as industry associations, disability advocacy organizations, and third-sector service providers, in creating pathways to scale. These organizations may serve both as alternative customer segments and as channels through which innovations eventually reach public sector organizations through demonstrated effectiveness and advocacy rather than through direct procurement processes alone.

When broader market-enabling structures remain weak, scaling is more likely to stay constrained to effort that is specific to cases and the transition from pilot to routine delivery keeps being individually negotiated when it could be more systematically supported. This again raises costs on both sides, because each new case requires too much repeated work instead of building on what has already been learned. More broadly, scaling requires higher-level coordination around market access, shared compliance interpretation, common standards, and cross-learning between contexts. These roles need to be designed and will not appear automatically.

These implications reinforce the central finding of the thesis. In institutionally complex settings such as GovTech, progression beyond the pilot does not follow automatically from technical success. It depends on whether continuation can be made institutionally supportable, whether delivery can be configured for routine use, and whether broader ecosystem structures begin to reduce the friction of market entry. The Finnish context demonstrates that while local innovation is possible, systemic barriers to scaling persist. Addressing these barriers requires coordinated action across multiple levels: startup

capability development, procurement practice evolution, institutional process harmonization, and ecosystem infrastructure creation. Without such coordination, the gap between pilot success and market scaling will continue to limit both innovation adoption and public sector modernization.

5.4 Limitations and future research

Multiple limitations define the generalizability of the findings. Firstly, the study is based on only a single case. While Company X provides rich access to the pilot stage configuration and the institutional dynamics that surround it, being a single-case design limits the extent findings can be generalized cross different GovTech ventures, domains or institutional contexts. The framework developed here would benefit from testing across multiple cases to analyze whether the scaling-relevant focus areas identified hold across different settings, or whether different configurations produce different bottlenecks. Future research with comparative case designs could examine if the gap in pilot towards continuation operates similar across service types, buyer organizations, or national regulatory environments. Comparing multiple counties would also reveal how variation at the county level affects the dynamics of scaling and whether some of certain county configurations enable progression more effectively than others.

Second, the case is at the pilot stage and scaling is analyzed as a feasibility question that looks forward in time rather than as an outcome that can be observed. This choice was deliberate and made intentionally, because the pilot stage is where the Valley of Death becomes structurally visible in this context. But it also means we cannot directly observe whether the identified focus areas here actually determine scaling success in practice. Future research should examine cases that have successfully moved beyond the pilot stage, and as well cases that attempted the transition and failed to validate if the focus areas identified here explain scaling outcomes in real situations. Following Company X as it attempts to move beyond the one pilot project would give value into the conversation of how the configuration evolves during the transition.

Third, the empirical material is composed of five semi-structured interviews and some supporting documents from the case company. While the interviews were substantial in length and covered multiple actor perspectives, the study would benefit from additional data sources. Future research could strengthen findings with longitudinal observation of these pilot to scale transitions, broader interviews with stakeholders that could include system providers and end-users, and analysis of pilot continuation rates and scaling patterns across the Finnish wellbeing services county context. Expanding the data would also allow examination of how the configuration evolves over time as ventures attempt to move from pilots toward routine delivery.

Fourth, actor level coverage is incomplete due to difficulty accessing stakeholders. The study captures perspectives well from the case company, the buyer organization, and an external expert but it does not include direct interviews with system suppliers, more national level regulators, end-users, or procurement professionals from other counties. It is likely that these actors have perspectives that could reveal additional dynamics. Particularly regarding system integration constraints, regulatory interpretation across contexts, user experience with the service, and variation in how different counties approach innovation procurement. Future research should expand actor coverage to capture these missing perspectives and examine how they shape ecosystem coordination challenges.

These limitations suggest directions for future research. Table 2 presents specific research questions that would extend and validate the findings of this thesis.

Analytical lens	Future research direction	Illustrative future research question
Business Model Design	**Value creation and delivery**	What makes a delivery model scalable across public sector contexts?
	Value capture	How do procurement rules affect viable pricing models?
	Business model development	How can startups use public sector entry to build pathways into private sector markets?

Ecosystem Coordination	**Emergence of broader ecosystem**	How does a broader ecosystem begin to form around a validated digital service innovation in a public sector context? Under what conditions do complementary actors engage early around a new digital service innovation in the public sector?
	Missing intermediaries	What kinds of intermediating actors are needed to reduce repeated entry, compliance, and integration work in public sector contexts?
Institutional Context	**Procurement decision processes**	How do procurement decision processes shape the relative position of startups, large incumbents, and in-house actors in public sector contexts?
	Procurement logic and innovation	How do current public procurement logics support or constrain innovation-oriented procurement?
	Public sector innovation capability	What kind of capabities are needed inside public sector organizations to procure, continue, and scale novel digital services?
	Wellbeing services county fragmentation	How do county-level differences in compliance interpretation, audit practice, and procurement requirements affect innovative procurement in B2G contexts? What forms of harmonization could better support innovative procurement across counties in B2G contexts?

Table 2. Future Research Directions & Questions

Future research addressing these questions would strengthen understanding of how digital service innovations navigate institutionally complex environments and would provide more robust guidance for ventures and buyers, and policymakers seeking to improve the conditions for scaling beyond the pilot stage. Longitudinal studies following ventures through the actual transition, comparative studies examining variation across counties and contexts, and focused investigation of intermediary roles and procurement pathway design would be particularly valuable for validating and extending the configurational framework developed in this thesis.

For transparency on AI LLM usage, Claude Sonnet 4.5 (Anthropic) and ChatGPT (OpenAI; GPT-5-family models) were used as supporting tools during the research and writing process. Their use was limited to language refinement, clarification of expression, simplification of sentence structures, testing of alternative phrasings, summarization of

already-selected literature for the author’s own assessment, and support in structuring written argumentation. The research design, literature selection, theoretical framing, data analysis, interpretation of findings, and final conclusions were all performed independently by the author. No new empirical data, academic sources or theoretical claims generated by these tools were included in the thesis without verification from original sources and full independent evaluation by the author. The tools therefore functioned only as an aid for writing and reflection, while responsibility for all content, interpretations, and academic judgments remains fully with the author.

6 Conclusion

This thesis examined what is essential for scaling digital service innovations by analyzing the dynamics and configuration of business model design, ecosystem coordination, and institutional context in a B2G pilot setting. This was done through a case study of a GovTech startup piloting an AI-assisted preliminary information collection service in a Finnish wellbeing services county. The focus was on the Valley of Death as a configurational challenge that becomes visible in the pilot phase.

The findings show that the Valley of Death in this setting is a configurational problem between two different states. The pilot represents a locally working and temporary configuration, whereas routine delivery requires a different configuration that is more repeatable and embedded. In total, three key scaling-relevant focus areas became apparent from the case analysis. First, moving past the pilot requires building a continuation pathway that helps make the outputs of the pilot relevant for procurement and connects them to a viable route forward. Second, progress past the pilot stage requires reconfiguring the delivery so that information moves more seamlessly into existing systems and processes and the service is not dependent on customized setup each time it is introduced into a new environment. Third, a broader ecosystem layer beyond the local pilot project arrangement, including intermediating actors that can reduce entry friction, support market access, and help the solution travel across counties is also needed.

The study responds to the fragmented VoD and pilotitis literature by showing how the problem can be analyzed in one integrated way and by bringing more attention to the manifestations of VoD in the B2G digital service context. It also addresses the underdeveloped startup-oriented perspective identified in the research gap by showing how progression past the pilot depends on how business model design and ecosystem coordination are configured under the institutional conditions. In addition, it further clarifies the role of institutional context by showing that procurement rules, evaluation

criteria, and participation barriers surround scaling and shape what elements become central in the business configuration.

These findings have practical implications. For startups, the pilot cannot function only as validation but also build a continuation route. For public sector buyers, the pilot should not be a separate technology experiment that is completely detached from later decisions and should test if routine use and continuation are feasible or possible at all. For policymakers, the results point to clearer continuation paths and less fragmented post-pilot conditions. This is more consistent practices across counties, less repeated reassessment of already validated solutions, and stronger structures for actually carrying successful pilots into routine use. This is important in order to make innovative startups able to contribute to the renewal of the public sector, reduce spending on pilots that do not continue forward, and improve the chances that public procurement leads to real, lasting improvements.

This study is limited by being a single-case design and focusing on only one pilot stage configuration at only one point in time. Future research would benefit from comparing B2G cases and longitudinal research following what happens after the pilot stage. In addition, further research is needed on the broader ecosystem that may form around solutions that are validated, the intermediary actors that could reduce repeated entry and compliance work, and the public sector and procurement processes that build the continuation in real contexts. Overall, this thesis shows that continuing past the pilot stage goes past whether the solution works and extends to whether the business configuration supports continuation.

References

Adner, R. (2017). Ecosystem as Structure: An Actionable Construct for Strategy. Journal of Management. https://doi.org/10.1177/0149206316678451

Adler, R. H. (2022). Trustworthiness in Qualitative Research. Journal of human lactation, 38(4), 598-602. https://doi.org/10.1177/08903344221116620

Akenroye, T. O., Oyedijo, A., Abubakre, A., Zsidisin, G., El Baz, J., & Vishnu, C. R. (2024). One size does not fit all: Deciphering the interdependence between barriers hindering SMEs' involvement in public sector procurement. The International journal of public sector management, 37(7), 935-965. https://doi.org/10.1108/IJPSM-06-2024-0184

Aksom, H., & Vakulenko, V. (2024). Revisiting the scope and suggesting novel domains of institutional theory in the public administration research. Teaching public administration, 42(2), 230-250. https://doi.org/10.1177/01447394231191935

Amit, R., & Zott, C. (2001). Value creation in e-business. Strategic Management Journal. https://doi.org/10.1002/smj.187

Arlbjørn, J. S., & Freytag, P. V. (2012). Public procurement vs private purchasing: Is there any foundation for comparing and learning across the sectors? The International journal of public sector management, 25(3), 203-220. https://doi.org/10.1108/09513551211226539

Autio, E., & Thomas, L. D. (2022). Researching ecosystems in innovation contexts. *Innovation & Management Review, 19*(1), 12-25. https://doi.org/10.1108/INMR-08-2021-0151

Baden-Fuller, C., & Haefliger, S. (2013). Business Models and Technological Innovation. Long Range Planning. https://doi.org/10.1016/j.lrp.2013.08.023

Bhatia, A., Matthan, R., Khanna, T., & Balsari, S. (2020). Regulatory sandboxes: A cure for mHealth pilotitis? *Journal of Medical Internet Research*. https://doi:10.2196/21276

Bharosa, N. (2022). The rise of GovTech: Trojan horse or blessing in disguise? A research agenda. Government Information Quarterly. https://doi.org/10.1016/j.giq.2022.101692

Bitektine, A., & Song, F. (2023). On the Role of Institutional Logics in Legitimacy Evaluations: The Effects of Pricing and CSR Signals on Organizational Legitimacy. Journal of management, 49(3), 1070-1105. https://doi.org/10.1177/01492063211070274

Bonnin Roca, J., & O'Sullivan, E. (2020). Seeking coherence between barriers to manufacturing technology adoption and innovation policy, International Journal of Production Economics. https://doi.org/10.1016/j.ijpe.2020.107818

Breaugh, J., McBride, K., Kleinaltenkamp, M., & Hammerschmid, G. (2021). Beyond Diffusion: A Systematic Literature Review of Innovation Scaling. Sustainability, 13(24), 13528. https://doi.org/10.3390/su132413528

Calza F., Ferretti M., Panetti E., Parmentola A. (2021). Moving drug discoveries beyond the valley of death: the role of innovation ecosystems. European Journal of Innovation Management. https://doi.org/10.1108/EJIM-11-2019-0342

Chamberlain, S., Dutt, P., Godfrey, A., Mitra, R., LeFevre, A. E., Scott, K., Mendiratta, J., Chauhan, V., & Arora, S. (2021). Ten lessons learnt: Scaling and transitioning one of the largest mobile health communication programmes in the world to a national government. BMJ global health, 6(Suppl 5), e005341. https://doi.org/10.1136/bmjgh-2021-005341

Chesbrough, H., & Rosenbloom, R. S. (2002). The role of the business model in capturing value from innovation: evidence from Xerox Corporation's technology spin-off companies. Industrial and Corporate Change. https://doi.org/10.1093/icc/11.3.529

Clarysse, B., Wright, M., Bruneel, J., & Mahajan, A. (2014). Creating Value in Ecosystems: Crossing the Chasm between Knowledge and Business Ecosystems. Research Policy. https://doi.org/10.1016/j.respol.2014.04.014

Currie, G., Lynch, A., Swan, J., Alderson, H., Friel, S., Harrop, C., Johnson, R., Kerridge, G., Smart, D., McGovern, R., & Munro, E. (2025). How to extend pilot innovation in public services: A case of children's social care innovation. Public Administration. https://doi.org/10.1111/padm.13028

Creswell, J. W., & Poth, C. N. (2023). Qualitative inquiry and research design: Choosing among five approaches. Thousand Oaks, California: Sage.

Crowe, S., Cresswell, K., Robertson, A., Huby, G., Avery, A., & Sheikh, A. (2011). The case study approach. *BMC medical research methodology, 11*(1), 100. https://doi.org/10.1186/1471-2288-11-100

Damschroder, L. J., Aron, D. C., Keith, R. E., Kirsh, S. R., Alexander, J. A., & Lowery, J. C. (2009). Fostering implementation of health services research findings into practice: A consolidated framework for advancing implementation science. Implementation Science. https://doi.org/10.1186/1748-5908-4-50

Denis, J. L., Hébert, Y., Langley, A., Lozeau, D., & Trottier, L. H. (2002). Explaining diffusion patterns for complex health care innovations. Health Care Management Review. Retrieved 6.12.2025 from https://journals.lww.com/hcmrjournal/abstract/2002/07000/explaining_diffusion_patterns_for_complex_health.7.aspx

DiCicco-Bloom, B., & Crabtree, B. F. (2006). The qualitative research interview. *Medical education, 40*(4), 314-321. https://doi.org/10.1111/j.1365-2929.2006.02418.x

Dubois, A., & Gadde, L. (2002). Systematic combining: An abductive approach to case research. Journal of business research, 55(7), 553-560. https://doi.org/10.1016/S0148-2963(00)00195-8

Earle, A., Leyva-de la Hiz, D. I., & Turell, Y. (2019). Hybrid business models and the public science-private industry interface. Journal of business models, 7(4), 20-26. https://doi.org/10.5278/ojs.jbm.v7i4.2921

Egermark, M., Blasiak, A., Remus, A., Sapanel, Y., & Ho, D. (2022). Overcoming pilotitis in digital medicine at the intersection of data, clinical evidence, and adoption. *Advanced Intelligent Systems.* https://doi.org/10.1002/aisy.202200056

Eisenhardt, K. M. (1989). Building Theories from Case Study Research. *The Academy of Management review, 14*(4), 532. https://doi.org/10.2307/258557

Ellwood, P., Williams, C., Egan, J. (2022). Crossing the valley of death: Five underlying innovation processes. Technovation. https://doi.org/10.1016/j.technovation.2020.102162

Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of advanced nursing, 62*(1), 107-115. https://doi.org/10.1111/j.1365-2648.2007.04569.x

Engin, Z., & Treleaven, P. (2019). Algorithmic Government: Automating Public Services and Supporting Civil Servants in using Data Science Technologies. The Computer Journal. https://doi.org/10.1093/comjnl/bxy082

European Commission. (2025). Europe's budget – Horizon Europe (2028-2034), Publications Office of the European Union,. Retrieved January 2, 2026, from https://research-and-innovation.ec.europa.eu/funding/funding-opportunities/funding-programmes-and-open-calls/horizon-europe_en Research and innovation

European Innovation Council. (2025). *EIC Accelerator*. European Innovation Council. Retrieved January 2, 2026, from https://eic.ec.europa.eu/eic-funding-opportunities/eic-accelerator_en

Gbadegeshin, S. A., Al Natsheh, A., Ghafel, K., Mohammed, O., Koskela, A., Rimpiläinen, A., Tikkanen, J., & Kuoppala, A. (2022). Overcoming the Valley of Death: A new model for high technology startups. Sustainable Futures. https://doi.org/10.1016/j.sftr.2022.100077

Gerhardt, U., Breitschwerdt, R. & Thomas, O. (2017). Engineering sustainable mHealth: the role of Action Research. https://doi.org/10.1007/s00146-015-0640-5

Greenhalgh, T., & Papoutsi, C. (2019). Spreading and scaling up innovation and improvement. The BMJ. https://doi.org/10.1136/bmj.l2068

Greenwood, R., Raynard, M., Kodeih, F., Micelotta, E. R., & Lounsbury, M. (2011). Institutional Complexity and Organizational Responses. The Academy of Management annals, 5(1), 317-371. https://doi.org/10.1080/19416520.2011.590299

Halinen, A., & Törnroos, J. (2005). Using case methods in the study of contemporary business networks. Journal of business research, 58(9), 1285-1297. https://doi.org/10.1016/j.jbusres.2004.02.001

Hartley, P. R., & Medlock, K. B., III. (2017). The Valley of Death for New Energy Technologies. The Energy Journal. https://doi.org/10.5547/01956574.38.3.phar

Hoekstra, M., Chideock, C., Bodea, G., Bharosa, N., Cave, J., Kokkeler, B., Oomens, I., Van den Berg, A., Van Veenstra, A. F., & Alleweldt, F. (2022). The Digital Single Market and the digitalisation of the public sector: GovTech and other innovations in public procurement. European Parliament. https://doi.org/10.2861/830794

Huang, F., Blaschke, S., & Lucas, H. (2017). Beyond pilotitis: Taking digital health interventions to the national level in China and Uganda. *Globalization and Health.* https://doi.org/10.1186/s12992-017-0275-z

Karjalainen, A.-M. (2020). *Digitalisation is a key enabler for public sector renewal*. Debt Management Annual Review 2020. Treasury of Finland. Derived 6.4.2026 from: https://www.treasuryfinland.fi/annualreview2020/digitalisation-is-a-key-enabler-for-public-sector-renewal/

Klitsie, J.B., Price, R.A. & De Lille, C.S.H. (2019). Overcoming the Valley of Death: A Design Innovation Perspective. Design Management Journal. https://doi.org/10.1111/dmj.12052

Kowalkowski, C., Wirtz, J., & Ehret, M. (2024). Digital service innovation in B2B markets. Journal of Service Management. https://doi.org/10.1108/JOSM-12-2022-0403

Lefebvre, V., Certhoux, G., & Radu-Lefebvre, M. (2022). Sustaining trust to cross the Valley of Death: A retrospective study of business angels' investment and reinvestment decisions. Technovation. https://doi.org/10.1016/j.technovation.2020.102159

Jucevicius, G., Juceviciene, R., Gaidelys, V., & Kalman, A. (2016). The Emerging Innovation Ecosystems and "Valley of Death": Towards the Combination of Entrepreneurial and Institutional Approaches. Inzinerine Ekonomika-Engineering Economics. https://doi.org/10.5755/j01.ee.27.4.14403

Linde, L., Sjödin, D., Parida, V., & Wincent, J. (2021). Dynamic capabilities for ecosystem orchestration A capability-based framework for smart city innovation initiatives. Technological forecasting & social change, 166, 120614. https://doi.org/10.1016/j.techfore.2021.120614

Markham, S. K. (2002). Moving Technologies from Lab to Market. Research-Technology Management. https://doi.org/10.1080/08956308.2002.11671531

Markham, S.K., Ward, S.J., Aiman-Smith, L. and Kingon, A.I. (2010), The Valley of Death as Context for Role Theory in Product Innovation. Journal of Product Innovation Management. https://doi.org/10.1111/j.1540-5885.2010.00724.x

Martin, GP., Weaver, S., Currie, G., Finn, R., McDonald, R. (2013). Innovation sustainability in challenging health-care contexts: Embedding clinically led change in routine practice. Health Services Management Research. https://doi.org/10.1177/0951484812474246

Mcintyre RA. (2014). Overcoming "The Valley of Death". Science Progress. https://doi.org/10.3184/003685014X14079421402720

Melander, L., & Arvidsson, A. P. (2020). Getting innovations out of interactions in the public procurement context. The Journal of business & industrial marketing, 35(12), 2051-2065. https://doi.org/10.1108/JBIM-09-2019-0414

Mergel, I., Ulrich, P., Kuziemski, M., & Martinez, A. (2022). Scoping GovTech dynamics in the EU. Publications Office of the European Union. https://doi.org/10.2760/700544

Merisalo, M., Oksanen, J., Valovirta, V. & Pihlajamaa, M. (2022). Innovatiivisten julkisten hankintojen yhteiskunnallisten ja yritysvaikutusten skaalautuminen. Business Finland. Retrieved 2.12.2025 from https://www.businessfinland.fi/globalassets/julkaisut/policy_brief_6-2022_innohankinnat-vaikutukset-ja-skaalautuminen_final.pdf#:~:text=innovaatiokumppanuudessa,T%C3%A4m%C3%A4%20koskee%20erityisesti%20tilanteita%2C%20joissa

Obwegeser, N., & Müller, S. D. (2018). Innovation and public procurement: Terminology, concepts, and applications. Technovation, 74-75, 1-17. https://doi.org/10.1016/j.technovation.2018.02.015

Pickernell, D., Kay, A., Packham, G., & Miller, C. (2011). Competing Agendas in Public Procurement: An Empirical Analysis of Opportunities and Limits in the UK for SMEs. Environment and planning. C, Government & policy, 29(4), 641-658. https://doi.org/10.1068/c10164b

Pollock, N., & Hyysalo, S. (2014). The Business of Being a User: The Role of the Reference Actor in Shaping Packaged Enterprise System Acquisition and Development. MIS Quarterly. Retrieved 15.12.2025 from https://www.jstor.org/stable/26634935

Popp, J., Matthews, D., Martinez-Coll, A., Mayerhöfer, T., Wilson, B. C. (2017). Challenges in translation: models to promote translation. Journal of Biomedical Optics. https://doi.org/10.1117/1.JBO.23.2.021101

Pratt, M. G. (2009). From the Editors: For the Lack of a Boilerplate: Tips on Writing Up (and Reviewing) Qualitative Research. *Academy of Management journal, 52*(5), 856-862. https://doi.org/10.5465/amj.2009.44632557

Saastamoinen, J., Reijonen, H., & Tammi, T. (2018). Should SMEs pursue public procurement to improve innovative performance? Technovation, 69, 2-14. https://doi.org/10.1016/j.technovation.2017.10.003

Scarbrough, H., & Kyratsis, Y. (2022). From spreading to embedding innovation in health care: Implications for theory and practice. Health Care Management Review. https://doi.org/10.1097/HMR.0000000000000323

Scarbrough, H., Sanfilippo, K., Ziemann, A., Stavropoulou, C. (2024). Mobilizing pilot-based evidence for the spread and sustainability of innovations healthcare: The role of innovation intermediaries. Social Science & Medicine. https://doi.org/10.1016/j.socscimed.2023.116394

Schiffelers TELM, Hochstenbach LMJ, Talboom-Kamp E, Sanders-van Wijk S, Kietselaer BLJH, Spreeuwenberg MM. (2025). Addressing pilotitis: A qualitative study on the implementation of eCoaches in hospital-based chronic care. Digital Health. https://doi.org/10.1177/20552076251374210

Shenton, A. K. (2004). Strategies for ensuring trustworthiness in qualitative research projects. *Education for information, 22*(2), 63-75. https://doi.org/10.3233/EFI-2004-22201

Suchman, M. C. (1995). Managing legitimacy: Strategic and institutional approaches. The Academy of Management review, 20(3), 571. https://doi.org/10.2307/258788

Teece, D. J. (2010). Business Models, Business Strategy and Innovation. Long Range Planning. https://doi.org/10.1016/j.lrp.2009.07.003

Tukiainen, J., Baulia, S., & Matinheikki, J. (2023). *Hyvinvointialueiden julkiset hankinnat tarvitsevat lisää kilpailua* (VATT Policy Brief 2/2023). Valtion taloudellinen tutkimuskeskus. Retrieved 2.12.2025 from https://vatt.fi/hyvinvointialueiden-julkiset-hankinnat-tarvitsevat-lisaa-kilpailua

Uyarra, E., Edler, J., Garcia-Estevez, J., Georghiou, L., & Yeow, J. (2014). Barriers to innovation through public procurement: A supplier perspective. Technovation, 34(10), 631-645. https://doi.org/10.1016/j.technovation.2014.04.003

Vickers, I., Lyon, F., Sepulveda, L., & McMullin, C. (2017). Public service innovation and multiple institutional logics: The case of hybrid social enterprise providers of health and wellbeing. Research policy, 46(10), 1755-1768. https://doi.org/10.1016/j.respol.2017.08.003

Vohora, A., Wright, M., & Lockett, A. (2004). Critical junctures in the development of university high-tech spinout companies. Research Policy. https://doi.org/10.1016/S0048-7333(03)00107-0

Wirtz, B. W., Pistoia, A., Ullrich, S., & Göttel, V. (2016). Business Models: Origin, Development and Future Research Perspectives. Long Range Planning. https://doi.org/10.1016/j.lrp.2015.04.001

Wu, X., Tan, X., & Wang, X. (2023). The institutional logics perspective in management and organizational studies. Journal of business research, 167, 114183. https://doi.org/10.1016/j.jbusres.2023.114183

Appendices

Appendix 1. Interview structure & interviewee-specific tailoring

Table A. Common interview structure

Section	Core content
Roolin tarkennus	Voisitko kertoa lyhyesti omasta roolistasi ja siitä, missä kohdassa olet mukana pilottiin, käyttöönottoon tai laajempaan kehittämiseen liittyvissä kysymyksissä? Miten oma roolisi liittyy uusien digitaalisten ratkaisujen arviointiin, käyttöönottoon tai jatkoon?
Liiketoimintamalli	**Pilotin tarkoitus, arvolupaus ja hyöty:** Mitä pilotilla ensisijaisesti haetaan? Mikä ongelma pyritään ratkaisemaan? Mistä huomataan, että ratkaisu ei ole vain kokeilu vaan oikeasti hyödyllinen? **Toimituslogiikka ja skaalautuvuus:** Onko toteutuksessa jotain liian raskasta, manuaalista tai vaikeasti toistettavaa? Mitä pitäisi muuttaa, jotta ratkaisu voisi toimia laajemmassa ja vakiintuneemmassa käytössä? **Kustannukset, hinnoittelu ja jatkuvuus:** Miten kustannus- ja hinnoittelulogiikka muuttuu, jos käyttöä laajennetaan? Millainen jatko on realistinen pilotin jälkeen?
Ekosysteemi	**Toimijat ja roolit:** Ketkä ovat pilotin kannalta keskeiset toimijat? Mitkä roolit ovat välttämättömiä myös laajemmassa käytössä? **Riippuvuudet, koordinointi ja päätöksenteko:** Missä riippuvuudet ovat kriittisimmät? Miten päätöksiä tehdään käytännössä? Missä kohdissa yhteistyöhön voi syntyä kitkaa tai pullonkauloja? **Käyttöönotto, juurtuminen ja integraatiot:** Miten ratkaisu kytkeytyy muihin järjestelmiin, prosesseihin ja toimijoihin? Mitä käyttöönotto ja juurtuminen vaativat? Mikä estää ratkaisun siirtymistä ympäristöstä toiseen?
Institutionaalinen konteksti	**Hankintalogiikka ja sääntely:** Miten hankintalogiikka, sopimusmallit ja sääntely vaikuttavat siihen, mitä voidaan tehdä pilotin jälkeen? Millaiset jatkoreitit ovat realistisia? **Riskit, hyväksyntä ja julkisen sektorin reunaehdot:** Millaisia tietosuojaan, tietoturvaan, auditointeihin tai tekoälyyn liittyviä ehtoja tai vaiheita etenemiseen liittyy? Miten julkisen sektorin riskienhallinta ja varovaisuus näkyvät tällaisissa ratkaisuissa? **Jatkopolku, kilpailutus ja siirrettävyys:** Miksi hyvä pilotti ei välttämättä etene pysyvämpään käyttöön? Mitkä tekijät vaikeuttavat suoraa jatkoa pilotista laajempaan käyttöönottoon? Mitä pitäisi tapahtua, jotta pilotista jäisi pysyvämpi hyöty?
Lopetus	Jos saisit muuttaa yhden asian tässä toimintaympäristössä, mikä se olisi?

Table B. Interviewee-specific tailoring

Interviewee	Emphasis	Illustrative example question
P1 (Founder)	Concrete pilot configuration, pricing logic, ecosystem dependencies, realistic continuation routes	Voitko kertoa yleisellä tasolla, mistä teidän pilotin juoksevat kustannukset syntyvät?
P2 (Procurement lead)	Procurement procedures, threshold logic, continuation constraints, participation conditions, repeated approvals	Miten teillä käytännössä päätetään, milloin voidaan edetä alle 60 tuhannen kynnysarvon hankintana ja milloin pitää valmistella laajempi kilpailutus?
P3 (Service area lead)	Reform needs, pilot usefulness, continuation difficulties, embedding into daily work	Mitä pitää olla kunnossa, että toimiala uskaltaa sitoutua laajempaan käyttöönottoon eikä jäädä vain kokeilutasolle?
P4 (External expert)	System-level pilot barriers, institutional decision logic, embedding problems, transferability across areas	Kun hyvinvointialue tekee pilotin tekoälystä tai digipalvelusta, missä kohtaa eteneminen useimmin hidastuu tai pysähtyy?
P5 (Board member)	Growth logic, commercialization, competitive strategy, proof of value	Mikä on pilotin jälkeen teidän tavoite ansaintamallissa ja logiikassa?

Appendix 2. Analytical coding by lens and dimension

Table A. Business Model Design (Micro)

Analytical dimensions (9)	Main themes (9)	Example Subthemes (27)	Example Codes (205)
Value Proposition	Teema 1: Pilotin tarkoitus, oppiminen ja hyötyjen todentaminen	Miksi pilotoidaan ja mitä pilotilla haetaan; Pilotissa haettava näyttö ja takaisinmaksulogiikka	“Pilottiprojekti – Sopimuksessa tunnistettiin hyötyhypoteesit, joista tehtiin mittareita tilaajan seurantaan”; “Pilotin tarkoitus – Pilotilla pitäisi kokeilla jotain, mitä ei ole aiemmin tehty samalla tavalla, ja samalla arvioida tuottavuutta, laatua, tehokkuutta, käytettävyyttä ja luotettavuutta”; “Mittarointi ja pilotin rooli – Asiakastyytyväisyys ja tehokkuus voidaan kellottaa, kysyä ja visualisoida, pilotin aikana mittarointi nähdään realistisena”
Value Creation & Delivery	Teema 4: Teknologia, tekoäly, järjestelmät ja integraatio-ongelmat	Pilotin tekniset ja toiminnalliset haasteet; Järjestelmien yhteensopivuus ja integraatiot	“Haasteet – Skaalauksen kannalta iso kysymys on, miten palvelun tuottama data saadaan uimaan sujuvasti ammattilaisten prosesseihin ja HVA:n järjestelmiin”; “Pilotin räätälöinti – Copy-paste-tyyppinen kirjaaminen kertoo siitä, että data ei vielä siirry prosessiin aidosti, vaan pilotissa joudutaan paikkaamaan integraation puutteita manuaalisesti”; “Haasteet – Eri hyvinvointialueilla on omat järjestelmänsä, joten sama ratkaisu ei siirry sellaisenaan ympäristöstä toiseen”
Value Capture	Teema 2: Pilotista jatkuvuuteen, skaalaamiseen ja jatkopolun katkokseen liittyvät ongelmat	Skaalaaminen ei ole automaattinen jatke pilotille; Vaihtoehtoiset tulonmuodostuksen ja kasvun polut	“Ansaintalogiikka – Suositeltavana mallina transaktiopohjaisuus”; “Hybridihinnoittelu – Kiinteä perusosa + muuttuva osa + marginaali”; “Vaihtoehtoinen bisnesmalli olisi myydä kehitystyötä eikä tuotetta”; “Julkinen sektori ei riitä liiketoiminnan pohjaksi”

Table B. Ecosystem Coordination (Meso)

Analytical dimensions (9)	Main themes (9)	Example Subthemes (27)	Example Codes (205)
Alignment	Teema 6: Ekosysteemi, toimijat, päätöksenteko ja toimittajariippuvuus	Ekosysteemin keskeiset toimijat; Ekosysteemin riippuvuudet ja päätöksenteon rakenne	“Ekosysteemin toimijat – Teknologiakumppani on kriittinen toimija, koska tekninen toteutus nojaa siihen”; “Päätöksenteko ekosysteemissä – Pilotin keskeiset peruspäätökset siitä mitä tehdään on tehty jo sopimuksessa”; “Päätöksenteko ekosysteemissä – HVA tekee pilotin sisällölliset ja toiminnalliset päätökset, case-yritys koordinoi, arvioi lisäykset ja vie lopullisen toteutuksen teknologiakumppanille”
Orchestration	Teema 5: Juurtuminen, työn muutos ja käyttöönoton arjen todellisuus	Juurtuminen vaatii johdettua muutosta; Käyttöönotto tarvitsee käytännön tukea ja sisältä tulevaa legitimiteettiä	“Johto ei voi vain päättää muutosta valmiiksi – Pelkkä johdon ymmärrys siitä, että ratkaisu on hyvä, ei riitä”; “Juurruttaminen vaatii käytännön tukea – Käyttöönotto vaatii konkreettista arjen tukea”; “Sisältä tuleva esimerkki on uskottavin tapa levittää uutta”; “Leviäminen ei tapahdu käskemällä – Laajempi muutos ei etene niin, että johto vain käskee”

Intermediation	Teema 4: Teknologia, tekoäly, järjestelmät ja integraatio-ongelmat; Teema 6: Ekosysteemi, toimijat, päätöksenteko ja toimittajariippuvuus	Järjestelmien yhteensopivuus ja integraatiot; Ekosysteemin riippuvuudet ja päätöksenteon rakenne	"Rajapintastrategia – Suosituksena yksi oma rajapinta"; "Integraatioiden varasuunnitelma – Jos rajapintoja ei saada, vaihtoehtona oma käyttöliittymä ja manuaalinen tiedonsiirto"; "Pääjärjestelmien toimittajia on vain muutama"; "Päätöksenteko ekosysteemissä – case-yritys koordinoi, arvioi lisäykset ja vie lopullisen toteutuksen teknologiakumppanille"

Table C. Institutional Context (Macro)

Analytical dimensions (9)	Main themes (9)	Example Subthemes (27)	Example Codes (205)
Interaction Conditions	Teema 3: Hankintalogiikka, sääntely ja institutionaalinen ympäristö; Teema 2: Pilotista jatkuvuuteen, skaalaamiseen ja jatkopolun katkokseen liittyvät ongelmat	Hankinnan rooli portinvartijana ja strategisena ohjaajana; Kynnysarvot, pilkkominen ja hankinnan tulkinnallisuus; Pilotin ja jatkon välinen katkos	"Hankinnan rooli portinvartijana – Hankinta ei ole vain tekninen tukitoiminto vaan portinvartija"; "60 000 euron pilotin etenemisehto – Pilotin voi tehdä ilman kilpailutusta, jos perusteet on kuvattu"; "Pilkkomisen kielto – Hyvinvointialue pyrkii estämään sen, että pilotti rakennetaan lainvastaisesti osiin"; "Pilotin jatkon puute liittyy hankintalakiin – Pilotin jälkeen ei synny automaattista jatkoa"; "Pilotin onnistuminen ei takaa jatkoa samalle toimijalle"
Legitimacy Conditions	Teema 3: Hankintalogiikka, sääntely ja institutionaalinen ympäristö; Teema 9: Toimialan uudistamistarve ja johtamisen näkökulma	Julkisen toimijan riskilogiikka; Uudistamisen ajurit toimialalla	"Julkinen näkyvyys lisää varovaisuutta – Pilottipäätökset näkyvät julkisina viranhaltijapäätöksinä"; "Pienen toimijan valinta vaatii rohkeutta – Startupin valitseminen isoon käyttöön on viranhaltijalle ja toimielimelle riskipäätös"; "Isojen toimijoiden etulyöntiasema legitimiteetissä – Suuret toimijat saavat enemmän anteeksi"; "Taloustilanne ohjaa uudistamista"; "Kaksoistavoite uudistamisessa – Uusilla ratkaisuilla haetaan samaan aikaan sekä tuottavuutta että asiakkaan näkökulmasta parempaa palvelua"
Resourcing Conditions	Teema 7: Startupin asema, markkinoille pääsy ja toimialan ymmärryksen merkitys; Teema 8: Julkisen sektorin laajempi kehittämisympäristö ja systeemitason ongelma	Startupin haavoittuva asema julkisessa ympäristössä; Julkisen sektorin kestävyysvaje, digi-investoinnit ja vaikutusten epäselvyys; IT:n asema ja julkisen organisaation kehittämiskyvykkyys	"Osallistumisehtojen aluekohtaisuus – Liikevaihto-, luottoluokitus- ja referenssivaatimukset voivat sulkea startupit kokonaan ulos"; "Kilpailutukseen osallistuminen vaatii muutakin kuin teknistä osaamista – Pelkkä hyvä tuote ei riitä"; "Julkisen digi-investoinnin vastine on vaikea osoittaa"; "IT näyttäytyy hyvinvointialueella tukena, ei ytimenä"