# Evaluating Impact Investing through a Systems Thinking Lens: Hallmarks of a Transformational Approach

By

**Yu (Sherry) Zhang**

M.S. Construction Economics and Management, University College London, 2017
B.B. International Business, University of Technology Sydney, 2016

SUBMITTED TO THE MIT SLOAN SCHOOL OF MANAGEMENT IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE IN MANAGEMENT STUDIES
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

MAY 2025



Signature of Author:
MIT Sloan School of Management
May 9, 2025

Certified by:
Jason Jay
Senior Lecturer, MIT Sloan School of Management
Director, MIT Sloan Sustainability Initiative
Thesis Supervisor

Accepted by:
Kelly Nixon
Senior Associate Director, MSMS Program
MIT Sloan School of Management

# Evaluating Impact Investing through a Systems Thinking Lens: Hallmarks of a Transformational Approach

By

**Yu (Sherry) Zhang**

Submitted to MIT Sloan School of Management
on May 9, 2025 in Partial Fulfillment of the
requirements for the Degree of Master of Science in
Management Studies.

**ABSTRACT**

As impact investing increasingly aspires to drive systemic change, the question of how to evaluate such efforts remains underexplored. Traditional evaluation approaches often grounded in linear causality and program-level outputs, and struggled to capture the complexity, interdependence, and emergent nature of systemic transformation. This thesis investigates how systemic investing can be evaluated by integrating systems thinking, evaluation theory, and investing practice. It develops a conceptual framework of thirteen hallmarks that characterize systemic investing evaluation across dimensions such as time horizons, stakeholder engagement, cross-sector collaboration, and capital dynamics. Drawing on 46 real-world cases, the research identifies 112 indicators to make these hallmarks observable and assessable in practice. To support practical application, the thesis also introduces an AI-assisted scoring tool that automates the evaluation of narrative content using the framework. Together, these contributions aim to support more reflective, adaptive, and system-aware evaluation practices in the emerging field of systemic investing.

**Thesis Supervisor:** Jason Jay
**Title:** Senior Lecturer, MIT Sloan School of Management
Director, MIT Sloan Sustainability Initiative

## Acknowledgements
First and foremost, I would like to express my deepest gratitude to my thesis advisor, Jason Jay, and my thesis reader, Emily Gates. Their support, encouragement, and guidance were invaluable throughout this journey. Jason has a remarkable ability to help me gather scattered thoughts, illuminate ideas stuck in a corner, and draw connections across seemingly unrelated threads. He helped me stay oriented when things felt uncertain and reminded me to focus on what truly matters. Working with Jason has been an experience far beyond writing a thesis—I have learned so much from how he thinks, mentors, and navigates complexity.

Emily opened the door to the field of evaluation for me. It was through reading her work that I first encountered this area, and her deep expertise helped me understand the field's richness and potential. Her thoughtful feedback and insights shaped many aspects of this thesis. From literature mapping to case analysis, from framework design to reviewing drafts, and even in our exchanges about life outside of research, I truly enjoyed the process of working under the guidance of Jason and Emily.

I would also like to thank the Owning Impact Project team—Hibah, Alessandro, Idongesit, Alban, and Jonathan. Under Jason's leadership, our weekly meetings were always thought-provoking and supportive. I'm also grateful to the MIT Sustainability Initiative. Although my time at MIT was brief, the courses and events offered by the Initiative gave me many ideas and connections. My sincere thanks go to the TransCap Initiative, the FEST, and the TWIST communities. This thesis drew heavily on case studies from TWIST and FEST—these communities generously shared their experiences and reflections, which became a key part of my research.

I want to thank my daughter, Yuxi, who was born during the process of writing this thesis. She gave me a new identity—not just as a student, but as a mother. Many late nights I typed with one hand while holding her in the other. You brought me strength I didn't know I had, and I hope one day you'll pursue whatever calls to you with your own kind of courage.

To my husband, Xing—thank you for being there through it all. From the early stages of my pregnancy to the emotional ups and downs, you stood by me with patience and unwavering support. You've always encouraged me to follow my heart, and that means the world to me. Finally, to my parents—thank you for flying across oceans to help care for me and our newborn. Your delicious food, quiet presence, and unconditional love have been my steady ground. I carry your strength with me in everything I do.

The decision to study this topic came quite naturally—from my past experiences in finance, to my long-standing interest in impact measurement, and the influence of systems thinking I encountered at MIT. Yet the research itself was far from easy. This is still an emerging field, filled with uncertainty and evolving practice. What I offer here is a first small attempt—a synthesis of many explorations.

# List of Content

# List of Figures

# List of Tables

# 1. Introduction

Over the past two decades, impact investing has gained traction as a strategy to direct private capital toward social and environmental good alongside financial returns. First coined by the Rockefeller Foundation in 2007 and championed by networks like the Global Impact Investing Network (GIIN), impact investing is defined as "investments made with the intention to generate positive, measurable social or environmental impact alongside a financial return". This approach embraced by a wide range of investors from pension funds and family offices to development agencies and foundations. Proponents view impact investing as a way to tackle "pressing social and environmental challenges" by harnessing market forces (Financier Worldwide, 2016).

Yet, even as impact investing flourished, it became clear that many of the world's most complex challenges, such as climate change, entrenched poverty, inequitable health, COVID and education outcomes, were not yielding to business-as-usual solutions. Early impact investments often targeted discrete projects or enterprises (for example, a single social business or a new technology) in hopes of outsized change. However, complex social and ecological problems are deeply systemic in nature and a single intervention can seldom unravel. Traditional investments that address only symptoms can even produce unintended consequences that spawn new problems (Daggers, Hannant, & Jay, 2023). In response to the limitations of isolated, piecemeal efforts, a growing concept in the impact sector calls for a more systemic approach to change. This gave rise to the concept of *systemic investing*, an emerging approach that explicitly applies systems thinking to impact finance (Daggers, Hannant, & Jay, 2023). Systemic investing seeks to go beyond one-off projects and instead transform the underlying structures and dynamics.

While the ambition of *investing for systems change* is high, it exposes a critical gap in evaluation practice. Traditional impact measurement frameworks typically rely on linear logic models and predefined programmatic outputs. They encourage investors to formulate a *theory of change* with specific, short-to-medium term outcomes and to track indicators that can be directly attributed to the intervention. Such frameworks work reasonably well for bounded projects instead of initiatives aiming for complex systems change, which unfold over long timeframes and involve multiple stakeholders. As Kania and Kramer (2011) observe, social solutions that "rely on a carefully worked out theory of change" and an evaluation attributing impact to a single program "rarely work under conditions of complexity," where unpredictable interactions of many players determine outcomes. In complex systems, cause and effect are not proportional and often not immediately observable, small actions can trigger big changes (Meadows, 2008). Patton (2011) argues that criteria like clear attribution, replicability, and upfront certainty of outcomes are "not appropriate or even meaningful" for "highly volatile environments, systems-change oriented interventions, and emergent social innovation".

Drawing on insights from systems thinking, complexity science, and evaluation theory, this thesis investigates how we might better understand, assess, and enhance the systemic investing evaluation. The research is guided by three key questions:

1. Why shift away from traditional evaluation approaches for systemic investing?
2. What does systemic investing evaluation look like and what are its characteristics?
3. How can practitioners begin to adopt and implement systemic investing evaluation in practice?

To address these, the thesis brings together theory and practice. A literature review situates this work within three intersecting fields: systems thinking, evaluation, and impact investing. Building on this, the research proposes a framework of 13 hallmarks, which are key characteristics observed in systemic investing evaluation. These hallmarks are derived from both literature and real-world cases. Then, drawing on 46 cases from two networks (TWIST and FEST), the study identifies 112 indicators that make the hallmarks observable and assessable. Finally, an AI-assisted scoring tool is developed to apply these indicators to actual projects, enabling rapid, context-aware evaluation.

In reviewing 46 cases of investments and initiatives aiming for systems-level change, several recurring patterns became evident: 1) Aspirations for transformational change. Funders and practitioners across these cases voiced ambitious goals to drive deep, systemic transformations (e.g. shifting a regional food system or reforming an education ecosystem), rather than just achieve incremental improvements. 2) Evaluation gaps. Despite these ambitions, few cases had a coherent evaluation framework aligned with system dynamics. Most relied on conventional metrics or ad-hoc storytelling, which did not capture feedback effects, changing network behaviors, or long-term capacity building. In other words, an evaluation gap persists – practitioners lack tools to assess progress toward transformation, making it hard to learn what's working in complex environments. 3) Common success factors. Notwithstanding the diversity of sectors and contexts, certain themes repeatedly mentioned in successful systemic investing efforts. Notably, many initiatives employed adaptive learning cycles, regularly reviewing data and context to iteratively adjust strategies. Many leveraged capitals in a catalytic way, for example, using grants or patient investment to attract additional resources or influence broader financial flows. And almost all stressed multi-stakeholder engagement, collaborating across nonprofits, governments, communities, and businesses to align goals and actions.

These field insights underscore a paradox in the current landscape: while the desire for transformational impact is high, the means to evaluate and manage such impact lag behind. This gap provides both a motivation and an opportunity for developing better frameworks that mirror the realities of systems change work.

# 2. Research Design and Methodology

## 2.1 Overall methodology of the thesis

This thesis adopts an approach that integrates theory and practice to explore how systemic investing can be meaningfully evaluated.

*Table 1: Research Questions and its related methodologies*

| Research Questions | Methodologies |
|---|---|
| (1) Why shift away from traditional evaluation approaches for systemic investing? | Literature review and the System Change Framework |
| (2) What does systemic investing evaluation look like and what are its characteristics? | 13 Hallmarks Framework |
| (3) How can practitioners begin to adopt and implement systemic investing evaluation in practice? | Indicator System and AI Tool |

The research is guided by three key questions. To address the first question, I conducted a literature review and drew on the System Change Framework (Kania et al., 2018), which helped articulate the limitations of conventional evaluation methods in the context of complex systems and informed the foundation for our evaluative lens. In response to the second question, I developed a framework of 13 Hallmarks that synthesize theoretical insights with observed features from systemic investment practices. These Hallmarks serve as a way to characterize what systemic investing evaluation looks like, which combined with the analysis of a series of real-world cases to test and refine these Hallmarks. For the third question, I derived a set of proposed indicators and a potential AI Tool.

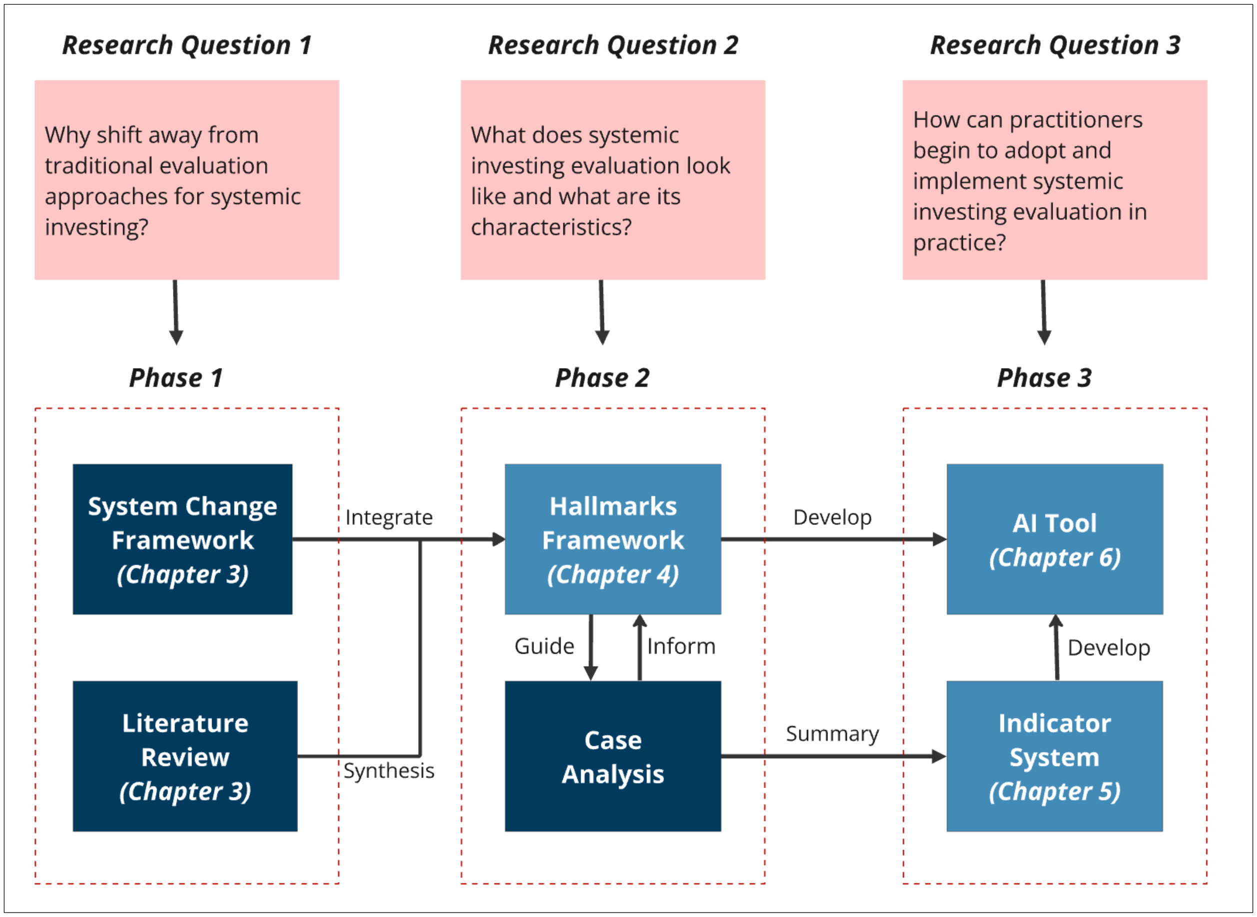


*Figure 1: Overall methodology/roadmap of the thesis*

## 2.2 Methodology for Literature Review

To identify emerging trends in systemic evaluation, this study adopted a synthesis-based approach to reviewing the literature. The review identifies recurring conceptual shifts; these shifts were then organized into thematic trends that reflect changes in how evaluation is understood. Figure 2 illustrates this synthesis process using one example. Drawing on Hargreaves and Podems (2012), it shows how the literature moves from advocating linear logic models toward approaches better suited for complex situations and interventions. The figure indicates how detailed textual descriptions from the literature can be abstracted into broader evaluative trends.

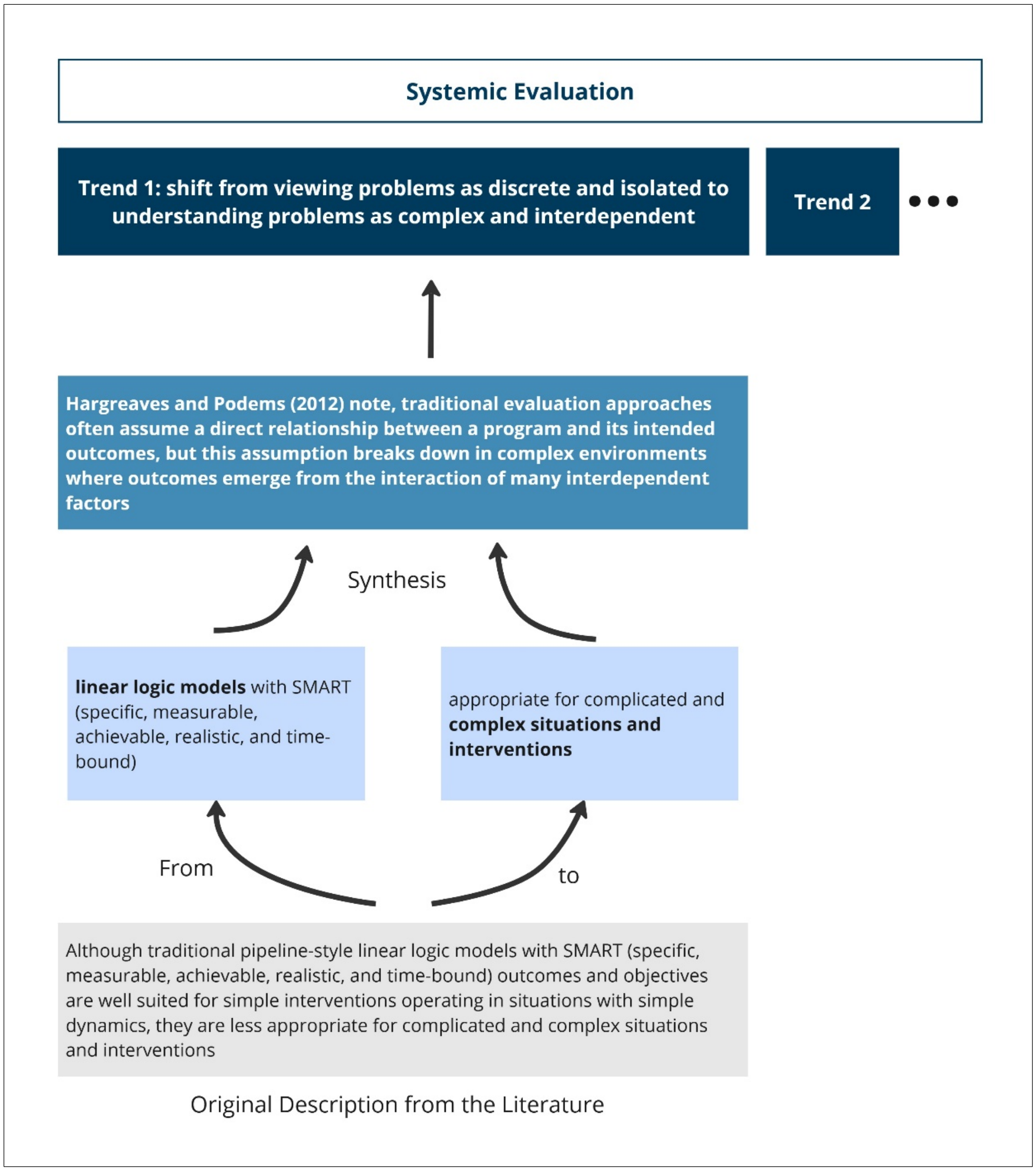


*Figure 2: Methodology for Literature Summary*

## 2.3 Methodology for Hallmarks Framework

The development of the Hallmarks Framework followed a multi-step method with both theoretical and empirical validation. Initially, the research drew upon two key inputs: a literature review on systems thinking, evaluation, and capital deployment, and the System Change Framework proposed by Kania et al. (2018), which outlines six interconnected conditions for systems change. These two inputs were synthesized and integrated to construct a preliminary version of the Hallmarks Framework. Case materials, including interviews, presentations, and project documentation, were analyzed to identify patterns, extract another relevant 2 Hallmarks (Multicapital Circulation and Capital Catalyzation).

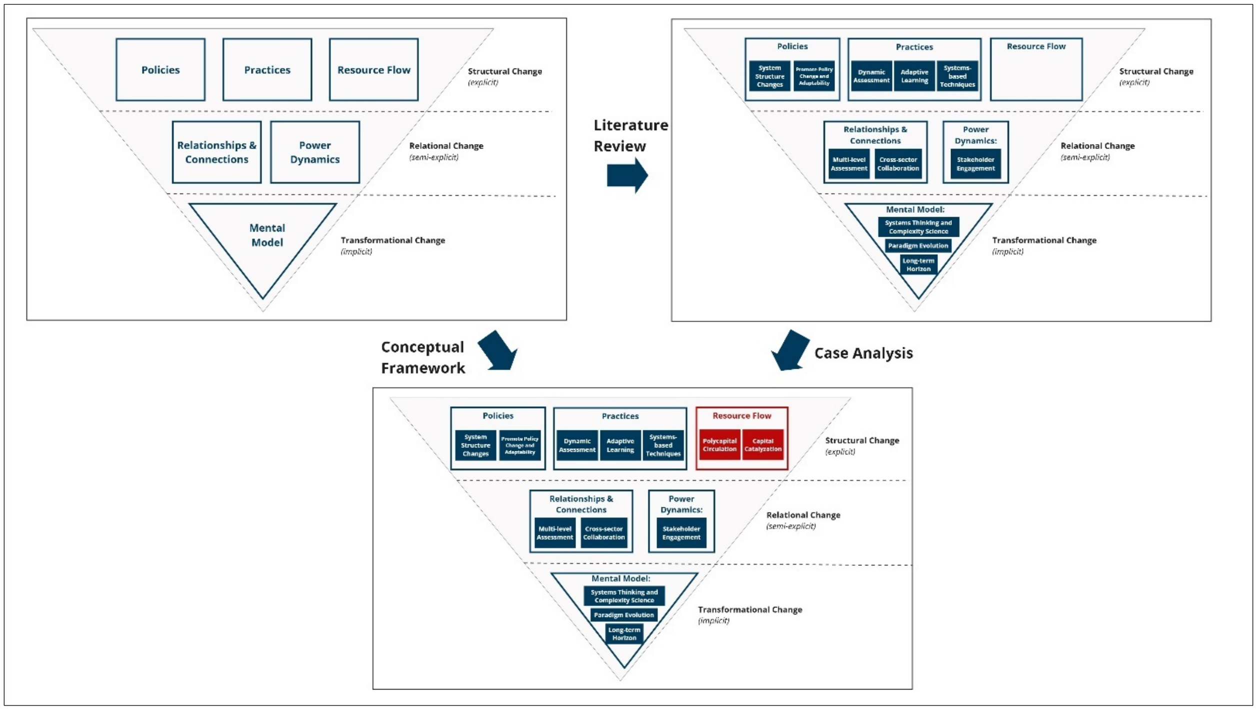


*Figure 3: Methodology for Hallmarks Framework*

## 2.4 Methodology for Case Analysis

Since systemic investing evaluation is a relatively new field without established measurement indicator frameworks, this research employs multiple case study designs to analyze possible measurement approaches. Learning from real practices described by interviewees, the approach aims to identify potential indicators that can be extracted from their descriptions (Figure 5). Through the analysis of multiple cases, patterns can be explored across different contexts and develop a potential set of measurement framework.

### *2.4.1 The Case Sampling*

In this thesis, TWIST (Together We Invest in Systems Transformation) and FEST (Financial Ecosystems for Systemic Transformation) are utilized as the source of cases. Of the 46 cases, 44 came from TWIST and 3 came from FEST.

TWIST is an international practitioner collective composed of asset owners, asset managers, advisors, and field-builders committed to advancing systemic investing practices. The network serves as a practical sampling pool for identifying initiatives that aim to foster systemic change through capital deployment. The sampled cases are self-identified as investors for system change. The Finance for Ecosystem and Systemic Transformation (FEST) initiative represents another community of practitioners, funders, and thought leaders working to accelerate this field's development. FEST frames finance not only as a resource allocation mechanism, but as a system shaped by underlying values, mental models, and institutional logics that must be reconfigured for transformative impact. 3 prototypes are being analyzed by 9 working groups, such as assessment, mapping, enabling policy environments, and narrative framing.

According to the TWIST White Paper (2024) and FEST Survey, the 46 Cases includes a diverse range of capital providers, with a breakdown indicating that approximately 32.9% are family offices, 22.4% are Enabling Platforms, 11.8% are foundations, 10.5% are consultancies, 7.9% are asset managers, 6.6% are Network Leaders, others including Coalition leaders (2.6%), Academic (2.6%) and Financial Institutes (2.6%).

The case studies also represent a broad geographic coverage, reflecting the network's global perspective on systemic investing: 29% of the cases address global or multi-regional issues, while 24% are situated in Europe and 19% in North America. Cases from emerging markets are also represented, with 10% of studies focused on Africa, 10% on South America, and 10% on Asia. Australia accounts for 2% of the cases.

In terms of domains of system change, the most commonly targeted systems are the economic system (28.6%), food system (21.4%), and financial system (14.3%). Social inequality is also a significant focus, accounting for 11.9% of the cases. Other systems such as oceans (4.8%), built environment (2.4%), energy (2.4%), health (2.4%), education (2.4%), electronics (2.4%), political governance (2.4%), mobility (2.4%), and media (2.4%) are represented to a lesser extent.

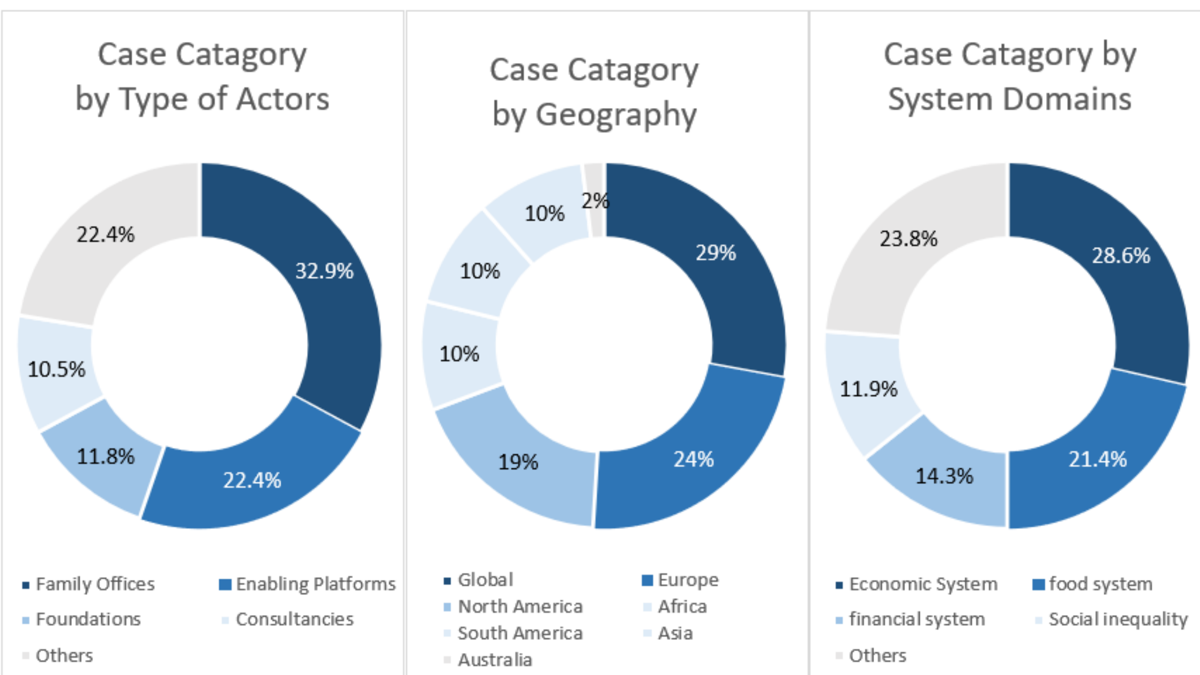


*Figure 4: Summary of Case characteristics*

TWIST and FEST cases offer insight into emerging practices among pioneering actors, however this composition suggests a concentration of actors operating within mission-driven and concessionary capital markets, a concentration of presence in Europe and North America, and a concentration of market-based and resource-related domains. The concentrations may introduce certain biases; therefore TWIST and FEST case provides a relevant but bounded source of systemic investment initiatives.

*Table 2: Overview of Case Studies*

| No. | Company/Project Name | Source |
|---|---|---|
| 1 | Tara Health Foundation | TWIST |
| 2 | Regenerative Alpine Region | TWIST |
| 3 | Library of Things | TWIST |
| 4 | Apolitical | TWIST |
| 5 | Osa Regenerative Rainforest Economy Lab | TWIST |
| 6 | KOKO Networks | TWIST |
| 7 | VisVires New Protein Fund II (VVNP II) | TWIST |
| 8 | Venture Ecosystem Building | TWIST |
| 9 | Fair By Design Fund | TWIST |
| 10 | The Fink Family and ReFED | TWIST |
| 11 | Women's World Banking Fund II | TWIST |
| 12 | FAES | TWIST |
| 13 | 1000 Ocean Startups Coalition | TWIST |
| 14 | Small Giants | TWIST |
| 15 | Regen Melbourne | TWIST |
| 16 | Connovo Venture Builder | TWIST |
| 17 | Transform Finance | TWIST |
| 18 | Purpose | TWIST |
| 19 | Tidal Impact | TWIST |
| 20 | Rainfall | TWIST |
| 21 | Hawaii Investment Ready | TWIST |
| 22 | Break Free From Plastic | TWIST |
| 23 | TransCap Initiative Net-Zero Mobility | TWIST |
| 24 | Coalition for Impact (C4i) | TWIST |
| 25 | Fresh Ventures | TWIST |
| 26 | New Capitalism Project | TWIST |
| 27 | Presencing Institute | TWIST |
| 28 | MyLand | TWIST |
| 29 | Aqua Spark | TWIST |
| 30 | Cinereach | TWIST |
| 31 | Metabolic Ecosystem | TWIST |
| 32 | Katapult | TWIST |
| 33 | Criterion Institute | TWIST |

| | | |
|---|---|---|
| 34 | Deep Transitions Lab | TWIST |
| 35 | Pioneers of Our Time | TWIST |
| 36 | Green Spark | TWIST |
| 37 | Dark Matter Labs | TWIST |
| 38 | Small Foundation | TWIST |
| 39 | Prime Coalition | TWIST |
| 40 | Finance Innovation Lab | TWIST |
| 41 | Social Innovation and Investment Foundation (SIIF) | TWIST |
| 42 | Meridian Institute | TWIST |
| 43 | Gatsby Africa | TWIST |
| 44 | Messina Foundation | FEST |
| 45 | Makaia | FEST |
| 46 | Kwaxala | FEST |

### *2.4.2 The Data Collection*

An important aspect of both communities is their openness to sharing works-in-progress and early-stage practices, and its commitment to group learning, which allowed us access to information not typically available elsewhere.

Data collection for the TWIST and FEST cases analysis was conducted through a combination of surveys, presentations, and discussions. Prior to each session, working group organizers distributed a template for participants to complete, which is used as a preliminary input for the study. During each sharing session, a representative from a participating organization presented their work, followed by an open discussion, in total lasting approximately one to one and a half hours. These sessions were recorded. The surveys, sharing session, and any supplementary materials provided by the presenters were subsequently synthesized to inform the case study analysis through a thematic coding process (Figure 5). Table 2 is a list of 46 Cases analyzed in this thesis.

### *2.4.3 The Data Analysis*

Figure 5 illustrates how to derive indicators through a case-based approach. I started by collecting original descriptions from interviews, then synthesized wordings from each case. Based on these syntheses, preliminary indicators are proposed to reflect, for example, how capital catalyzation was described in practice. I then grouped similar indicators under a common measurement concept, in this example, the capital leverage ratio. Finally, each indicator was classified under a broader theme or objective and linked to the corresponding hallmark (in this case, Capital Catalyzation). This method allowed us to connect real-world case and evidence to potential systemic evaluation indicators.

During a first phase of this analysis, I conducted by hand this extraction of the practice description, synthesis into a concise description, inference of an indicator, and summary of a theme. I then used the Cursor software package to generate Python code to help automate this analysis using ChatGPT 4.0mini through an API. As a result, I was able to collect a total of 616 data points from 46 cases, and summarize 112 indicators, categorized by 13 Hallmarks.

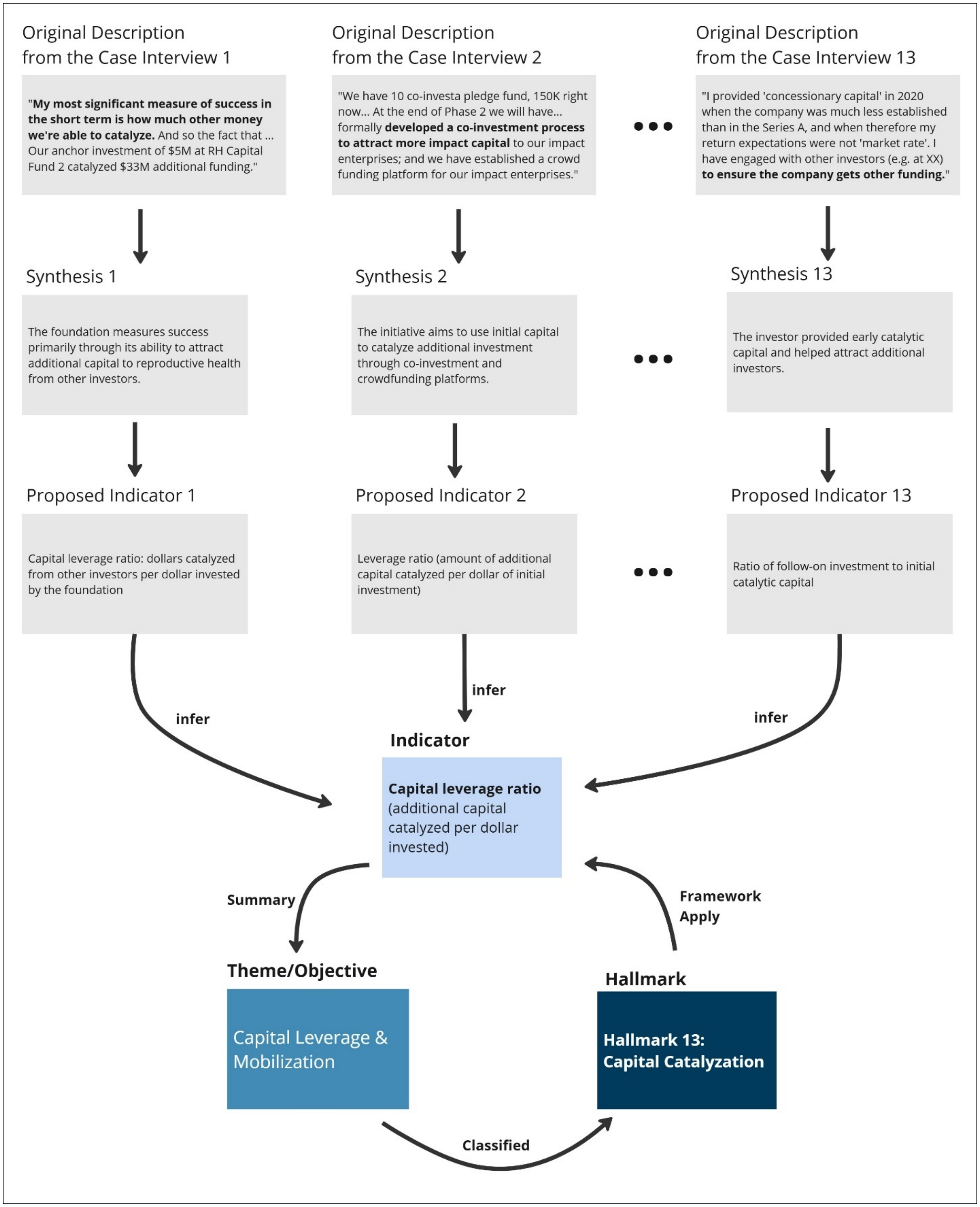


*Figure 5: Case Summary Methodology*

## 2.5 Limitations of Methodology

There are several limitations to the methodological approach adopted in this research. First, much of the case analysis relies on interview data, although this way rich in context and practitioner insight, while mainly shaped by the perspectives and framing of the interviewees. This introduces the possibility of “recall” bias, or aspirational descriptions that may not fully reflect actual outcomes or practices or possibilities. Second, the process of synthesizing qualitative data into indicators involved interpretive judgment. While efforts were made to increase the transparency in how indicators were derived from the data, the step from description to indicators is subjective. It reflects the researcher’s understanding of what counts as relevant evidence and how it aligns with the broader evaluation framework. Third, the construction of the 13 Hallmarks Framework, though based on literature and field cases, also carries some degree of normative influence, particularly in selecting which characteristics to emphasize as “hallmarks” of systemic evaluation. These limitations do not undermine the value of the framework, but they do point to the importance of viewing the results as just a starting point for this comprehensive topic.

# 3. Rationale for a Systemic Approach to Investing Evaluation (Responding to RQ1)

This chapter responds to the first research question:
**RQ1: Why shift away from traditional evaluation approaches for systemic investing?**

To address this question, the chapter examines the limitations of conventional evaluation methods in capturing the dynamic, interconnected, and adaptive nature of systemic change efforts. Traditional approaches often emphasize linear causality, short-term outputs, and isolated interventions. However, systemic investing initiatives seek to address root causes, enable feedback loops, and influence broader systems through distributed and long-term interventions. These characteristics pose challenges for evaluation frameworks over transformation.

Building on literature from systems thinking, complexity science, and the impact investing field, this chapter develops the rationale for a systemic evaluation lens. It discusses why existing tools may be insufficient and articulates the conceptual foundations necessary for evaluating investments that aim not only to do good but to shift the underlying conditions of complex systems. In doing so, it lays the groundwork for the next two chapters, which explore what systemic investing evaluation should look like (Chapter 4) and how it can be implemented in practice (Chapter 5).

## 3.1 Research Scope and Framework

The scope of this research project lies at the intersection of three domains: systemic thinking and complexity science (STCS); evaluation; and capital deployment through impact investing. It explores how systemic-thinking-based evaluation can be effectively applied in the domain of capital deployment.

While each of the three domains has been extensively studied, the intersection of all three remains relatively underexplored. To address this gap, this research conducts a literature review based on the bilateral intersections between these fields. Specifically, it draws upon the following conceptual groupings:

(1) Systems Thinking and Impact Investing, hereby referred to as *Systemic Investing* (SI). This intersection examines how capital can be deployed to influence whole systems rather than isolated issues, recognizing that financial flows are embedded within and can reshape broader social and environmental systems.

(2) Systems Thinking and Evaluation, hereby referred to as *Systemic Evaluation* (SE). This integration explores how evaluation methods can capture system dynamics that occur in complex interventions.

(3) Impact Investing and Evaluation, often termed *Impact Measurement and Management* (IMM). This field focuses on assessing the social and environmental outcomes of investments while informing ongoing decision-making and strategy refinement.

These combinations serve as foundational lenses to better understand the emerging integrated field of **Systemic Investing Evaluation**, where all three areas converge. By integrating these perspectives, we can develop frameworks for tracking how financial capital influences system dynamics and contributes to lasting, transformative change. The relationships among these domains and their intersections are illustrated in Figure 6.

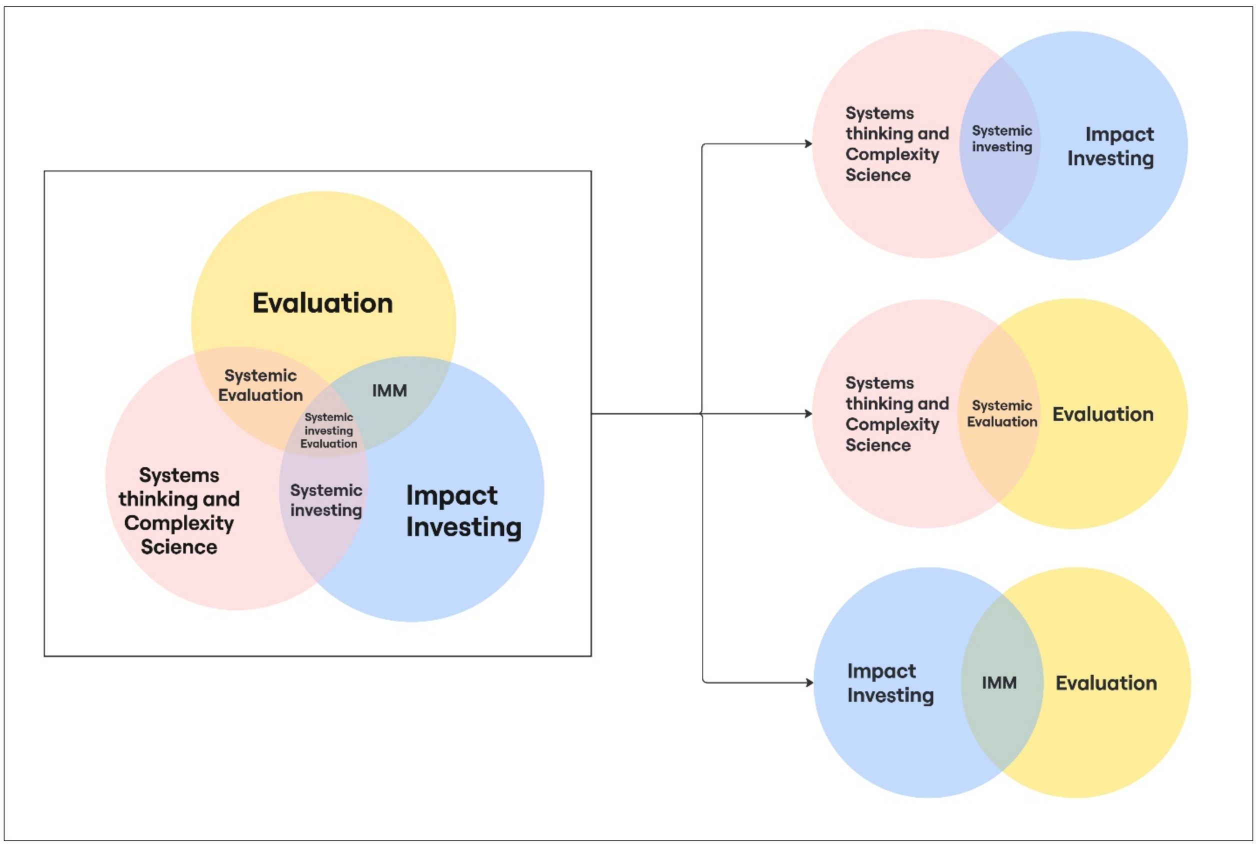


*Figure 6: Research Scope: Intersections of Three Domains*

## 3.2 Systemic Evaluation (SE)

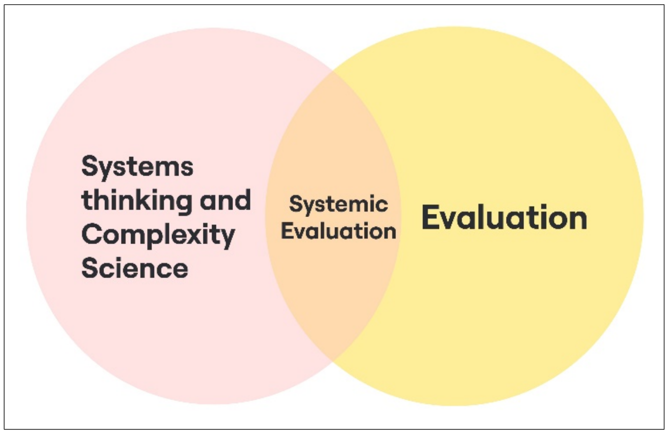


*Figure 7: Intersection 1- Systemic Evaluation*

Systemic Evaluation has emerged as a response to the limitations of conventional, linear approaches to evaluation when addressing complex, multi-stakeholder, and interdependent social challenges. Drawing from systems thinking and complexity science (STCS), Systemic Evaluation reflects a shift in systems-level understanding, adaptive learning, and transformational intent.

The introduction of systems thinking and complexity science into the field of evaluation can be traced back several decades, originating from disciplines such as general systems theory, cybernetics, and complexity science (Nicolis & Prigogine, 1989; Kauffman, 1995; Guba & Lincoln, 1989). These foundational theories challenged linear and reductionist approaches, laying the groundwork for evaluation frameworks that could better address the dynamic, emergent, and multi-layered nature of social change. Over time, scholars such as Patton (2011), Midgley (2000), and Byrne and Callaghan (2014) further articulated how complexity concepts, such as non-linearity, emergence, adaptation, feedback loops, and attractor states, can be incorporated into evaluation theory and practice. This intellectual evolution has led to the development of approaches like Developmental Evaluation, Realist Evaluation, and systemic theories of change that support learning and adaptation in dynamic systems (Walton, 2014; Hargreaves & Podems, 2012).

Eleven trends are evident in the development of Systemic Evaluation theories and approaches (refer to Table 3 for the summary).

**First, there is a growing shift in evaluation practice from viewing problems as discrete and isolated to understanding them as complex and interdependent.** As Hargreaves and Podems (2012) note, traditional evaluation approaches often assume a direct relationship between a program and its intended outcomes, but this assumption breaks down in complex environments where outcomes emerge from the interaction of many interdependent factors. Walton (2016) similarly observes that interventions increasingly operate within open, adaptive systems where causal chains are nonlinear, boundaries are fluid, and effects unfold over multiple scales. In such contexts, efforts to isolate and evaluate a single program risk can miss the broader system dynamics that shape results. Beer (2017) further highlights how funders and evaluators struggle to reconcile the demands for short-term, measurable results with the reality that systemic change is slow, unpredictable, and often influenced by forces outside any single actor's control. These insights reflect a broader shift in evaluation practice: as the problems themselves have become more complex, approaches shall be adjusted to understanding and responding to them. Several scholars underscore the practical value of systems thinking in evaluation. Patton (2016) explicitly defines complexity as a guiding principle for both innovation and evaluation, suggesting that evaluators must navigate uncertainty, adaptation, and emergence throughout the process. Similarly, Walton (2016) presents complexity as a property not only of interventions but of the social systems they engage with, stressing the importance of mapping relationships, identifying feedback structures, and being attentive to unintended consequences. Cabaj (2019) builds on this by defining systems as composed of interrelated elements whose purpose emerges through their interactions.

**Second, there is a shift from determining fixed value based on predetermined criteria to developing evolving value**. Schwandt and Gates (2021) argues that we should shift from determining value to developing value, and systemic evaluation should move beyond traditional models that separate means and ends. Donella Meadows (1999) identified paradigms as deeply held beliefs about how systems work, as among the most powerful leverage points for transformation. Systemic evaluation, by this logic, becomes not just a technical tool but a way of surfacing and potentially reshaping those underlying assumptions. Hargreaves and Podems (2012) add that traditional evaluation model are based on SMART goals and linear logic, often fall short in contexts marked by ambiguity and emergence. In these cases, a more flexible, principles-based evaluation becomes necessary, one that can support emergent learning and iterative decision-making (Patton, 2016; Gates, 2021). A trend in the field is the recognition that evaluating transformation requires transforming evaluation itself (Patton, 2020). This idea is exemplified by Blue Marble Evaluation (Patton, 2020), which focuses on global systems transformations and posits that traditional evaluation approaches may be insufficient for addressing complex, interconnected global challenges. These developments are reflected in global evaluation events and networks, from the Australasian Evaluation Society 2018 and IDEAS 2019 conference on "transformative change" to the "Visionary Evaluation" by American Evaluation Association in 2020.

**Third, there is a shift from short-term metrics and immediate outcomes to long-term timeframes and extended impact pathways.** Carney (2015) referred to climate change as "the tragedy of the horizon," highlighting how long-term risks are often invisible to short-term actors. Gates (2021) and Balakrishnan et al. (2024) point out that focusing only on near-term outputs can obscure the trajectory of deeper, more structural changes. Many interventions require persistent efforts and slow-building shifts that are difficult to detect in early years.

Ivashina and Lerner (2019) advocate for patient capital as a necessary condition for addressing the world's most pressing systemic problems.

**Fourth, there is a shift in evaluative focus from assessing individual programs to examining pathways toward system change.** Rather than evaluating the outcomes of individual programs designed to modify parts of systems, systemic evaluation strives to understand how interventions might change the underlying structures, behaviors, and patterns within systems or even design entirely new systems (Beer, 2017). This perspective recognizes that meaningful change often requires transformation at the system level rather than merely adjusting components within an existing framework. The ReThink Health initiative, as described by Gates and Fils-Aime (2021), exemplifies this approach through tools such as the Stewards' Pathway and Well-being Portfolio, which help regional stewards visualize and navigate system dynamics. These tools support a shift from evaluating isolated programs to understanding how various interventions contribute to broader system transformation. Gates and Fils-Aime (2021) highlight a methodological shift from attribution to contribution in systemic evaluation. Rather than attempting to isolate the effects of specific interventions, systemic evaluators recognize that change emerges from the interaction of multiple factors within complex systems. This shifts the focus from measuring readily quantifiable outputs to tracking more diffuse and long-term changes in system conditions, such as shifts in mental models or power dynamics. Davies and Goldie (2023) further elaborate on this shift, arguing for a more plausible, contribution-based understanding of causality that acknowledges the slow pace of change, the emergent nature of outcomes, the opacity of causal pathways, the evolving nature of interventions, and the multiplicity of definitions of both systems and success in complex environments.

**Fifth, there is a shift from single-sector approaches and siloed evaluation to multi-stakeholder, boundary-spanning collaborations.** Traditional evaluation examines initiatives within sectoral silos, whereas systemic evaluation recognizes that complex problems require coordinated efforts across government, business, nonprofit, and community sectors, each with distinct resources, constraints, and capabilities. Systemic problems, like climate or inequality, span sectors and require coordinated action. Walton (2016) frames this as network governance, where public, private, and civic actors co-create solutions. Trujillo (2018) notes that such cross-sectoral efforts are now common strategies for tackling "wicked" problems. In systemic investing, collaboration enables impact at scale. Daggers et al. (2023) advocate for ecosystem investing, supporting actors across sectors to collectively shift conditions.

**Sixth, there is a shift from treating stakeholders as data sources or passive recipients to engaging them as co-evaluators and knowledge co-creators.** Gates (2021) highlights the use of dialogic methods, where stakeholders actively help shape the evaluation focus, interpretation, and learning processes. This reflects a broader shift toward participatory evaluation traditions. Walton (2016) also connects stakeholder engagement to system governance. He observes that networked governance models create better conditions for adaptation and learning. In practice, stakeholder engagement is often difficult. However, systemic evaluators increasingly recognize that addressing these problems is itself a form of system change (Hargreaves & Podems, 2012; Gates, 2021). In systemic investing, engaging stakeholders means listening to communities, frontline workers, and integrating their views into strategy and learning cycles.

**Seventh, there is a shift from measuring programmatic outputs and direct outcomes to assessing modifications in underlying system structures.** Cabaj (2019) proposes evaluating system change by looking for shifts in key system drivers, actor behaviors, and broader patterns of system functioning. This moves evaluation beyond program-level results toward assessing structural influence. Similarly, Meadows (1999) describes system structure, such as rules, feedback loops, and flows of information, as deeper levers of change compared to surface-level outcomes. Evaluators have developed practical ways to observe structural change. For example, Walton (2016) discusses how changes in governance models, feedback loops, and decision-making patterns can signal structural shifts.

**Eighth, there is a shift from rigid, compliance-oriented funding frameworks to adaptive, co-created policy environments that enable systemic change.** Systemic change efforts often push for policy reform to change formal rules and build systems that can adapt over time. Patton (2016) emphasizes that developmental evaluation must provide timely, actionable feedback to inform ongoing policy shifts. Rather than only evaluating outcomes at the end, evaluators need to embed themselves in the policy cycle, enabling continuous adaptation. Cabaj (2019) links adaptive learning to policy change, arguing that when actors use evidence to revise their strategies and policies, it is a sign of systemic responsiveness. In practice, evaluating policy adaptability means looking at whether institutions can course-correct, respond to feedback, and integrate learning into their rules or practices. It also means recognizing that not all change comes from official policy.

**Nineth, there is a shift from treating evaluation as a bounded activity to embracing it as a continuous, embedded process.** Walton, Gates, and Vidueira (2021) advocate for a shift away from evaluation as a bounded, summative activity toward a continuous, embedded process that supports systemic change. Systemic approach means departure from fixed intervention approaches and instead embraces adaptation (Beer, 2017; Patton, 2011). This involves using evaluation to surface and reshape mental models, redistribute power, and inform adaptive learning cycles. In this view, systemic evaluation becomes a form of system stewardship, contributing to both intentional system redesign and institutional learning. Davies and Goldie (2023) reinforce this perspective with their concept of "evaluation as the engine of system change," which emphasizes embedding evaluative thinking into strategic planning and balancing a directional "North Star" vision with continuous iteration. This aligns with what Picciotto (2020) describes as "triple-loop learning," which moves beyond adjusting strategies (single loop) or goals (double loop) to interrogating the underlying values and power structures that define what is considered valuable in the first place.

**Tenth, there is a shift from single-loop learning focused on program improvement to double and triple-loop learning that questions fundamental assumptions.** While dynamic assessment focuses on timely information flow, adaptive learning refers to how that information is used. Gates (2021) argues that evaluation should support double- and triple-loop learning within organizations, helping actors reassess not just actions, but goals and mental models. Hargreaves and Podems (2012) describe practical strategies for increasing adaptability, such as using early warning indicators, flexible data collection methods, and decoupled evaluation components. Walton (2016) also emphasizes adaptive governance structures, particularly networked approaches that allow for distributed sense-making and iterative decision-making. Evaluators working within such structures can help participants not only learn from data but also act on that learning in a timely and collective way. Adaptive learning is especially crucial in policy and investment contexts, so that we can capture real-time shifts in risk, opportunity, or impact.

**Eleventh, there is a shift from standardized evaluation tools and linear methodologies to context-specific, systems-oriented methodologies.** Traditional evaluation relies heavily on logic models, experimental designs, and quantitative metrics, whereas systemic evaluation draws on diverse techniques to capture complex dynamics. Bustamante et al. (2021) provide examples of systems mapping, network analysis, and scenario modeling used in food system evaluations. These tools help evaluators identify leverage points, visualize system dynamics, and anticipate unintended effects. Walton (2016) and Patton (2016) further highlight that the choice of tools matters and using complexity-aware methods are suitable for the complexity of the context.

## 3.3 Systemic Investing (SI)

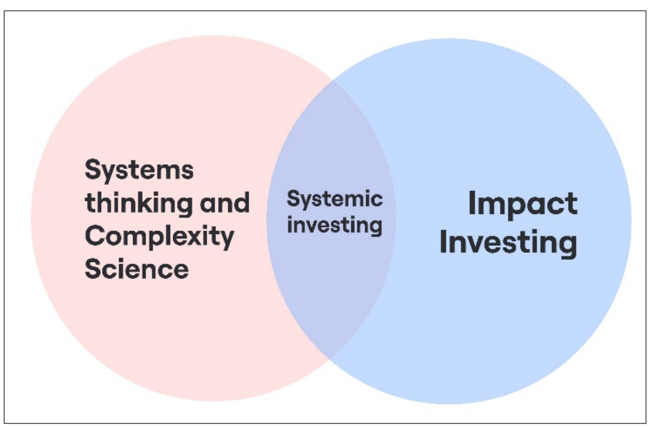


*Figure 8: Intersection 2 - Systemic Investing*

The concept of systemic investing has emerged alongside a broader shift in the financial sector which described as the “systems turn in finance” by Jay (2024). The TransCap Initiative (2024) defines systemic investing as "the deployment of financial capital to transform human and natural systems with the intention of advancing environmental sustainability and social justice". This shift encompasses two orientations (Jay, 2024). One focuses on managing systemic risks that affect the stability of markets and long-term asset values, while the other seeks to contribute to systemic change by influencing the underlying structures of environmental and social systems.

The first orientation is represented by a group of actors including the Principles for Responsible Investment (PRI), the Investment Integration Project (TIIP), and the Shareholder Commons, etc. Often associated with large fiduciary investors, these organizations have advanced the notion of "system-level investing" or "ESG 2.0" (TIIP, 2024). This approach aims to address systemic risks, such as climate change, income inequality, and biodiversity loss, by incorporating them into investment decisions at the portfolio level. It reflects the view that long-term financial performance is increasingly shaped by system-wide externalities rather than the performance of individual firms.

A second orientation is evident among mission-driven capital providers such as Rockefeller Philanthropy Advisors, the Center for Strategic Philanthropy (CSP), and the initiative "Together We Invest in Systems Transformation" (TWIST). These organizations' work can be framed as "Investing for Systems Change" or "Impact Investing 2.0" (Jay, 2024). Their strategies tend to focus on enabling structural transformation within systems, such as in health, education, food, or energy, and often by supporting cross-sector collaboration, infrastructure development, or narrative shifts. The emphasis here is less on risk reduction and more on enabling coordinated interventions that support systemic outcomes over time.

Between these two orientations lies a group of initiatives that apply a multicapital perspective to systemic investing. These efforts include the TransCap Initiative, Deep Transitions and MIT Sustainability Initiative. A defining feature is the intentional use of multiple forms of capital, such as financial, natural, social, and political, mobilized in ways that support coordinated action among stakeholders toward shared objectives (Daggers et al., 2023). Rather than adhering to fixed categories, these initiatives emphasize flexibility in defining system boundaries. As noted in the TWIST white paper (2024), system boundaries are determined through negotiation among coalitions, reflecting both the nature of the challenge and the strategic priorities of participating investors. They may be spatial (e.g. a city), sectoral (e.g. food systems), or thematic (e.g. just transition), and are often redefined over time as understanding evolves (Hofstetter, 2024). Positioned outside the two primary orientations, the FEST network aims to be inclusive for both sides by bridging insights and actors across the spectrum, fostering shared learning to advance systemic investment practices.

One compelling case study demonstrating the possibilities of systemic investing is the work of Betsy and Jesse Fink in addressing food waste in the United States. Rather than simply funding individual companies with promising solutions, the Finks were combining philanthropy, direct investments, and foundation grants while building a network of stakeholders around this previously neglected issue (Yau & Jay, 2023). This approach led to the creation of ReFED, a backbone organization that provides data infrastructure and coordinates action across the food system.

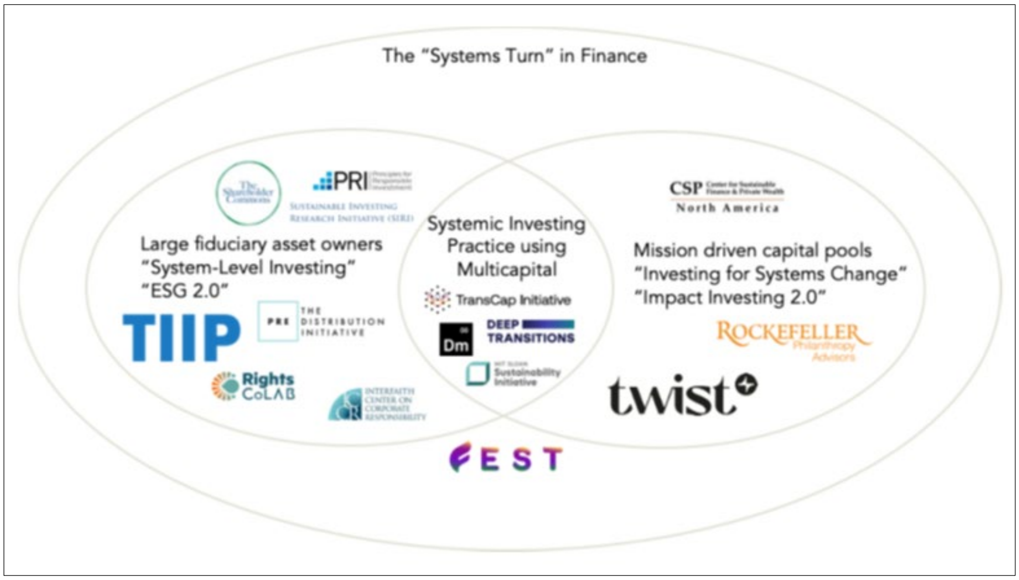


*Figure 9: The "Systems Turn" in Finance*

## 3.4 Impact Measurement and Management (IMM)

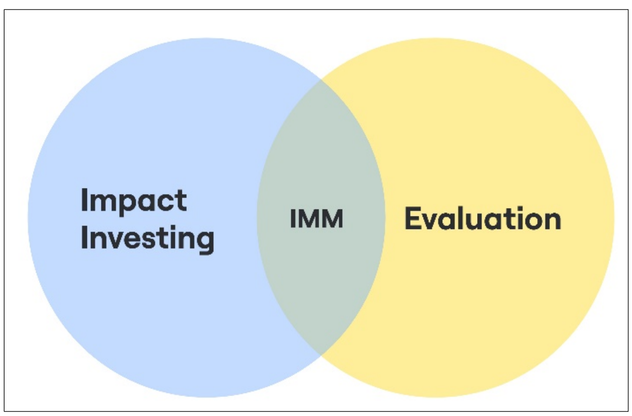


*Figure 10:  Intersection 3 - Systemic Investing*

Impact Measurement and Management (IMM) is important for organizations seeking to understand, demonstrate, and enhance their social and environmental contributions. The field has evolved significantly from traditional evaluation approaches to more sophisticated methodologies that integrate impact considerations throughout organizational decision-making processes. Early frameworks such as the Logic Model and Theory of Change laid the groundwork by establishing causal pathways between organizational activities and intended outcomes, emphasizing the importance of articulating how change occurs (Ebrahim & Rangan, 2014). These models have been complemented by the Social Return on Investment (SROI) methodology, which attempts to monetize social value creation by assigning financial proxies to non-financial outcomes, thus enabling cost-benefit comparisons across diverse interventions (Nicholls et al., 2012). Despite criticisms regarding the subjectivity of financial proxy selection, SROI has gained traction for its ability to translate complex impact concepts into the familiar language of financial returns.

The emergence of standardized metrics and reporting frameworks has further advanced the IMM field by enabling comparability across organizations and sectors. The Impact Reporting and Investment Standards (IRIS+) catalog, developed by the Global Impact Investing Network (GIIN), offers a comprehensive set of standardized metrics for assessing social, environmental, and financial performance across various impact dimensions (Bouri et al., 2018). Similarly, the United Nations' Sustainable Development Goals (SDGs) have provided a universal framework that organizations increasingly use to align their impact measurement practices with globally recognized development priorities. The SDGs' 169 targets and 232 indicators offer a common language for articulating impact, though organizations often struggle with selecting appropriate indicators and establishing meaningful connections between organizational activities and high-level SDG outcomes (Sachs et al., 2019). The B Impact Assessment, developed by B Lab, represents another significant standardization effort, providing a holistic evaluation tool that examines organizational impacts across governance, workers, community, environment, and customers dimensions (Stubbs, 2017).

Recent methodological innovations have sought to address limitations in traditional approaches by incorporating stakeholder perspectives more centrally and leveraging technological advances. Participatory impact assessment methods, drawing on principles from participatory action research, engage beneficiaries and other stakeholders in defining meaningful outcomes and designing appropriate measurement approaches (Chambers, 2017). These methods challenge power imbalances in conventional evaluation practices by privileging the voices of those most affected by interventions. Developmental evaluation offers another innovative approach, providing real-time feedback to inform adaptive management in complex, dynamic contexts where outcomes cannot be fully predetermined (Patton, 2011). Meanwhile, technological advances in data analytics, machine learning, and remote sensing are enabling more sophisticated impact measurement by processing larger datasets, identifying patterns, and reducing data collection costs. Blockchain applications have also emerged to enhance transparency and verification in impact reporting chains (Kshetri, 2018).

The integration of impact considerations into organizational management systems represents perhaps the most significant evolution in the field. Impact management frameworks move beyond post-hoc evaluation to embed impact thinking throughout strategic planning, operational decision-making, and governance structures. The Impact Management Project (IMP) has developed a widely adopted framework that examines impact across five dimensions: What, Who, How Much, Contribution, and Risk (Impact Management Project, 2018). This multidimensional approach helps organizations develop more nuanced understandings of their impacts while informing strategic decisions about impact maximization.

Despite significant methodological advances, the IMM field continues to grapple with several persistent challenges. The attribution problem remains particularly vexing, as establishing causal relationships between specific interventions and observed outcomes is inherently difficult in complex social and environmental systems (White, 2010). Experimental and quasi-experimental impact evaluation designs, including randomized controlled trials, difference-in-difference methods, and propensity score matching, offer rigorous approaches to attribution but face practical limitations including cost, ethical considerations, and limited generalizability across contexts. Furthermore, the field continues to struggle with balancing standardization needs against contextual relevance, as highly standardized metrics may sacrifice important nuances while wholly customized approaches undermine comparability (Ebrahim et al., 2014). The data quality challenge also persists, with many organizations lacking sufficient resources, expertise, or systems to collect reliable impact data consistently. Looking forward, emerging approaches such as systems thinking and complexity-aware evaluation methods hold promise for addressing these challenges by better capturing dynamic interactions, feedback loops, and emergent outcomes in complex systems (Hargreaves & Podems, 2012).

## 3.5 The "Water of System Change" Framework

This thesis adopts the "Water of Systems Change" framework developed by Kania, Kramer, and Senge (2018), which conceptualizes systems change as the process of shifting the underlying conditions that hold social problems in place. The framework identifies six interrelated conditions that operate across three levels of visibility: structural, relational, and transformative. At the structural level, change involves formal policies, practices, and resource flows, which are the explicit rules. The semi-explicit, relational layer includes relationships and connections as well as power dynamics, emphasizing how authority is distributed and how different actors interact. At the most implicit level lie mental models, which means the deeply held beliefs, assumptions, and narratives that influence how people make sense of the world. These six conditions interact dynamically: progress at the structural level is difficult to sustain without shifts in relationships and mental models, and changes in the latter are often needed to unlock structural reform. This framework provides the foundation for the analysis that follows, offering a way to understand not only what needs to change in a system, but also how and where to intervene.

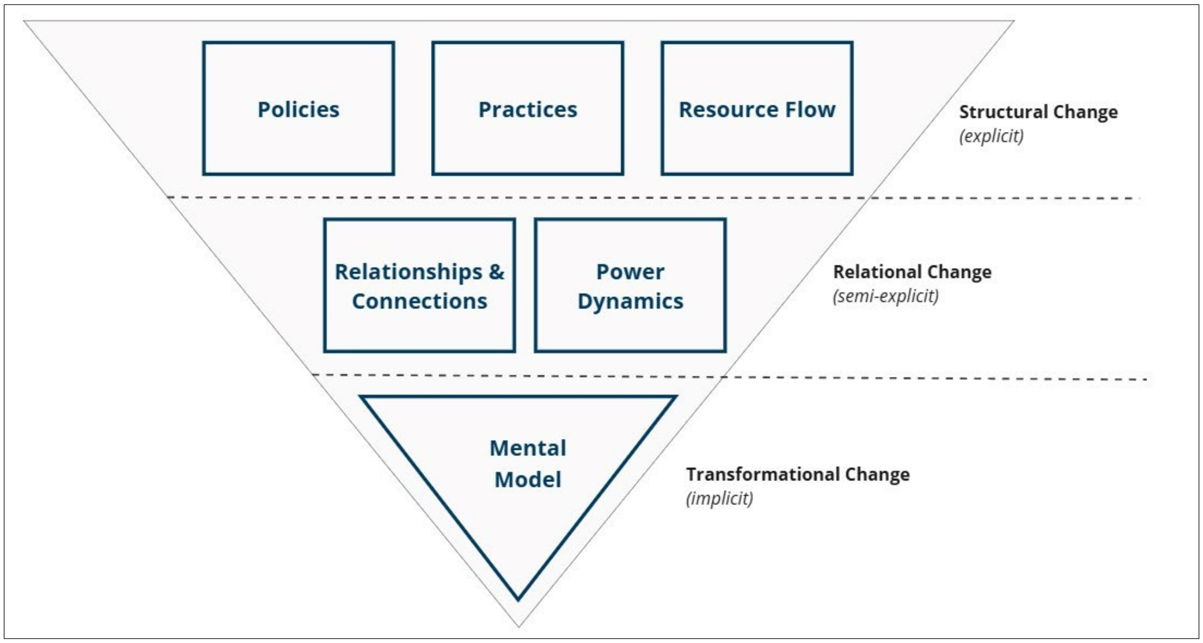


*Figure 11: System change conceptual framework (Kania et al., 2018)*

# 4. Articulating the Hallmarks of Systemic Investing Evaluation (Responding to RQ2)

This chapter responds to the second research question:
**RQ2: What does systemic investing evaluation look like and what are its characteristics?**

To address this question, the chapter introduces and defines a set of **13 hallmarks**, those are the characteristics that collectively describe how systemic investing evaluation differs from conventional approaches. These hallmarks were developed through an iterative process that combines theories from systems thinking, evaluation literature, and systemic investing practice.

## 4.1 Introduction of Hallmarks

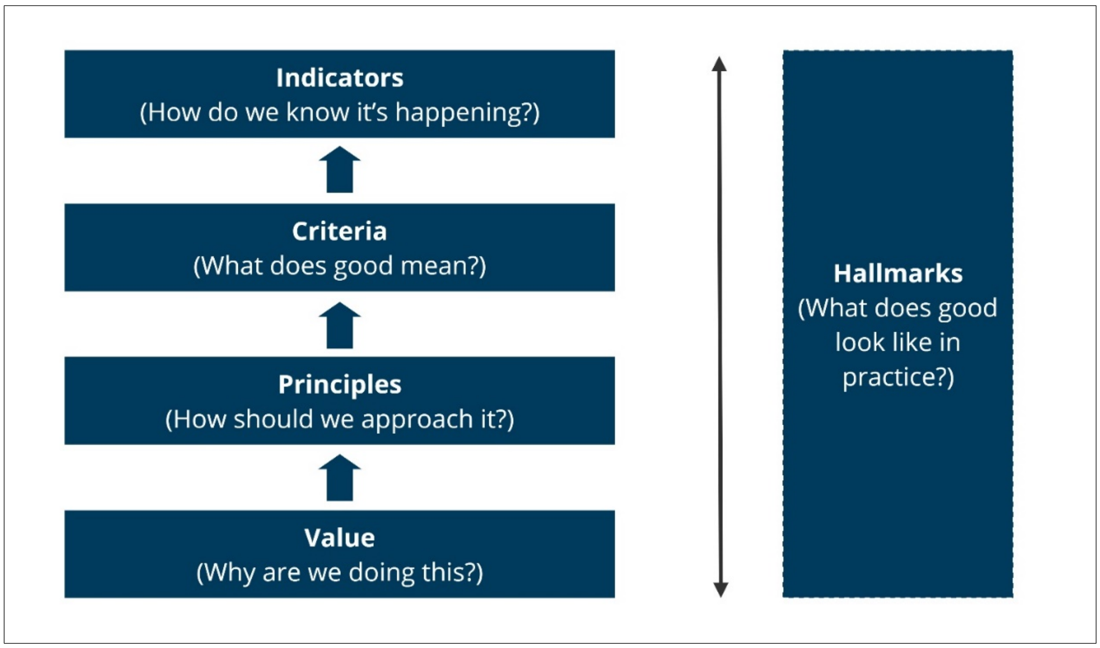


*Figure 12: Evaluation concepts*

In evaluation theory, values, principles, criteria, and indicators are different concepts. Figure 12 is trying to summary those concepts. At the foundation of this structure are *values*, which refer to deeply held beliefs about what is important, worthwhile, or desirable (Teasdale, 2021; House, 2015). These values are often implicit but play a critical role in shaping what is ultimately evaluated and how. Building on values, *principles* provide normative guidance for evaluation practice. As Buckton et al. (2024) argue in their review of transformation-focused evaluation, principles help evaluators approach complex systems with appropriate orientations such as reflexivity, power awareness, and adaptive learning. They answer the question: How should we conduct evaluations that are aligned with transformation?

Translating principles into practice requires the development of *criteria*, which is the specific dimensions or standards against which the quality of an intervention is assessed. Criteria serve as the operational expressions of principles and demonstrate the logic of evaluation (Scriven, 1991; Gates et al., 2024). Teasdale (2021) further refines this view by distinguishing between the domain of a criterion (what is being judged) and its source (whose values it reflects), reinforcing that criteria are always socially constructed and context dependent. Finally, *indicators* are concrete, often measurable signals tied to specific criteria. They enable evaluators to gather evidence and render judgments.

This thesis builds on that layered logic but introduces *hallmarks as "what does good look like in practice"*, Hallmarks are defined as recognizable features or patterns for systemic investing evaluation in practice. They make the principle/criteria/value observable. By focusing on *hallmarks* and their corresponding *indicators*, this study aims to provide a practical and theoretically grounded framework for evaluating whether and how systemic investing practices reflect transformational intent.

## 4.2 11 Hallmarks emerged from literature review

As described in Chapter 3, which identified eleven key trends in evaluation approaches, the table below summarizes the transition from traditional evaluation methods to systemic evaluation approaches across multiple dimensions. Each row represents a shift in evaluation practice, highlighting how conventional approaches have evolved to address the complexity, interconnectedness, and dynamic nature of modern social and environmental challenges. The right column identifies the corresponding hallmark that characterizes each aspect of this systemic transformation. These hallmarks collectively represent a framework for understanding and implementing evaluation approaches that are better suited to capturing and influencing systemic change in complex contexts.

*Table 3: Summary of trends*

| Trend | Traditional Approach | Systemic Approach | Hallmark |
|---|---|---|---|
| 1 | Viewing problems as discrete and isolated | Understanding problems as complex and interdependent | Systems Thinking and Complexity Science |
| 2 | Determining fixed value based on predetermined criteria | Developing evolving value through continuous reflection and adaptation | Paradigm Evolution |
| 3 | Short-term metrics and immediate outcomes | Multi-generational timeframes and extended impact pathways | Long-Term Horizon |
| 4 | Assessing individual programs in isolation | Examining pathways toward system change across micro, meso, and macro levels | Multi-Level Assessment |
| 5 | Single-sector approaches and siloed evaluation | Multi-stakeholder, boundary-spanning collaborations | Cross-Sector Collaboration |
| 6 | Treating stakeholders as data sources or passive recipients | Engaging stakeholders as co-evaluators and knowledge co-creators | Stakeholder Engagement |

| | | | |
|---|---|---|---|
| 7 | Measuring programmatic outputs and direct outcomes | Assessing modifications in underlying system structures | System Structure Changes |
| 8 | Rigid, compliance-oriented funding frameworks | Adaptive, co-created policy environments that enable systemic change | Promote Policy Change and Adaptability |
| 9 | Treating evaluation as a bounded activity with fixed beginnings and endpoints | Embracing evaluation as a continuous, embedded process | Dynamic Assessment |
| 10 | Single-loop learning focused on program improvement | Double and triple-loop learning that questions fundamental assumptions | Adaptive Learning |
| 11 | Standardized evaluation tools and linear methodologies | Context-specific, systems-oriented methodologies | Systems-Based Techniques |

## 4.3 2 Hallmarks emerged from case studies

Systemic investing involves not only deploying capital but transforming the way capital flows through a system. The twelfth hallmark, **multicapital circulation**, emphasizes the deliberate use of multiple forms of capital, such as financial, social, natural, human, to enable systems-level change. Across the 46 cases analyzed, this hallmark was strongly expressed in initiatives that wove together public funds, philanthropic dollars, and community capital, to create reinforcing effects.

For instance, the Tara Health Foundation explicitly described its strategy as “pulling all of our programs together”, integrating grants, investments, and partnerships to advance reproductive equity. Rather than treating each capital type as separate, the foundation designed funding mechanisms to facilitate capital for greater systemic influence. Several initiatives also used creative financial designs: Tidal Impact mentioned that “Our plans include providing different types of capital to adapt to the different needs of companies in our pipeline”. Presencing Institute pointed out that “all the resources to decide and implement these things came from within the system, including the time, the infrastructure, and the finances”. These examples demonstrate that multicapital circulation is being operationalized by practitioners.

Closely linked is the thirteenth hallmark: **capital catalyzation**. Where multicapital circulation focuses on diversity, capital catalyzation emphasizes momentum, it means whether initial capital investments are structured in ways that attract further capital, expand reach, or influence larger financial flows. This was one of the most consistently observed patterns across cases. In the Tara Health Foundation, leaders explicitly defined success as “being catalytic”, in another word, not just investing their own funds but crowding in co-investors, shifting market narratives, and inspiring policy change. Several initiatives used specific mechanisms to generate follow-on investment. In the Regenerative Alpine Region case, early-stage investors created “pledge funds” to bring in co-investors. The Library of Things designed a franchise model that scaled through replication, attracting municipal and private funding in new geographies. Apolitical provided concessionary capital to reduce risk in early stages, enabling later commercial participation.

Together, these two hallmarks show how systemic investing is not just about how much capital is deployed, but how intelligently it flows and grows. Initiatives that circulate diverse forms of capital while catalyzing additional contributions tend to generate more resilient and scalable systems change. Evaluators can support this by tracking capital dynamics over time, including both diversity of inputs and the extent to which early capital unlocks new possibilities.

## 4.4 Integrated Hallmarks Framework

Building upon the hallmark development discussed earlier, both from academic research and practical applications, and in alignment with the "Water of System Change" Framework outlined in Chapter 3, the figure below presents 13 hallmarks framework. This framework organizes the hallmarks according to the six conditions for systems change identified by Kania, Kramer, and Senge (2018): mental models, relationships and connections, power dynamics, policies, practices, and resource flows. Each condition represents a different level of intervention, ranging from explicit structural elements to implicit transformative factors. The hallmarks are positioned within these conditions to illustrate how systemic evaluation should operate across multiple dimensions. This approach recognizes that effective evaluation in complex systems requires attention to both visible elements (like policies and practices) and less visible factors (such as mental models and power relationships) that collectively determine a system's behavior and potential for transformation.

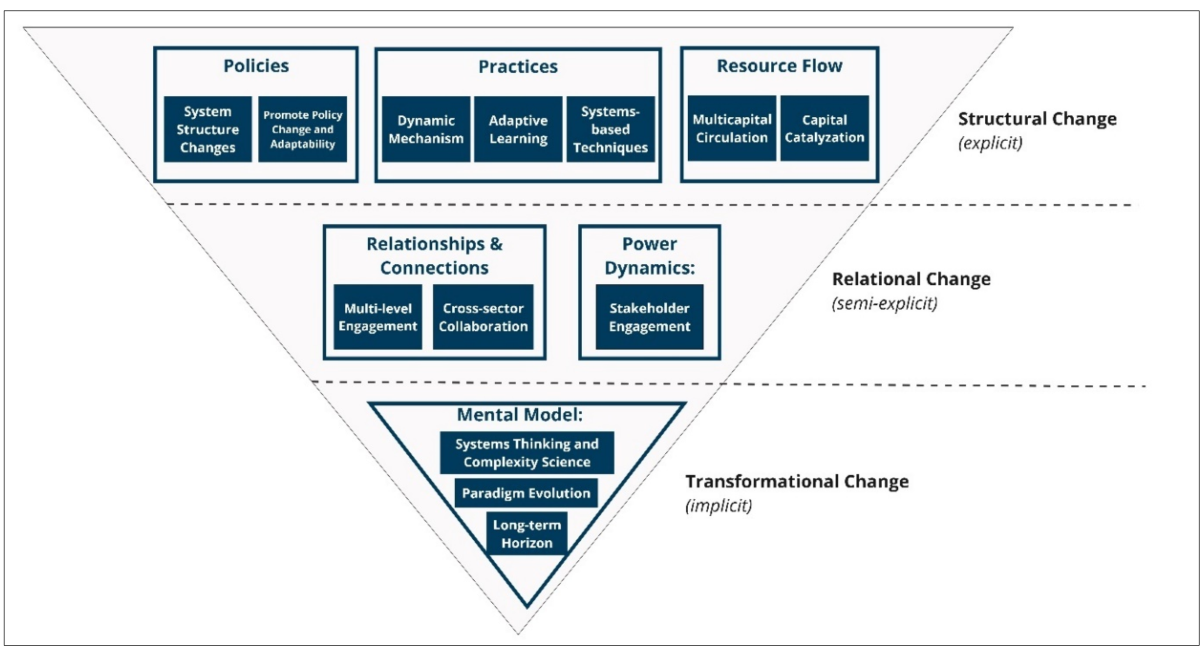


*Figure 13: 13 Hallmarks Framework*

In this research, I conceptualize systemic investing and systemic evaluation as interconnected but distinct fields that together form the foundation of the Systemic Investing Evaluation approach. Based on an iterative review of theoretical literature and practitioner insights, I developed a set of hallmarks that reflect attributes of systemic change. These hallmarks span both investment and learning practices, recognizing the interdependence between how capital is deployed and how impact is assessed and understood.

The table *"Definition and Implication of Hallmarks Framework"* reflects this integrated perspective. For each hallmark, I provide a definition, followed by two application columns, one describing implications for investment practice, and another for evaluation practice. This structure makes explicit how each hallmark can manifest in both domains, while also recognizing that some hallmarks may be more applicable or observable in one domain than the other although most of the time there is a weaving of capital deployment and learning processes and blurring boundaries.

## 4.5 Funding of Systemic Evaluation

Despite the emphasis on continuous learning and adaptive evaluation in systemic investing, the funding of evaluation itself remains a persistent and structural challenge. Early-stage ventures or systemic initiatives often lack the financial capacity to support robust evaluation, while investors frequently expect investees to assume responsibility for demonstrating impact. This disconnect leads to chronic underinvestment in evaluation and undermines the infrastructure needed to generate credible, system-level insights.

A systemic investing approach should therefore recognize the underfunding of measurement and evaluation as a systemic issue in its own right. Just as capital is allocated to finance programmatic interventions, there should also be intentional and structured investment in the evaluation and learning processes that support them. Without this, efforts to assess transformation across complex systems risk being superficial or misaligned with the true dynamics of change. Accordingly, this thesis explicitly incorporates the need for sustained, flexible, and systemically embedded funding of evaluation as an implication under Hallmark 12 (Multicapital Circulation) and Hallmark 13 (Capital Catalyzation), highlighting it as an essential component of effective systemic investing practice.

*Table 4: The Definition and Implication of Hallmarks Framework*

| System Change Levels | System Change Conditions | Hallmarks | Definition | Implications for Investment Practice (Systemic Investing is characterized by...) | Implications for Evaluation (Systemic Investing Evaluation is characterized by...) |
|---|---|---|---|---|---|
| Transfor-mational Change (implicit) | Mental Model | **Hallmark 1: Systems Thinking and Complexity Science** | *Embracing systems thinking and complexity science means recognizing that social, environmental, and financial issues are interrelated parts of complex adaptive systems. We should adopt holistic perspectives, examine nonlinear cause-effect relationships, and use concepts from complexity science (e.g. feedback loops, emergence) to inform decisions.* | Investors map the system they seek to understand how to better deploy capital, identify key actors, relationships, and feedback loops that maintain current conditions. They analyze potential ripple effects of investments beyond immediate targets and design investment strategies that address root causes rather than symptoms. | Evaluators use systems mapping and complexity frameworks to design evaluation approaches, looking beyond linear causality to capture emergent, unexpected outcomes. They assess how interventions influence system dynamics and identify changes in feedback loops, not just discrete outcomes. |
| | | **Hallmark 2: Paradigm Evolution** | *Paradigm evolution refers to shifts in the fundamental mindset or worldview guiding an evaluation approach. It involves questioning and evolving the underlying assumptions (or "mental models") about how change happens, moving beyond traditional paradigms (e.g. short-term profit maximization or linear problem-solving) to new paradigms aligned with systemic change.* | Investors examine their own assumptions about how change happens and are willing to question traditional investment models. They develop investment approaches that challenge status quo thinking and measure value beyond financial returns, embracing new frameworks that recognize systemic impact. | Evaluators examine the mental models underlying both the intervention and evaluation approach. They track shifts in mindsets and narratives among stakeholders as a key outcome and employ methods that capture paradigm shifts rather than only measuring predefined outcomes. |
| | | **Hallmark 3: Long-Term Horizon** | *A long-term horizon means planning, investing, and evaluating with a far-reaching timeframe in mind. Instead of focusing on quick wins or short-term outputs, systemic approaches consider long-term outcomes and impacts, recognizing that systemic change often unfolds over years or decades.* | Investors commit to patient capital timeframes aligned with the pace of systemic change, often 10+ years. They design blended return expectations that account for long-term value creation and establish intermediate milestones to track progress toward long-term goals. | Evaluators design evaluation frameworks that span appropriate time horizons for systemic change, using developmental approaches that track progress over extended periods. They identify meaningful leading indicators that signal progress toward long-term outcomes and balance assessment of short-term wins with longer-term transformation. |
| Relational Change (semi-explicit) | Relation-ships & Connections | **Hallmark 4: Multi-Level Engagement** | *Multi-level Engagement refers to enabling change at multiple levels of a system -- for example, at the level of individual participants, organizations, networks, and the broader system or society. It recognizes that systemic change involves interactions across micro, meso, and macro levels.* | Investors design investment strategies that target multiple levels of a system simultaneously, recognizing that portfolio-level impact often exceeds the sum of individual investments. They balance direct investments in organizations with investments in enabling infrastructure, field-building, and policy change. | Evaluators assess outcomes at multiple levels (individual, organizational, network, field, system) and analyze interactions between these levels. They employ methods that capture both "zoom in" (specific organizations/projects) and "zoom out" (system-wide) perspectives, examining how change at one level influences other levels. |
| | | **Hallmark 5: Cross-Sector Collaboration** | *Cross-sector collaboration is the cooperative effort among actors from different sectors -- typically the public (government), private (business), and social (non-profit, civil society) sectors -- toward a shared goal. In systemic initiatives, complex problems often span sector boundaries, so collaboration is needed to leverage diverse resources, align policies and practices, and address issues holistically.* | Investors actively build partnerships across public, private, and social sectors rather than operating in silos. They participate in or initiate collaborative funding mechanisms (e.g., blended finance vehicles or donor collaboratives) and coordinate strategies with other capital providers. | Evaluators bring together diverse perspectives from public, private, and social sectors to establish shared evaluation frameworks and collective definitions of success. Cross-sector evaluation approaches leverage complementary expertise, resources, and data sources from different stakeholders, enabling more comprehensive assessment of complex initiatives that span traditional sector boundaries. |

| | | | | | |
|---|---|---|---|---|---|
| | Power Dynamics | **Hallmark 6: Stakeholder Engagement** | *Stakeholder engagement means actively involving those who have a stake in or are affected by an initiative in its design, implementation, and evaluation. This often translates to participatory approaches that include diverse voices (beneficiaries, community members, implementing partners, etc.), ensuring that the process empowers stakeholders.* | Investors engage affected communities and stakeholders in investment decision-making, especially those typically excluded from financial decisions. They establish feedback mechanisms that give voice to beneficiaries and design participatory processes that share power with communities being served. | Evaluators employ participatory evaluation methods that meaningfully involve diverse stakeholders throughout the evaluation process. They create opportunities for beneficiaries and affected communities to define what success looks like and interpret findings and ensure evaluation processes promote equity and inclusion. |
| Structural Change (explicit) | Policies | **Hallmark 7: System Structure Changes** | *System structure changes refer to alterations in the fundamental structural elements of a system -- such as policies, regulations, incentive structures, organizational forms, or information flows. In systemic evaluation, this hallmark looks for evidence that an intervention is changing the "rules of the game" or the architecture of the system, not just delivering incremental benefits within the existing structure.* | Investors target investments that have potential to change underlying system structures, not just operate within them. They balance funding for direct services with investments in advocacy, policy change, and new institutional arrangements that shift incentives and rules. | Evaluators assess changes in formal and informal rules governing the system (policies, regulations, norms). They track shifts in institutional arrangements, incentive structures, and information flows that may indicate deeper structural change, and document how structural changes enable broader system transformation. |
| | | **Hallmark 8: Promote Policy Change and Adaptability** | *Policies encompass both internal and external dimensions of systemic change. Internally, funders and investors shape system behavior by establishing agreements, funding structures, timelines, scopes, and accountability mechanisms when deploying capital. These parameters influence how implementing partners act and adapt. Externally, policy also refers to broader institutional, regulatory, and governmental rules that structure the operating environment.* | Investors design flexible funding mechanisms that adapt to evolving contexts and partner needs. They coordinate with other funders to create coherent, aligned requirements that reduce administrative burden, and support policy advocacy and reform efforts that create enabling conditions for systemic solutions. | Systemic evaluation thrives when supported by thoughtful policies at both organizational and macro levels. Within organizations, evaluation policies that encourage flexibility, innovation, and appropriate risk-taking create space for evaluators to employ adaptive methods suited to complex contexts. At broader levels, an enabling policy environment helps evaluators to move beyond conventional methods. |
| | Practices | **Hallmark 9: Dynamic Mechanism** | *Dynamic mechanism means continuously or iteratively renewing an initiative in real-time, acknowledging that conditions change, and the intervention itself may need to adapt. It contrasts with one-off, static evaluations. This hallmark is about monitoring system changes and program effects as they emerge and using feedback loops to inform ongoing decision-making.* | Investors establish continuous feedback mechanisms rather than relying solely on annual reporting cycles. They create regular learning touchpoints with investees and remain flexible about shifting strategies and metrics as contexts change, treating investment plans as living documents. | Evaluators design iterative, ongoing assessment processes rather than point-in-time evaluations. They implement real-time data collection and sense-making routines that inform adaptive management and remain flexible about evaluation questions and methods as the initiative evolves. |
| | | **Hallmark 10: Adaptive Learning** | *1) Adaptive learning is the process by which organizations and initiatives systematically learn from experience and evidence and then adapt their strategies and actions accordingly. 2) In addition, adaptive learning also encompasses structured education and training initiatives that build capacity within organizations. By equipping with the necessary knowledge to understand and engage with complex systems, these training efforts foster a mindset shift and enable more effective adaptation and long-term systemic change.* | Investors build learning routines that go beyond compliance monitoring to question assumptions, daily behaviors, and theory of change. They provide capacity-building support for systems thinking and adaptive management within the organisation itself and investee organizations, this involves programs such as workshops, board training sessions, and skill development courses that familiarize key stakeholders with systems science concepts, tools, and practices. | Evaluators facilitate reflective practice and learning processes, not just measurement. They design evaluation approaches that support double and triple-loop learning (questioning assumptions and paradigms), and help organizations build their internal evaluation and learning capacity. |

| | | | | | |
|---|---|---|---|---|---|
| | | **Hallmark 11: Systems-based Techniques** | *Technical application in systemic evaluation refers to the use of specialized methods, tools, and analytical techniques (often from systems science or related technical fields) to better understand and influence complex systems. This can include systems mapping, modeling, network analysis, use of data science, scenario planning, and other techniques that complement the soft skills of facilitation and learning to assist systemic evaluation.* | Investors employ technical tools from systems science to inform investment decisions and portfolio design. They utilize methods like system mapping, network analysis, or system dynamics to identify leverage points and potential impacts and build capacity to apply these tools themselves or through partnerships. | Evaluators incorporate system-based analytical techniques alongside traditional evaluation methods. They apply appropriate technical tools (e.g., system dynamics modeling, network analysis, contribution analysis) to understand complex change processes, and help stakeholders interpret and use insights from these technical approaches. |
| | Resource Flow | **Hallmark 12: Multicapital Circulation** | *Multicapital circulation refers to the mobilization and flow of multiple forms of capital within a system, beyond just financial capital. This includes natural capital, social capital, human capital, cultural capital, symbolic capital, and other "capitals" that accumulate over time.* | Investors recognize and intentionally deploy multiple forms of capital beyond financial resources. They track how their investments impact various capital stocks (natural, social, human, etc.) and design strategies that regenerate rather than deplete these capitals. | Systemic evaluation under multicapital logic requires the capacity to observe not only the deployment but also the accumulation and distribution of diverse forms of capital over time. This involves tracing how financial, social, intellectual, human, and cultural capital are mobilized within the system and assessing who gains access to and benefits from these resources. Evaluators should be able to detect whether capital circulation aligns with the intended impact goals, and whether unintended accumulation patterns emerge through processes like stakeholder convenings, capacity building, or visibility campaigns. A robust evaluation examines how relationships are forged, knowledge is produced, and reputations evolve, in addition to tracking financial flows.<br><br>At the same time, the evaluation process itself should be sufficiently resourced to carry out this task. This includes access to financial capital (adequate funding), social capital (networks and trust), intellectual capital (appropriate methods and contextual knowledge), and symbolic capital (credibility and perceived legitimacy). Without these forms of capital, evaluators may struggle to engage stakeholders meaningfully, interpret multicapital dynamics, or hold actors accountable to their systemic commitments. |
| | | **Hallmark 13: Capital Catalyzation** | *Capital catalyzation refers to the use of capital in a way that catalyzes or leverages additional resources for impact. This often means employing investment strategies that crowd in other investors, de-risk innovative models, or kick-start market mechanisms - using capital as a spark plug for larger systemic change.* | Investors structure capital to mobilize additional resources and unlock new funding streams. They employ catalytic tools like first-loss capital, guarantees, or anchor investments that reduce risk for others, and measure success partly by how much additional capital they attract to the issue. | Evaluating capital catalyzation requires attention to additionality and the sequencing of capital flows. A systemic evaluation approach asks whether the investment enabled actions or attracted capital that would not have occurred otherwise. This often involves constructing plausible counterfactuals and identifying catalytic moments (such as when philanthropic or public capital de-risks a venture that subsequently draws in commercial investors). Capturing these dynamics entails evaluating both ex-ante investment decision processes (e.g., selection for catalytic potential) and ex-post outcomes (e.g., whether follow-on investment occurred due to the initial action), thereby tracing whether and how catalytic leverage was achieved. |

| | | | | | |
|---|---|---|---|---|---|
| | | | | | At the same time, the evaluation of catalytic capital mechanisms often suffers from a lack of dedicated funding. Systemic evaluation acknowledges that measurement and learning infrastructure itself is a systemic concern. Many early-stage ventures or systemic initiatives lack the capacity to fund robust evaluation, while investors often expect investees to take responsibility for impact claims. This misalignment leads to underinvestment in evaluation and a reversion toward finance-only logics. Therefore, systemic investing architectures should explicitly include sustained, flexible, and appropriately structured funding for evaluation as part of the capital being catalyzed. This ensures that learning cycles are resourced and embedded in the broader transformation effort. |

# 5. Developing Indicators for Systemic Investing Evaluation (Responding to RQ3)

This chapter responds to the third research question:
**RQ3: How can practitioners begin to adopt and implement systemic investing evaluation in practice?**

Having established a set of thirteen hallmarks in the previous chapter to characterize systemic investing evaluation, this chapter focuses on their practical application. Specifically, it develops a corresponding indicator system that translates each hallmark into observable and assessable elements. These indicators are intended to support evaluators, funders, and practitioners in identifying whether and to what extent a given initiative embodies the principles of systemic change.

## 5.1 Introduction

As outlined in the methodology, the indicator system was developed through an analysis of 46 cases sourced from the TWIST and FEST networks, resulting in a total of 616 data points and 112 indicators. The 616 data points are the original description came from interviews, surveys, presentation slides, and other available materials. Each piece of information was first synthesized to capture its core meaning, from which proposed indicators were generated. Where multiple indicators pointed to similar evaluative concerns or practices, they were grouped under shared themes or objectives and consolidated into unified indicators. Each indicator was classified as either qualitative or quantitative, depending on its nature. The "support case number" indicates how many of the 46 cases explicitly reflect or support a given indicator. It is important to note that this classification process was not intended to be exhaustive; rather, it focused on identifying and clustering the most salient features that emerged across the cases. The goal was to find recurring patterns and insights, while some contextual nuances may not be fully captured.

Five patterns identified through the review of the indicators across the 13 hallmarks.

First, the integration of systems thinking and multi-level analysis throughout the framework. Most initiatives do not limit themselves to single-issue or enterprise-level metrics; instead, they employ system mapping, identification of leverage points, and assessment across individual, organizational, and ecosystem levels. This is reflected in indicators such as the comprehensiveness of systems maps, the number of system levels assessed, and the explicit tracking of cross-sectoral pathways. These practices demonstrate a shared commitment to understanding and influencing the broader context in which investments operate, rather than focusing solely on isolated outcomes.

Another pattern is the emphasis on adaptability, learning, and long-term perspective. Indicators frequently capture the frequency of strategic reviews, the presence of iterative assessment frameworks, and the documentation of strategic pivots or adaptations. Many initiatives track not only short-term outputs but also the durability and evolution of outcomes over time, using long-term metrics and phased approaches. This reflects a recognition that systemic change should not be linear or immediate (as traditional approaches indicate), and that ongoing learning is necessary to respond to emerging insights and shifting system dynamics.

Third, the circulation of multiple forms of capital, both financial and non-financial. Indicators in this area focus on the diversity of capital types deployed, the integration of blended finance structures, and the mobilization of technical assistance and network resources alongside investment capital. The measurement of capital leverage ratios and the tracking of follow-on investment further highlight the catalytic intent behind many systemic investing strategies. These approaches are often complemented by efforts to democratize capital flows and broaden participation in investment processes.

Fourth, stakeholder engagement and participatory governance also emerge as the central theme. Many indicators measure the breadth and depth of stakeholder involvement in decision-making, the inclusion of marginalized or unconventional voices, and the presence of feedback loops that inform adaptation. There is a clear trend toward embedding co-creation, power-sharing, and community-led mechanisms within both investment and evaluation practices, with the aim of the sustainability of outcomes.

Finally, the hallmarks collectively highlight the challenge of measuring paradigm shifts and structural change. While some indicators attempt to track changes in mental models, organizational missions, or policy environments, these areas remain difficult to quantify and standardize. As a result, qualitative evidence, such as narrative change, demonstration of new models, and shifts in governance or policy, is often used along quantitative metrics.

## 5.2 Synthesized Indicators

This framework leans more toward assessing the “systemicness of practice”—the degree to which initiatives reflect systemic intent and methods—rather than producing a comprehensive view of systemic outcomes. This emphasis reflects several underlying factors. First, the systemic investing efforts examined through the TWIST network are largely in process and emergent; most have not yet operated long enough for broader system-level effects to be fully observable. Second, across the cases, dedicated investment in outcome-level evaluation remains limited, and many initiatives are still exploring how to assess long-term change. Third, evaluation activities and impacts are not always visible or explicitly communicated in the source materials. In several instances, important insights only surfaced during informal conversations or conference settings when deeper probing occurred.

As a result, the indicators in this framework largely reflect how systemic intent is operationalized—through strategies, relationships, and experimentation—rather than the full arc of systems transformation. This orientation should not be interpreted as a dismissal of outcomes, but rather as a reflection of the current developmental stage of the field, as well as a methodological constraint shaped by available data.

*Table 5: Synthesized Indicator from 46 Case Studies*

| Hallmarks | Theme/Objective | Synthesized Indicators | Indicator Type | Support Case Volume |
|---|---|---|---|---|
| **1. Systems Thinking and Complexity Science** | Comprehensiveness of Systems Mapping | Number of system dimensions/nodes identified | Quantitative | 8 |
| | | Number of feedback loops or causal relationships mapped | Quantitative | |
| | Identify Leverage Points | Number of intervention points or leverage points identified through analysis | Quantitative | 5 |
| | Breadth of Stakeholder Representation in Mapping | Number/diversity of stakeholder groups engaged in systems mapping | Quantitative | 4 |
| | | Extent and quality of stakeholder engagement in system mapping and analysis | Qualitative | 4 |
| | Depth of Systems Analysis | Number of system elements and relationships explicitly considered in strategy documents | Quantitative | 4 |
| | Rigor and Sophistication of Systems Techniques | Qualitative assessment of systems mapping rigor (e.g., use of causal loop diagrams, scenario modeling, network analysis) | Qualitative | 9 |
| | Integration of Systems Thinking in Decision-making | Evidence that systems analysis directly informs investment or strategy decisions | Qualitative | 6 |
| | Boundary Definition and System Framing | Clarity and rationale for system boundaries and scope; explicit articulation of system elements and relationships | Qualitative | 4 |
| | Recognition and Management of Complexity/Emergence | Evidence of tracking emergent properties, unintended consequences, or adaptation to system surprises | Qualitative | 4 |
| **2. Paradigm Evolution** | Stakeholder Mindset Change | Percentage of stakeholders (staff, partners, community) reporting a shift in mindset or mental models | Quantitative | 4 |
| | | Depth and diversity of stakeholder involvement in co-creating new paradigms | Qualitative | 5 |
| | Adoption of New Paradigm Practices | Number/percentage of portfolio organizations adopting new paradigm-aligned practices or models | Quantitative | 5 |
| | | Evidence of new language, narratives, or framing in organizational communications | Qualitative | 5 |
| | Demonstration of Alternative Models | Existence and visibility of pilot projects or case studies that embody the new paradigm | Qualitative | 5 |
| **3. Long-Term Horizon** | Strategic Planning Horizon | Number of years encompassed in strategic planning | Quantitative | 9 |
| | | Presence of explicit 5+, 10+, or 20+ year plans | Quantitative | |
| | Investment Holding Period | Average holding period of investments (years) | Quantitative | 4 |
| | Sustainable Entities | Number of organizations/incubated entities designed to outlive the funder/foundation | Quantitative | 3 |
| | Long-Term Metrics in Evaluation | Presence and frequency of long-term (5+ or 10+ year) outcome metrics in evaluation frameworks | Quantitative | 5 |
| | Long-Term Vision and Commitment | Evidence of explicit long-term vision in strategy documents, interviews, or public communications | Qualitative | 5 |
| | Phase-Based or Multi-Stage Approach | Use of phased strategies or multi-stage plans to achieve system change over time | Qualitative | 4 |

| | | | | |
|---|---|---|---|---|
| | Legacy Planning and Exit Strategy | Qualitative evidence of planning for organizational legacy, handover, or sustainability beyond the funder | Qualitative | 3 |
| | Recognition of Temporal Trade-Offs | Explicit discussion of short-term vs. long-term impact trade-offs and prioritization of long-term gains | Qualitative | 3 |
| **4. Multi-level Assessment** | System Levels Assessed | Number of distinct system levels (e.g., individual, organizational, ecosystem, policy) included | Quantitative | 6 |
| | Metrics Addressing Multiple Levels | Percentage of impact metrics/KPIs that explicitly track change at more than one system level | Quantitative | 6 |
| | Aggregation Capability Across Levels | Percentage of KPIs that can be aggregated from enterprise to portfolio to ecosystem level | Quantitative | 4 |
| | Breadth of Stakeholder Types | Number of different stakeholder groups whose outcomes are measured at multiple levels | Quantitative | 5 |
| | Multi-level Theory of Change | Existence and clarity of a theory of change that links outcomes across system levels | Qualitative | 5 |
| | Integration of Micro, Meso, Macro Perspectives | Evidence of assessment frameworks or narratives that deliberately integrate micro (individual), meso (organizational/community), and macro (system/policy) perspectives | Qualitative | 5 |
| | Adaptation of Assessment Tools for Multi-level Use | Use of customized or flexible tools (e.g., dashboards, frameworks) to track impact at different levels | Qualitative | 4 |
| | Stakeholder Participation in Multi-level Assessment | Degree to which stakeholders at different levels are engaged in defining, collecting, and interpreting impact data | Qualitative | 4 |
| **5. Cross-sector Collaboration** | Number of Cross-sector Partnerships | Number of formal partnerships involving organizations from at least two distinct sectors | Quantitative | 10 |
| | Diversity of Sectors Represented | Number of different sectors (e.g., business, government, civil society, academia) engaged in collaboration | Quantitative | 8 |
| | Frequency of Collaboration | Number of joint initiatives, meetings, or projects conducted per year across sectors | Quantitative | 5 |
| | Cross-sector Governance Structures | Number of cross-sector steering committees or shared governance bodies established | Quantitative | 4 |
| | Depth of Collaboration | Evidence of co-design, shared goal-setting, and joint ownership of outcomes | Qualitative | 5 |
| | Integration of Unconventional Partners | Inclusion of non-traditional actors (e.g., lived experience panels, grassroots organizations, startups) | Qualitative | 5 |
| | Adaptability of Collaborative Structures | Evidence of evolving partnership models or flexible collaboration arrangements | Qualitative | 4 |
| | Collaborative Influence on System Change | Qualitative evidence that cross-sector collaboration led to new policies, practices, or market shifts | Qualitative | 4 |
| **6. Stakeholder Engagement** | Diversity of Stakeholder Groups Engaged | Number of distinct stakeholder groups actively engaged (e.g., by sector, demographic, role) | Quantitative | 10 |
| | Breadth of Stakeholder Input in Decision-Making | Percentage of key decisions (e.g., investment, program design) involving direct stakeholder input | Quantitative | 7 |

| | | | | |
|---|---|---|---|---|
| | Stakeholder Engagement Frequency | Number of engagement events (workshops, forums, consultations) per year | Quantitative | 5 |
| | Stakeholder Representation in Governance | Percentage of governance or advisory roles held by stakeholders from target communities | Quantitative | 4 |
| | Depth and Quality of Engagement Processes | Evidence of co-creation, participatory design, or empowered stakeholder roles | Qualitative | 6 |
| | Inclusion of Marginalized or Unconventional Voices | Inclusion of lived experience panels, grassroots groups, or underrepresented communities | Qualitative | 6 |
| | Power Redistribution and Shared Decision-Making | Qualitative evidence of shifting authority to stakeholders (e.g., community-led investment, voting) | Qualitative | 4 |
| | Feedback Loops and Responsiveness | Mechanisms for stakeholders to provide feedback and see it reflected in adaptations | Qualitative | 4 |
| **7. System Structure Changes** | Number of Policy or Regulatory Changes Influenced | Number of policy, regulatory, or legislative changes directly attributed to the initiative or investment | Quantitative | 6 |
| | Creation of New Institutional Arrangements | Number of new governance bodies, platforms, or market mechanisms established | Quantitative | 5 |
| | Adoption of New Organizational Practices | Number of organizations adopting new practices, standards, or business models as a result of the initiative | Quantitative | 4 |
| | Structural Barriers Addressed | Number of specific system-level barriers (e.g., tariffs, licensing, access bottlenecks) removed or reduced | Quantitative | 3 |
| | Depth of Structural Change | Evidence of fundamental shifts in system architecture, power dynamics, or value flows | Qualitative | 5 |
| | Role in Enabling Systemic Replication | Qualitative evidence that new structures or models are designed for replication or scaling across contexts | Qualitative | 4 |
| | Stakeholder Participation in Structural Change | Degree to which affected stakeholders are involved in designing or governing new structures | Qualitative | 4 |
| | Alignment with Systemic Goals | Evidence that structural changes are explicitly linked to broader systemic outcomes (e.g., equity, resilience, sustainability) | Qualitative | 3 |
| **8. Promote Policy Change and Adaptability** | Number of Policy Changes Influenced | Number of policy, legislative, or regulatory changes directly influenced or enacted | Quantitative | 7 |
| | Policy Engagement Activities | Number of policy advocacy, engagement, or consultation activities conducted per year | Quantitative | 4 |
| | Public Opinion Shift Toward Policy Goals | Number or percentage of documented shifts in public opinion supporting policy change | Quantitative | 3 |

| | | | | |
|---|---|---|---|---|
| | Frequency of Policy Framework Updates | Frequency of updates or adaptations to policy-related frameworks, strategies, or tools | Quantitative | 2 |
| | Depth of Policy Influence | Evidence of shaping not just policy content but also policy-making processes and stakeholder participation | Qualitative | 4 |
| | Adaptability in Policy Engagement | Evidence of timing, flexibility, and responsiveness in policy advocacy strategies | Qualitative | 3 |
| | Cross-sector and Multi-stakeholder Policy Work | Quality and diversity of cross-sector and stakeholder involvement in policy change efforts | Qualitative | 4 |
| | Policy Narrative and Framing Shifts | Qualitative evidence of new narratives, frames, or language influencing policy debates and perceptions | Qualitative | 4 |
| **9. Dynamic Assessment** | Frequency of Formal Strategy or Impact Reviews | Number of formal strategy or impact assessment review sessions per year | Quantitative | 8 |
| | Update Rate of KPIs/Assessment Frameworks | Percentage of KPIs or impact metrics updated or refined annually based on new data or learning | Quantitative | 4 |
| | Variance Analysis Between Projected and Actual Impact | Number of variance analyses conducted comparing projected vs. actual impact metrics | Quantitative | 3 |
| | Implementation of Progress Tracking Systems | Existence and utilization of formal progress tracking systems for ongoing monitoring | Quantitative | 3 |
| | Integration of Feedback Loops | Evidence of structured feedback mechanisms (e.g., stakeholder or community input) informing strategy adjustments | Qualitative | 6 |
| | Responsiveness to Emerging Insights | Documentation of strategic pivots or adaptations made in response to new information or system changes | Qualitative | 5 |
| | Evolution of Assessment Frameworks | Qualitative evidence of evolving or iterative assessment frameworks (not static, but regularly improved) | Qualitative | 3 |
| | Learning from Failure and Success | Evidence that both positive and negative outcomes are systematically analyzed and inform future strategy | Qualitative | 4 |
| **10. Adaptive Learning** | Frequency of Strategic Adaptations | Number of documented strategic pivots or adaptations made per year based on new insights or feedback | Quantitative | 5 |
| | Implementation of Learning Mechanisms | Number of formal knowledge-sharing mechanisms (e.g., learning sessions, cross-case reviews) established | Quantitative | 4 |
| | Rate of Methodology/Framework Revisions | Number of times impact measurement or investment frameworks are revised based on new learning | Quantitative | 3 |
| | Evidence of Knowledge Integration Across Portfolio | Percentage of portfolio companies or projects that have adopted new practices based on shared learning | Quantitative | 3 |
| | Documentation of Learning from Failure and Success | Qualitative evidence that both positive and negative outcomes are systematically analyzed and inform future strategy | Qualitative | 6 |

| | | | | |
|---|---|---|---|---|
| | Integration of Stakeholder and Community Feedback | Evidence that stakeholder or community feedback is directly used to adjust strategy or operations | Qualitative | 5 |
| | Learning Loops Between Organizational Units | Evidence of knowledge exchange and feedback between different divisions, entities, or partners | Qualitative | 3 |
| | Openness to External Learning and Benchmarking | Evidence of incorporating external lessons, peer feedback, or benchmarking into strategy | Qualitative | 4 |
| **11. Systems-based Techniques** | Application Frequency of Systems Tools | Number of formal systems-based analytical techniques (e.g., systems mapping, network analysis, causal loop diagrams) applied per project | Quantitative | 9 |
| | Depth of Systems Analysis in Decision-making | Percentage of investment or project decisions explicitly informed by systems-based techniques | Quantitative | 6 |
| | Identification of Leverage Points | Number of leverage points or intervention nodes identified and prioritized using systems techniques | Quantitative | 5 |
| | Integration of Systems Tools Across Portfolio | Percentage of portfolio projects or investments utilizing at least one formal systems-based technique | Quantitative | 4 |
| | Sophistication of Systems Techniques Employed | Qualitative assessment of the rigor, complexity, and appropriateness of systems-based techniques used (e.g., depth of systems mapping, use of feedback analysis, scenario modeling) | Qualitative | 9 |
| | Customization and Adaptation of Tools | Evidence of adapting or customizing systems-based techniques to local context, sector, or problem type | Qualitative | 5 |
| | Integration in Strategic Planning | Extent to which systems-based techniques are embedded in strategic planning processes (not just analysis) | Qualitative | 4 |
| | Capacity Building in Systems Approaches | Evidence of training or capacity-building efforts to enable teams and partners to use systems-based techniques | Qualitative | 3 |
| **12. Multicapital Circulation** | Diversity of Capital Types Deployed | Number of distinct capital types (e.g., grants, equity, debt, technical assistance, social, natural) used per initiative | Quantitative | 13 |
| | Blended Finance Structures Created | Number of blended finance vehicles or structures (e.g., combining grants and investments) established | Quantitative | 8 |
| | Frequency of Capital Type Integration | Percentage of portfolio projects or investments using more than one capital type | Quantitative | 5 |
| | Volume of Non-financial Capital Mobilized | Value or number of non-financial capital contributions (e.g., technical assistance, knowledge, network support) | Quantitative | 5 |
| | Strategic Alignment of Capital Types | Evidence that multiple capital types are intentionally aligned to address different leverage points or system needs | Qualitative | 5 |
| | Flexibility and Responsiveness of Capital | Qualitative evidence of adapting capital mix in response to evolving needs or opportunities | Qualitative | 5 |

| | | | | |
|---|---|---|---|---|
| | Integration of Non-financial Capital | Depth and quality of technical assistance, knowledge sharing, or network support provided alongside financial capital | Qualitative | 5 |
| | Democratization and Accessibility of Capital | Evidence of democratized capital access (e.g., crowdfunding, community ownership, participatory grantmaking) | Qualitative | 5 |
| **13. Capital Catalyzation** | Capital Leverage & Mobilization | Capital leverage ratio (additional capital catalyzed per dollar invested) | Quantitative | 13 |
| | Follow-on/Multiplier Effect | Ratio of follow-on investment to initial catalytic capital | Quantitative | 10 |
| | | Number of follow-on investments catalyzed | Quantitative | |
| | Ecosystem/Market Expansion | Number of new investors/partners | Quantitative | 7 |
| | | Number of new market entrants | Quantitative | |
| | | Geographic expansion rate (new markets per year) | Quantitative | |
| | Resource Pooling and Partnerships | Amount of additional funding catalyzed via pooled/matched funds or partnerships | Quantitative | 5 |
| | Long-term Impact | Ratio of initial investment to long-term capital generated | Quantitative | 3 |
| | Value Redistribution | Percentage/value of returns redistributed to stakeholders/data creators | Quantitative | 1 |
| | Technology/Non-financial Resource Catalyzation | Value of resources catalyzed through partnerships relative to direct investment | Qualitative | 1 |
| | Strategic/Field Influence | Evidence of improved investment terms or field influence (e.g., more favorable terms, advocacy impact) | Qualitative | 2 |
| | Systemic/Loop Creation | Evidence of reinforcing capital loops (qualitative description of how initial capital creates cycles of additional resource flows) | Qualitative | 1 |
| | Systemic Intent | Evidence of catalytic intent in investment design (e.g., explicit system change pathways, complementary actions) | Qualitative | 2 |
| | Non-financial Resource Mobilization | Quality and extent of non-financial support mobilized alongside financial capital | Qualitative | 2 |
| | Innovative Mechanisms/Approaches | Development or diversity of catalytic capital mechanisms (e.g., matching funds, innovative financial mechanisms, capacity grants) | Qualitative | 2 |
| | Narrative/Perception Shift | Evidence of narrative or perception change among mainstream investors or stakeholders | Qualitative | 2 |

# 6. Leveraging AI to Support Systemic Investing Evaluation (Responding to RQ3)

This chapter also responds to the third research question:
**RQ3: How can we move toward systemic investing evaluation in practice?**

Building on the indicator system introduced in the previous chapter, this section explores how artificial intelligence (AI) can support the practical implementation of systemic investing evaluation. In particular, it presents an AI-assisted scoring tool designed to operationalize the 13 hallmarks and their associated indicators through automated document analysis. This tool aims to increase evaluation consistency and scalability.

## 6.1 Introduction

Artificial intelligence (AI), with its ability to process large volumes of unstructured information, identify patterns, and adapt to emerging contexts, presents a promising complement to human judgment in the field of systemic evaluation. Systemic problems are inherently complex. They are characterized by interdependent variables, non-linear dynamics, feedback loops, and adaptive behaviors that evolve over time. Traditional evaluation tools often struggle to capture such complexity, particularly when assessing the systemic nature of impact-oriented initiatives.

I developed an AI-based scoring tool to enhance the scalability of systemic investing evaluations. Details of the model and prompting are in the Appendix below. It is grounded in a structured framework of 13 hallmarks, which articulate key characteristics of systemic change such as systems thinking, long-term orientation, stakeholder engagement, and cross-sector collaboration. To operationalize these hallmarks, a library of 112 synthesized evaluation indicators was constructed, capturing both quantitative and qualitative dimensions of systemic change across multiple domains including policies, practices, resource flows, power dynamics, relationships, and mental models (as stated by the table of indicator system in Chapter 5).

The methodology behind the tool involves pre-training the AI system with contextual foundation. This includes definitions and distinguishing features of each hallmark, relevant theoretical insights from systems thinking literature, and the identified indicator database. With this contextual understanding, just upload materials (such as project documents, interview recordings, or organizational impact reports) and generate hallmark scores. Additionally, it proposes a set of context-appropriate indicators to the specific characteristics and strategies described in the document.

## 6.2 Assumptions Underlying the AI-Based Indicator Assessment

The AI-based evaluation presented in this section is grounded in a set of methodological assumptions.

First, the evaluation assumes that the presence of an indicator, i.e., whether the corresponding concept or practice appears in the written documentation of a case, serves as a proxy for that dimension being actively engaged in the investment process. This presence is inferred through natural language processing techniques applied to case descriptions, reports, or interviews. Absence of an indicator does not necessarily mean the dimension is missing in practice, but rather that it is not sufficiently represented in the available documentation.

Second, the assessment adopts a coverage-based approach: the more indicators present in a given case and the more the context relevant to the hallmark definition, the higher score it is in that hallmark. This reflects an assumption that systemic investing and evaluation are multidimensional in nature, and that broader coverage across these dimensions (as defined by the 13 hallmarks) suggests greater systemic intentionality and execution. At this stage, all indicators are treated as equally weighted, and the analysis does not account for the relative importance, depth, or integration of each indicator.

Third, the method does not currently assess the appropriateness or contextual relevance of each indicator to the specific investment strategy or sector. Instead, it focuses on surface-level alignment: whether the case documentation addresses or reflects the identified features of systemic investing. As such, this initial version of the tool prioritizes breadth over depth.

These assumptions will need to be revisited and tested in future iterations, especially as the framework is applied across a wider range of cases and refined with stakeholder input. Additional layers of evaluation such as assessing quality of implementation, interconnections among indicators, or stakeholder perceptions could complement this initial approach.

## 6.3 AI Scoring sample

| Hallmark | Score (0-10) | Justification | Suggested Indicators |
|---|---|---|---|
| Systems Thinking and Complexity Science | 7.5 | The demonstrates a good understanding of systems thinking by addressing interrelated social and environmental issues through community engagement and resource sharing. However, it lacks comprehensive systems mapping. | - Number of system dimensions/nodes identified; - Number of feedback loops or causal relationships mapped; - Evidence that systems analysis directly informs investment or strategy decisions |
| Paradigm Evolution | 8.0 | The initiative challenges traditional consumption paradigms by promoting borrowing over buying, aligning with new economic models. It shows evidence of shifting mindsets among stakeholders. | - Number of formal changes in mission statement, focus areas, or investment thesis over time; - Percentage of stakeholders reporting a shift in mindset or mental models; - Qualitative evidence of new language or narratives in communications |
| Long-Term Horizon | 6.0 | While the initiative has a long-term vision of community connection and reduced consumption, it lacks explicit long-term planning and metrics to track progress over extended periods. | - Presence of explicit 5+, 10+, or 20+ year plans; - Evidence of explicit long-term vision in strategy documents; - Presence and frequency of long-term outcome metrics in evaluation frameworks |
| Multi-Level Engagement | 7.0 | The engages multiple community levels, but the assessment of impact across these levels could be more robust. The initiative shows potential for broader systemic change through community involvement. | - Number of distinct system levels included in impact assessment; - Percentage of impact metrics that track change at more than one system level; - Existence and clarity of a theory of change that links outcomes across system levels |
| Cross-Sector Collaboration | 6.5 | The initiative collaborates with local councils and community organizations, but the depth and breadth of cross-sector partnerships could be improved to enhance systemic impact. | - Number of formal partnerships involving organizations from at least two distinct sectors; - Number of different sectors engaged in collaboration; - Evidence of co-design and shared goal-setting in collaborative efforts |
| Stakeholder Engagement | 8.5 | The actively engages diverse stakeholder groups, ensuring their voices are heard in decision-making processes, which enhances community ownership and empowerment. | - Number of distinct stakeholder groups actively engaged; - Percentage of key decisions involving direct stakeholder input; - Evidence of co-creation and participatory design in engagement processes |
| System Structure Changes | 5.0 | The initiative does not significantly alter existing structural elements of the system, focusing more on service delivery rather than systemic policy or regulatory changes. | - Number of policy or regulatory changes influenced; - Evidence of fundamental shifts in system architecture or power dynamics; - Degree to which affected stakeholders are involved in designing new structures |
| Promote Policy Change and Adaptability | 6.0 | While the initiative has the potential to influence policy through community engagement, there is limited evidence of direct policy advocacy or adaptability in response to changing contexts. | - Number of policy changes influenced; - Frequency of updates to policy-related frameworks; - Evidence of timing and flexibility in policy advocacy strategies |
| Dynamic Mechanism | 7.0 | The shows some capacity for adaptive learning through community feedback, but formal mechanisms for continuous strategy adjustments are not well established. | - Number of formal strategy or impact assessment review sessions per year; - Evidence of structured feedback mechanisms informing strategy adjustments; - Documentation of strategic pivots made in response to new information |
| Adaptive Learning | 6.5 | The initiative demonstrates some learning from experience, but there is limited evidence of systematic knowledge-sharing mechanisms or capacity-building efforts within the organization. | - Number of formal knowledge-sharing mechanisms established; - Evidence of stakeholder feedback used to adjust strategy; - Documentation of learning from both positive and negative outcomes |
| Systems-based Techniques | 7.0 | The employs some systems-based techniques, but the depth and sophistication of these methods could be enhanced to better inform decision-making and strategy development. | - Number of formal systems-based analytical techniques applied per project; - Percentage of investment decisions informed by systems-based techniques; - Evidence of training or capacity-building efforts to enable teams to use systems-based techniques |
| Multicapital Circulation | 6.0 | The initiative recognizes the importance of multiple forms of capital, but there is limited evidence of intentional deployment and tracking of these capitals beyond financial resources. | - Number of distinct capital types used per initiative; - Evidence of aligning multiple capital types to address system needs; - Quality and extent of non-financial support mobilized alongside financial capital |
| Capital Catalyzation | 5.5 | The has potential for catalyzing additional resources, but there is limited evidence of structured capital mobilization strategies or leveraging additional funding streams. | - Capital leverage ratio (additional capital catalyzed per dollar invested); - Number of follow-on investments catalyzed; - Evidence of improved investment terms or field influence |

*Figure 14: The result: AI scoring sample for a case*

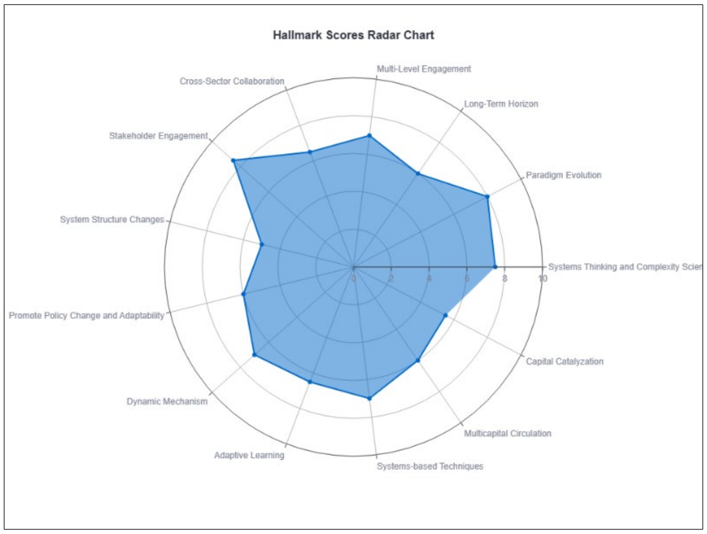


*Figure 15: The Radar Chart for 13 Hallmarks of the case*

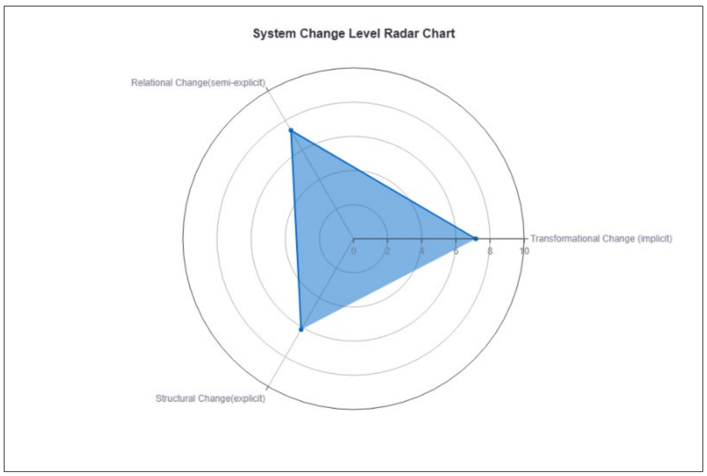


*Figure 16:  The Radar Chart for 3 levels of the case*

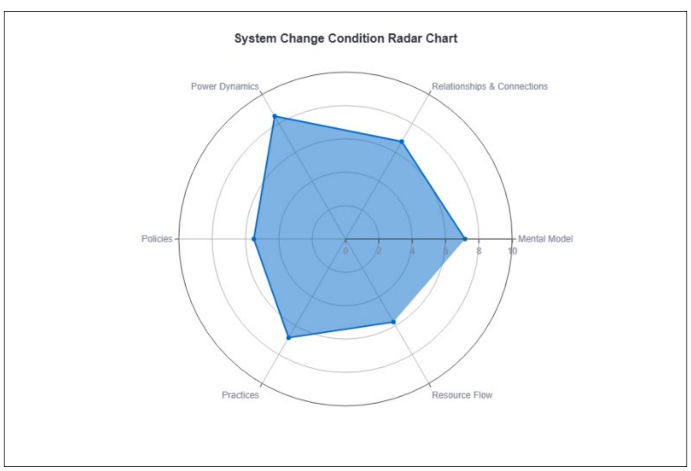


*Figure 17: The Radar Chart for 6 conditions of the case*

## 6.4 AI Scoring tool result for the 46 cases

| | Systems Thinking and Complexity Science | Paradigm Evolution | Long-Term Horizon | Multi-Level Engagement | Cross-Sector Collaboration | Stakeholder Engagement | System Structure Changes | Promote Policy Change and Adaptability | Dynamic Mechanism | Adaptive Learning | Systems-based Techniques | Multicapital Circulation | Capital Catalyzation |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Case 1 | 9.0 | 8.0 | 7.5 | 8.5 | 9.0 | 8.0 | 7.0 | 8.0 | 7.5 | 8.0 | 8.5 | 9.0 | 8.0 |
| Case 2 | 8.5 | 7.0 | 9.0 | 8.0 | 8.5 | 9.0 | 7.5 | 6.5 | 8.0 | 7.0 | 8.0 | 7.5 | 6.0 |
| Case 3 | 7.5 | 8.0 | 6.0 | 7.0 | 6.5 | 8.5 | 5.0 | 6.0 | 7.0 | 6.5 | 7.0 | 5.5 | 6.0 |
| Case 4 | 6.5 | 7.0 | 5.5 | 6.0 | 7.5 | 6.5 | 5.0 | 6.0 | 5.5 | 6.0 | 6.5 | 7.0 | 6.0 |
| Case 5 | 8.5 | 7.0 | 9.0 | 8.0 | 8.5 | 9.0 | 7.5 | 8.0 | 8.0 | 7.5 | 8.0 | 7.0 | 8.0 |
| Case 6 | 8.0 | 7.5 | 9.0 | 6.5 | 7.0 | 8.5 | 7.0 | 6.0 | 7.5 | 8.0 | 7.0 | 6.5 | 7.0 |
| Case 7 | 8.5 | 7.0 | 9.0 | 8.0 | 7.5 | 8.0 | 6.5 | 7.0 | 8.0 | 7.5 | 8.5 | 7.0 | 8.0 |
| Case 8 | 8.5 | 7.0 | 9.0 | 8.0 | 7.5 | 8.0 | 6.5 | 7.0 | 8.0 | 7.5 | 8.5 | 7.0 | 6.0 |
| Case 9 | 8.5 | 7.0 | 9.0 | 8.0 | 8.5 | 9.0 | 7.5 | 8.0 | 7.0 | 8.0 | 8.5 | 7.5 | 8.0 |
| Case 10 | 8.5 | 7.0 | 9.0 | 8.0 | 8.5 | 9.0 | 7.5 | 8.0 | 7.0 | 8.0 | 8.5 | 7.5 | 8.0 |
| Case 11 | 7.5 | 6.0 | 8.0 | 7.0 | 8.0 | 7.5 | 6.5 | 7.0 | 6.0 | 6.5 | 7.0 | 8.0 | 7.5 |
| Case 12 | 8.5 | 7.0 | 6.5 | 8.0 | 9.0 | 8.5 | 7.5 | 7.0 | 6.0 | 6.5 | 7.0 | 7.5 | 8.0 |
| Case 13 | 8.5 | 7.0 | 9.0 | 8.0 | 8.5 | 9.0 | 7.5 | 7.0 | 8.0 | 8.5 | 8.0 | 7.5 | 7.0 |
| Case 14 | 8.5 | 9.0 | 8.0 | 8.5 | 9.0 | 9.0 | 7.5 | 8.0 | 8.0 | 8.5 | 8.0 | 7.5 | 8.0 |
| Case 15 | 9.0 | 8.5 | 9.0 | 8.0 | 9.0 | 9.0 | 8.0 | 8.5 | 8.0 | 8.5 | 9.0 | 8.0 | 8.5 |
| Case 16 | 8.5 | 7.0 | 9.0 | 8.0 | 8.5 | 9.0 | 7.5 | 7.0 | 8.0 | 8.5 | 8.0 | 7.5 | 8.0 |
| Case 17 | 8.5 | 7.0 | 9.0 | 8.0 | 7.5 | 8.5 | 6.5 | 7.0 | 8.0 | 7.5 | 8.0 | 7.0 | 6.0 |
| Case 18 | 8.5 | 9.0 | 8.0 | 7.5 | 8.0 | 9.0 | 8.5 | 8.0 | 7.0 | 8.0 | 8.5 | 7.5 | 8.0 |
| Case 19 | 8.5 | 7.0 | 9.0 | 8.0 | 8.5 | 9.0 | 7.5 | 8.0 | 7.0 | 8.0 | 8.5 | 7.5 | 8.0 |
| Case 20 | 8.0 | 7.5 | 6.5 | 7.0 | 8.0 | 9.0 | 6.0 | 5.5 | 7.0 | 6.5 | 8.0 | 7.5 | 6.0 |
| Case 21 | 8.5 | 7.0 | 9.0 | 8.0 | 8.5 | 9.0 | 7.5 | 7.0 | 8.0 | 8.5 | 8.0 | 7.5 | 8.0 |
| Case 22 | 9.0 | 8.0 | 7.5 | 8.5 | 9.0 | 8.0 | 7.0 | 8.0 | 8.5 | 7.5 | 8.0 | 7.0 | 8.0 |
| Case 23 | 8.5 | 7.0 | 9.0 | 8.0 | 8.5 | 9.0 | 7.5 | 8.0 | 8.0 | 7.5 | 8.0 | 7.0 | 8.0 |
| Case 24 | 8.5 | 7.0 | 6.5 | 8.0 | 9.0 | 8.5 | 7.5 | 6.0 | 7.0 | 7.5 | 8.0 | 6.5 | 7.0 |
| Case 25 | 8.5 | 7.0 | 9.0 | 8.0 | 7.5 | 8.5 | 6.5 | 7.0 | 7.5 | 8.0 | 8.0 | 7.0 | 6.5 |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Case 26** | 8.5 | 7.0 | 9.0 | 8.0 | 8.5 | 9.0 | 7.5 | 8.0 | 8.0 | 7.5 | 8.0 | 7.0 | 8.0 |
| **Case 27** | 8.5 | 7.0 | 9.0 | 8.0 | 7.5 | 8.5 | 7.0 | 6.5 | 8.0 | 7.5 | 8.0 | 7.0 | 6.5 |
| **Case 28** | 8.5 | 7.0 | 9.0 | 8.0 | 7.5 | 8.5 | 7.0 | 6.5 | 8.0 | 7.5 | 8.0 | 7.0 | 6.5 |
| **Case 29** | 8.5 | 8.0 | 9.0 | 7.5 | 8.0 | 9.0 | 7.0 | 7.5 | 8.0 | 8.5 | 8.0 | 7.5 | 8.0 |
| **Case 30** | 8.5 | 9.0 | 7.5 | 8.0 | 7.0 | 8.5 | 6.5 | 7.0 | 8.0 | 7.5 | 8.0 | 6.0 | 7.0 |
| **Case 31** | 9.0 | 8.5 | 9.0 | 8.0 | 9.0 | 8.5 | 7.5 | 8.0 | 8.0 | 8.5 | 9.0 | 8.0 | 8.5 |
| **Case 32** | 8.5 | 7.0 | 9.0 | 8.0 | 8.5 | 9.0 | 7.5 | 7.0 | 8.0 | 8.5 | 8.0 | 9.0 | 8.0 |
| **Case 33** | 8.5 | 9.0 | 7.5 | 8.0 | 8.5 | 9.0 | 7.0 | 8.0 | 7.5 | 8.0 | 8.5 | 7.0 | 8.0 |
| **Case 34** | 8.5 | 9.0 | 8.0 | 7.5 | 8.0 | 9.0 | 7.0 | 7.5 | 8.0 | 8.5 | 8.0 | 7.5 | 8.0 |
| **Case 35** | 9.0 | 8.0 | 9.0 | 8.0 | 7.0 | 9.0 | 8.0 | 7.0 | 8.0 | 8.0 | 9.0 | 8.0 | 7.0 |
| **Case 36** | 8.5 | 7.0 | 9.0 | 8.0 | 8.5 | 9.0 | 8.0 | 8.5 | 7.5 | 8.0 | 8.0 | 7.5 | 8.0 |
| **Case 37** | 9.0 | 8.5 | 9.0 | 8.0 | 9.0 | 9.0 | 8.5 | 8.0 | 8.5 | 8.0 | 9.0 | 8.5 | 9.0 |
| **Case 38** | 8.5 | 7.0 | 9.0 | 8.0 | 8.5 | 9.0 | 7.5 | 8.0 | 7.0 | 8.0 | 8.5 | 7.5 | 8.0 |
| **Case 39** | 8.5 | 7.0 | 9.0 | 8.0 | 8.5 | 9.0 | 7.5 | 8.0 | 8.0 | 7.5 | 8.0 | 8.5 | 9.0 |
| **Case 40** | 8.5 | 7.0 | 9.0 | 8.0 | 8.5 | 9.0 | 7.5 | 8.0 | 7.0 | 6.5 | 8.0 | 7.5 | 8.0 |
| **Case 41** | 8.5 | 7.0 | 9.0 | 8.0 | 6.5 | 8.5 | 7.5 | 7.0 | 8.0 | 7.5 | 8.0 | 6.0 | 7.0 |
| **Case 42** | 8.5 | 7.0 | 6.5 | 8.0 | 9.0 | 8.5 | 7.5 | 7.0 | 6.0 | 6.5 | 8.0 | 7.5 | 7.0 |
| **Case 43** | 8.5 | 7.0 | 9.0 | 8.0 | 7.5 | 8.5 | 7.0 | 6.5 | 7.5 | 8.0 | 8.0 | 7.0 | 7.5 |
| **Case 44** | 8.5 | 9.0 | 8.0 | 8.5 | 9.0 | 9.0 | 8.0 | 8.5 | 8.0 | 8.5 | 8.0 | 8.0 | 8.5 |
| **Case 45** | 7.5 | 8.0 | 6.5 | 7.0 | 6.0 | 7.5 | 5.5 | 6.0 | 6.5 | 7.0 | 6.0 | 5.0 | 6.0 |
| **Case 46** | 8.5 | 9.0 | 8.0 | 8.5 | 7.5 | 9.0 | 8.0 | 8.5 | 7.0 | 8.0 | 8.0 | 7.5 | 8.0 |

*Figure 18: The AI scoring result for 46 cases*

# 7. Applying the Framework to a Case (Builders Vision)

## 7.1 Case Analysis Method and Rationale

This chapter applies the 13-hallmark framework to a real-world case to explore how systemic investing and evaluation manifest in practice. Specifically, I analyze Builders Vision, an impact-driven investment and philanthropy platform with publicly available impact reports.

The case is analyzed in 3 ways:

1. **External assessment of investment practice**: I apply the hallmarks to assess the extent to which Builders Vision's strategies and activities reflect systemic investing characteristics.
2. **Review of internal evaluation system**: I examine how the organization itself approaches evaluation, and whether its methods align with the principles of systemic evaluation (as reflected in the right-hand column of the hallmarks framework).
3. Following this human-led analysis, I also apply the AI-assisted scoring tool to the Builders Vision case documentation and compare the automated evaluation with human judgment to explore alignment and limitations.

*Table 6: Comparison of Section A and B*

| Item | **Section A: Evaluating Investment Practice** | **Section B: Evaluating Evaluation Practice** |
|---|---|---|
| Evaluation Lens | Myself as an external evaluator | Builders Vision as an internal evaluator |
| Object of Evaluation | Its investment practices | Its evaluation system |
| Framework Basis | Column of Hallmarks (Table 4): Implications for Investment Practice | Column of Hallmarks (Table 4): Implications for Evaluation |

## 7.2 Case Analysis

Builders Vision is an impact-driven investment and philanthropy platform dedicated to addressing some of the most pressing environmental challenges. It operates across three interconnected sectors: Oceans, Food & Agriculture, and Energy, leveraging investment, philanthropy, and advocacy to catalyze systemic change. The organization integrates diverse forms of capital, including grants, direct investments, and fund investments, to support innovators and changemakers working towards a more sustainable and equitable future.

*7.2.1 Transformative Level of Systemic Investing - Long-term Horizon (Hallmark 3)*

*"We are committed to long-term partnerships with fund managers and entrepreneurs to help them develop, scale, and succeed."*

- *Builders Vision Impact Report 2023*

*A. Evaluating Investment Practice*

Builders Vision emphasizes extended timeframes for its work and its relationships with partners. It describes a commitment to long-term partnerships with fund managers and entrepreneurs, with the intention of contributing to their development and growth over time. In practice, this involves support for organizations and companies at early or formative stages, even when outcomes are not immediate or fully known.

In addition to individual investments, Builders Vision applies long-term thinking in how it engages with field-level challenges. In the Oceans portfolio, the organization supports not only direct activities such as sustainable aquaculture but also broader field infrastructure, including data tools, community-building efforts, and coordination platforms. The organization describes its work as part of a long-term effort to build a healthier ocean economy and to support sectors that require sustained engagement across multiple actors and time horizons. Rather than seeking short-term outputs, the organization provides patient capital and aims to stay involved over multiple stages of a partner's development.

*B. Evaluating Evaluation Practice*

One way this long-term orientation is expressed is through the articulation of outcome timelines. Builders Vision identifies both near-term and long-term goals to guide its strategic direction. For example, in the Oceans portfolio, it states an objective that by 2030, critical ocean solutions, such as regenerative seaweed farming and sustainable aquaculture, will be "demonstrated and de-risked," with supporting community infrastructure in place to facilitate broader adoption. Looking further ahead, by 2050, the organization envisions a "new blue economy" where healthy ocean systems support human well-being, climate resilience, and sustainable industry. This dual timeframe approach allows Builders Vision to focus on actionable progress in the present while aligning its investments with more systemic and intergenerational change. It also reflects a structured effort to manage complexity by linking intermediate steps to longer-term sector transformation.

*7.2.2 Relational Level of Systemic Investing - Stakeholder Engagement (Hallmark 6)*

*"By asking partners to select impact indicators that will bring the most value to them, we aim to reinforce the link between IMM and strategy, wherein data is used as an input to critical operational decisions"*

- *Builders Vision Impact Report 2023*

***A. Evaluating Investment Practice***

The organization actively involves partners and stakeholders in the development of solutions, strategies, and implementation processes. A clear example is its participation in the 1,000 Ocean Startups coalition, where Builders Vision contributed to the development of the Ocean Impact Navigator. This tool was created collaboratively by stakeholders across the field to help organizations measure impact in the blue economy. Rather than applying a pre-designed measurement approach, the organization worked with others to co-develop a shared resource that could be adapted to different contexts.

Another example of stakeholder engagement is the OceansSummit, where over 130 stakeholders, including grantees, fund managers, portfolio companies, and other key actors, gathered to connect and share perspectives. The event served as a space for different actors to exchange experiences, align on challenges, and explore opportunities for future collaboration. This approach reflects Builders Vision's philosophy expressed in their report: "We stand proudly behind those on the forefront of change. Our community of partners includes nearly 500 grantees, portfolio companies, and public and private fund managers."

In their investment practices, Builders Vision emphasizes meaningful engagement rather than transactional relationships, noting, "Our commitment to them will always go beyond the check. We advocate for them, connect them to our networks, and help them to develop, scale, and succeed." This approach aligns with the hallmark's focus on actively involving stakeholders in strategic decisions rather than treating them as passive recipients.

***B. Evaluating Evaluation Practice***

In terms of evaluation practices, Builders Vision exemplifies the stakeholder engagement hallmark by employing participatory evaluation methods that involve diverse stakeholders throughout the assessment process. As described in their report, in their impact measurement work, "organizations are asked to select the indicators that they find most valuable." This approach is explicitly intended to support decision-making processes within partner organizations and to ensure that measurement is relevant to their goals and operating context. Their evaluation approach is designed to create opportunities for beneficiaries and affected communities to define what success looks like and interpret findings.

Builders Vision has also built feedback mechanisms that give voice to a diverse range of stakeholders. The organization notes that they "regularly examining our choices, asking ourselves what is working and reflecting on the intended and unintended outcomes, so we can apply these learnings to future opportunities."

### ***7.2.3 Structural Level Systemic Investing - Capital Catalyzation (Hallmark 12)***

*“Builders Initiative saw the company’s potential even at its early stage. This led to subsequent funding and growth.”*

- *Builders Vision Impact Report 2023*

Builders Vision provides capital with the goal of not only supporting specific companies or funds, but also encouraging broader flows of investment into areas it sees as important. In multiple cases, early investments were followed by additional rounds of funding from other actors.

Capital is deployed not only to companies but also to investment funds and field-building initiatives. For example, the organization supports fund managers aligned with its focus areas and values. The relationship is described as extending beyond financial transactions. Builders Vision offers non-financial support such as strategic input, access to networks, and operational advice. Fund managers identified convenings, introductions to capital partners, and operational guidance as among the most helpful forms of support. These forms of engagement aim to strengthen the capacity of fund managers and portfolio companies, potentially increasing the flow of capital into the targeted sectors.

In the Oceans portfolio, the organization also supports collaborative initiatives that seek to improve coordination and data access in the field. These activities are intended to address barriers to investment and participation in ocean-related sectors. Through this combination of early capital, field-building support, and collaborative platforms, Builders Vision contributes to conditions that may encourage further investment from other sources.

One example of this approach is Builders Vision's early-stage support for Matter, a UK-based startup developing technology to prevent microplastics from entering wastewater through household and commercial laundry systems. Builders Initiative's investment team first learned about Matter in 2020 and recognizing the potential of its innovation at an early stage, made an equity investment through its impact LLC in 2021. This initial investment supported the company in continuing to develop its product, strengthening its internal operations, and preparing for broader market entry. Following this early support, additional investors began to engage. In 2022, S2G Ventures — another entity within the Builders Vision platform — provided a $1 million convertible note investment to help Matter advance toward its Series A funding round. This step allowed S2G to better understand the company and assess its growth potential. By 2023, S2G Ventures co-led Matter's Series A, which totaled $10 million and was raised from a syndicate of market-rate investors. This sequence illustrates how early capital, combined with relationship-building and internal coordination, can help position emerging companies to access larger, more conventional sources of investment.

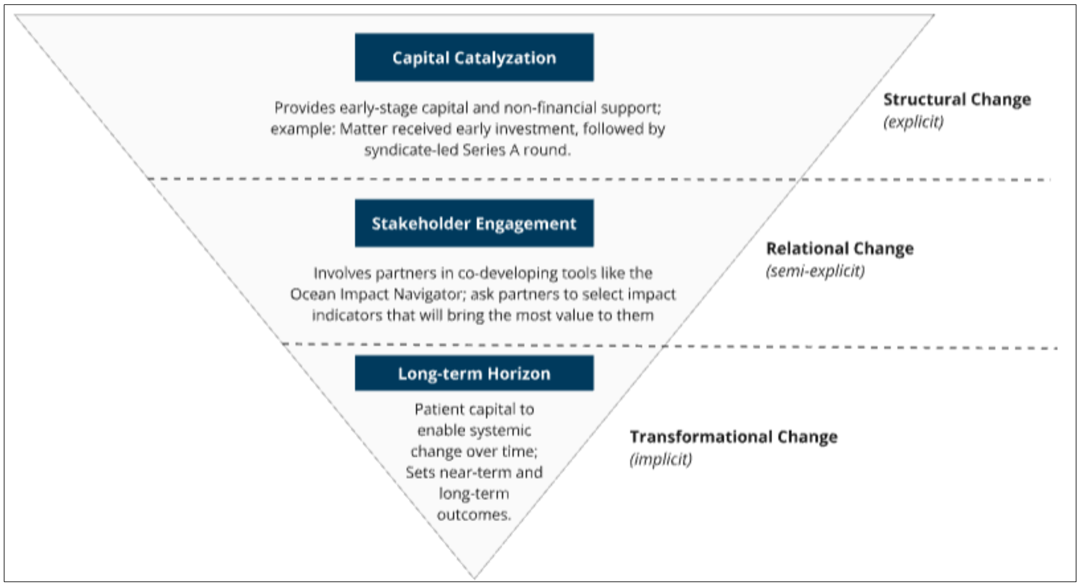


*Figure 19: Hallmark analysis for Builders Vision*

## 7.3 AI scoring tool result for Builders Vision

*Table 7: AI scoring tool result for Builders Vision*

| No. | **Hallmarks** | **Score (0-10)** | **Rationale for the Score** | **Suggested Indicator** |
|---|---|---|---|---|
| 1 | Systems Thinking and Complexity Science | 9 | BV demonstrates strong systems thinking in combining philanthropy and investment across food, oceans, and energy. | 1.1 Number of system dimensions/nodes identified<br>1.2 Number of feedback loops or causal relationships mapped |
| 2 | Paradigm Evolution | 8.5 | Clear evidence of paradigm evolution with explicit rejection of short-termism in favor of ecosystem-level value creation. | 2.1 Percentage of stakeholders (staff, partners, community) reporting a shift in mindset or mental models<br>2.2 Depth and diversity of stakeholder involvement in co-creating new paradigms |
| 3 | Long-Term Horizon | 9.2 | Long-term orientation is embedded in portfolio strategies and multi-decade impact aspirations. | 3.1 Number of years encompassed in strategic planning<br>3.2 Presence of explicit 5+, 10+, or 20+ year plans |
| 4 | Multi-Level Assessment | 8.8 | Impact data collected and synthesized across grantees, funds, and companies reflects multi-level assessment. | 4.1 Number of distinct system levels (e.g., individual, organizational, ecosystem, policy) included in impact assessment<br>4.2 Percentage of impact metrics/KPIs that explicitly track change at more than one system level |
| 5 | Cross-Sector Collaboration | 9 | Cross-sector collaboration is evident in coalitions like 1,000 Ocean Startups and partnership with UN bodies. | 5.1 Number of formal partnerships involving organizations from at least two distinct sectors;<br>5.2 Number of different sectors (e.g., business, government, civil society, academia) engaged in collaboration |
| 6 | Stakeholder Engagement | 9.5 | Robust engagement with 500+ partners through strategic support, technical assistance, and convenings. | 6.1 Number of distinct stakeholder groups actively engaged (e.g., by sector, demographic, role)<br>6.2 Percentage of key decisions (e.g., investment, program design) involving direct stakeholder input |
| 7 | System Structure Changes | 8.7 | Initiatives contribute to changing sectoral systems, including regulatory frameworks and incentive models. | 7.1 Number of policy, regulatory, or legislative changes directly attributed to the initiative or investment<br>7.2 Number of new governance bodies, platforms, or market mechanisms established |

| | | | | |
|---|---|---|---|---|
| 8 | Promote Policy Change and Adaptability | 8.4 | Policy adaptability is evident in funder collaborations and flexible use of philanthropic capital to shape policy contexts. | 8.1 Number of policy, legislative, or regulatory changes directly influenced or enacted<br>8.2 Number of policy advocacy, engagement, or consultation activities conducted per year |
| 9 | Dynamic Assessment | 9.1 | Annual impact reporting, iteration on IMM strategy, and data-driven portfolio adjustments suggest dynamic assessment. | 9.1 Number of formal strategy or impact assessment review sessions per year;<br>9.2 Percentage of KPIs or impact metrics updated or refined annually based on new data or learning |
| 10 | Adaptive Learning | 9.3 | Evidence of learning loops, including real-time updates to frameworks and responsive IMM adaptations. | 10.1 Number of documented strategic pivots or adaptations made per year based on new insights or feedback<br>10.2 Number of formal knowledge-sharing mechanisms (e.g., learning sessions, cross-case reviews) established |
| 11 | Systems-based Techniques | 8.9 | Use of tools like Ocean Impact Navigator indicates application of systems-based techniques. | 11.1 Number of formal systems-based analytical techniques (e.g., systems mapping, network analysis, causal loop diagrams) applied per project<br>11.2 Percentage of investment or project decisions explicitly informed by systems-based techniques |
| 12 | Multicapital Circulation | 8.6 | Capital allocation strategy spans grants, equity, debt, and catalytic funds to circulate multiple forms of capital. | 12.1 Number of distinct capital types (e.g., grants, equity, debt, technical assistance, social, natural) used per initiative<br>12.2 Number of blended finance vehicles or structures (e.g., combining grants and investments) established |
| 13 | Capital Catalyzation | 9 | Demonstrates catalytic function by anchoring blended finance vehicles and supporting scale-up investments. | 13.1 Capital leverage ratio (additional capital catalyzed per dollar invested)<br>13.2 Ratio of follow-on investment to initial catalytic capital |
| | Total Average | 8.92 | | |

## 7.4 Comparative Analysis of Human Evaluation and AI Assessment

After presenting the AI-assisted scoring tool results for Builders Vision, we can do a comparative analysis between the human evaluation conducted across selected hallmarks and the comprehensive assessment provided by the AI tool.

The AI-based assessment largely aligns with the human analysis in identifying Builders Vision's core strengths. The tool assigned particularly high scores to Long-term Horizon (9.2/10), Stakeholder Engagement (9.5/10), and Capital Catalyzation (9.0/10), which mirrors the qualitative analysis that highlighted these as areas where Builders Vision demonstrates exemplary practice.

For Long-term Horizon, both approaches recognized Builders Vision's explicit articulation of 2030 and 2050 goals and its patient capital approach. Similarly, with Stakeholder Engagement, both the AI assessment and human recognized the breadth of stakeholder involvement.

# 8. Future Directions for Research and Practice

This chapter outlines future directions emerging from this thesis, organized into three parts. Section 8.1 summarizes the key reflections for each of the three guiding research questions. Section 8.2 discusses potential extensions for academic research. Section 8.3 offers reflections on how the frameworks and tools developed here might be applied or adapted in practical settings. Together, these sections highlight that systemic investing evaluation is not a fixed model, but an evolving approach that invites ongoing exploration.

## 8.1 Reflections by Research Question

The first research question asked: *“Why shift away from traditional evaluation approaches for systemic investing?”*. In response, this thesis reviewed literature, finding that conventional evaluation methods often grounded in assumptions of linear causality, discrete interventions, and time-bound results, which are poorly suited to assessing the effects of systemic investing. Systemic interventions frequently operate across multiple domains, rely on indirect and emergent pathways to change, and unfold over extended periods. Traditional methods struggle to capture such complexity and tend to prioritize measurable outputs over deeper system shifts. As a result, this study supports a broader transition toward complexity-aware, contribution-based, and learning-oriented evaluation approaches that are better aligned with the nature of systems change.

The second research question explored: “What does systemic investing evaluation look like and what are its characteristics?”. To answer this, the thesis introduced a framework of thirteen hallmarks. These hallmarks synthesize insights from systems thinking, evaluation theory, and real-world investing practices, offering a view of what distinguishes systemic evaluation from conventional approaches. Rather than prescriptive metrics or checklists, the hallmarks are conceptual reference points, and highlight such features as long-term timeframes, adaptive learning processes, cross-sector collaboration, power-aware stakeholder engagement, and the intentional flow of multiple forms of capital. By identifying these patterns, the framework helps practitioners and evaluators determine whether and how their work reflects transformational intent.

The third research question asked: “How can practitioners begin to adopt and implement systemic investing evaluation in practice?” The thesis responded by developing two practical tools. First, a set of 112 qualitative and quantitative indicators was constructed by analyzing 46 cases from the TWIST and FEST networks. These indicators make the hallmarks observable and assessable. Second, an AI-based scoring tool was introduced to support the application of the framework to textual materials such as reports, memos, and case studies. This tool enables faster, more consistent assessments across projects, while allowing for context-aware suggestions tailored to specific investment goals. These tools, taken together, offer a methodologically grounded path from conceptual framing to practical use.

## 8.2 Directions for Future Research

Future research could strengthen the foundation and expand the scope of systemic investing evaluation in several ways. One key direction involves longitudinal studies that track systemic investing initiatives over time. Such studies would offer insights into how evaluation practices evolve and how early indicators correlate with long-term systems change. By capturing the unfolding of transformation processes, researchers could better understand which evaluative signals are most predictive or meaningful over different time horizons.

Another important area for inquiry is comparative analysis. Evaluations that examine traditional and systemic approaches within the same investment context could help clarify their respective strengths and limitations. This could also inform decisions about when systemic evaluation is most useful, and what trade-offs exist in terms of feasibility, stakeholder engagement, and learning value.

There is also room to expand the role of technology in systemic evaluation. The AI tool introduced in this thesis represents an early prototype for automating evaluative analysis. Future work could explore more advanced machine learning models, natural language processing techniques, and system mapping algorithms that help identify causal patterns, feedback dynamics, or structural shifts. Such tools might support real-time monitoring or help detect weak signals of change in large, unstructured datasets.

Lastly, the general framework of hallmarks and indicators could benefit from sector-specific adaptation. While this thesis offers general analysis, tailoring the framework to domains such as energy transitions, food systems, healthcare, or education would increase its practical relevance. Sectoral adaptations would allow for greater alignment with domain-specific language, regulatory environments, and impact pathways, which makes the framework more actionable for specialized investors and policymakers.

## 8.3 Directions for Practice

Beyond research, there are several ways the findings of this thesis can inform practice. Investors can begin by learning exercise and applying the hallmarks and selected indicators during investment strategy design, due diligence, or portfolio reviews. Rather than attempting to quantify all dimensions, they may use the framework to surface questions about system dynamics, leverage points, and systemic intent, prompting more reflective capital deployment decisions.

For philanthropic funders, the framework offers a way to guide grantee selection, structure learning-oriented grant conditions, or assess alignment with long-term system change goals. It also provides language and tools that can help funders move beyond project-level accountability and toward more adaptive, field-level stewardship.

Field builders, including networks and intermediary organizations, may find the indicators and AI tool useful for cross-case comparison, community learning, or the design of evaluation support platforms. These tools could facilitate the emergence of shared evaluation norms, create incentives for experimentation, and enable funders and practitioners to align around more coherent understandings of systemic impact.

In all of these applications, systemic evaluation should be seen not as a fixed standard, but as a reflective practice. Its purpose is not only to measure impact, but also to enhance the capacity of organizations and networks to learn, adapt, and contribute to deeper structural transformation. This thesis aims to offer a starting point for that work.

# 9. Conclusion

This research explored how systemic investing can be meaningfully evaluated, addressing the intersection of systems thinking, evaluation practice, and capital deployment. Through the examination of literature and case studies, the thesis has developed a 13 Hallmarks Framework, which provides an approach to understanding systemic investing evaluation across 6 conditions for systems change. The indicator system derived from 46 real-world cases offers practitioners a set of metrics and qualitative assessment approaches with 112 indicators to implement this framework. The application of the framework to the Builders Vision case with AI scoring tool provide one possibility for practical uses.

Several implications emerge from this work. First, systemic evaluation should be embedded as an ongoing, learning-oriented process, rather than as a bounded exercise. The complexity of the systems involved requires flexible feedback mechanisms that inform decision-making in real time. Second, capacity-building around systems thinking remains essential, not just for evaluators but also for investors and implementing partners. Tools such as systems mapping, feedback loop analysis, and dynamic scenario planning should be more widely understood and integrated into practice. Third, while the indicator framework offers structure, it is not intended as a checklist. Also, the hallmarks and associated metrics should serve as prompts for reflection, dialogue, and continuous improvement. Practitioners are encouraged to adapt them thoughtfully to local context and evolving realities. Finally, the AI tool developed in this research demonstrates the potential for augmenting human judgment through technology, particularly in rapidly synthesizing complex documents and generating relevant evaluative insights. However, such tools should be used combined with human interpretation, particularly when addressing qualitative dimensions such as mental models and power dynamics.

In closing, this research does not offer a definitive answer to how systems change should be evaluated, it proposes an evolving framework. As systemic investing continues to evolve, the evaluation should also move toward methods that embrace complexity, support learning, and enhance our collective capacity to create transformative change.

# AI Use Disclosure

This thesis was written by the author with selective assistance from ChatGPT (OpenAI GPT-4o and GPT-4o Mini, accessed throughout April and early May 2025). Research design decisions, analytical interpretations, and final wording were conducted and verified by the author. The AI model was used to support writing tasks, including:

- Refining paragraph clarity and consistency of language in headings across chapters.
- Assistance with qualitative data analysis from case studies, as described in section 2.4.3.
- The development of AI Scoring Tool in Chapter 6 (described in the Appendix below).

# Appendix

## Appendix 1 The prompt behind AI Tool in generating result

In this study, I used Cursor to generate code for the AI tool and integrated GPT-4o-mini via API.

Also, the evaluation prompt consists of two components: a system message and a user message. The system message defines the model’s role as a “systemic investing evaluation expert” and instructs it to assess cases based on 13 predefined hallmarks, assigning a score (0–10), providing justifications, and suggesting relevant indicators for each hallmark. The output is required in a structured JSON format. The user message includes two parameters: the hallmarks (serving as the evaluation framework) and the user_doc (the investment case to be evaluated). This prompt structure increases consistency in generating standardized evaluations across multiple cases. The following prompt was used to generate evaluations of investment cases based on the above framework:

{"role": "system", "content": '''You are a professional systemic investing evaluation expert. Your task is to assess investment cases based on the 13 hallmarks framework.

IMPORTANT: You MUST return your response in the following JSON format:

```
{
  "table": "A markdown formatted table with columns: Hallmark, Score (0-10), Justification, Suggested Indicators",
  "overall_score": "The average score as a number with one decimal point",
  "scores": {
    "Systems Thinking and Complexity Science": score1,
    "Paradigm Evolution": score2,
    ...
  }
}
```

For each hallmark, you should:

1. Provide a score from 0 to 10 (with one decimal point)

2. Give a brief justification for the rating

3. Suggest relevant indicators from the provided indicator set

The table should be formatted in markdown with clear column headers.

The overall_score should be a number with one decimal point.

The scores object should use the full Hallmark title as the key, matching the Hallmark column in the table, and contain all 13 hallmark scores as numbers.

Ensure your evaluation is thorough, objective, and well-justified.'''} ,

{"role": "user", "content": """Please evaluate the following case using the provided framework:

Evaluation Framework (Hallmarks):

{hallmarks}

Case Document:

{user_doc}"""

}

## Bibliography


Balakrishnan, S., Krehbiel, M., & Lee, Y. (2024). Better Philanthropy for Systems Change. Retrieved from https://www.omidyarnetwork.in/publications/better-philanthropy-for-systems-change

Beer, T. (2017). Systems Change Evaluation Forum: Executive Summary. Center for Evaluation Innovation.

Bouri, A., Mudaliar, A., Schiff, H., Bass, R., & Dithrich, H. (2018). Roadmap for the future of impact investing: Reshaping financial markets. Global Impact Investing Network.

Buckton, S. J., Fazey, I., Ball, P., Ofir, Z., Colvin, J., Darby, M., Hejnowicz, A. P., Leicester, G., Newman, R., Page, G. G., Parsons, K., & van Mierlo, B. (2024). Twelve principles for transformation-focused evaluation. University of York.

Bustamante, A., Tebbens, R. J. D., Lahey, A., & Teddlie, C. (2021). Systems thinking and complexity science informed evaluation frameworks. New Directions for Evaluation, 2021(170), 57–70. https://doi.org/10.1002/ev.20475

Byrne, D, Callaghan, G, 2014, Complexity theory and the social sciences: The state of the art, Oxford: Routledge

Cabaj, M. (2019). Evaluating systems change results: An inquiry framework. Tamarack Institute.

Carney, M. (2015). Breaking the tragedy of the horizon – climate change and financial stability. Speech given at Lloyd's of London. Retrieved from https://www.bankofengland.co.uk/speech/2015/breaking-the-tragedy-of-the-horizon-climate-change-and-financial-stability

Chambers, R. (2017). Can we know better? Reflections for development. Practical Action Publishing.

Cursor. (2024). *Cursor code editor* [Software]. https://www.cursor.sh

Daggers, J., Hannant, A., & Jay, J. (2023, September 13). Systemic investing for social change. Stanford Social Innovation Review. https://ssir.org/articles/entry/systemic_investing_for_social_change

Davies, J., & Goldie, M. (2023). Navigating system change evaluation: A white paper. Social Finance. https://www.socialfinance.org.uk/assets/documents/Navigating-system-change-evaluation.pdf

Ebrahim, A., & Rangan, V. K. (2014). What impact? A framework for measuring the scale and scope of social performance. California Management Review, 56(3), 118-141. https://doi.org/10.1525/cmr.2014.56.3.118

Ebrahim, A., Battilana, J., & Mair, J. (2014). The governance of social enterprises: Mission drift and accountability challenges in hybrid organizations. Research in Organizational Behavior, 34, 81-100. https://doi.org/10.1016/j.riob.2014.09.001

Financier Worldwide. (2020, March). The emergence of impact investing. Financier Worldwide. https://www.financierworldwide.com/the-emergence-of-impact-investing

Gates, E. F., & Fils-Aime, F. (2021). System change evaluation: Insights from The Rippel Foundation and its ReThink Health initiative. New Directions for Evaluation, 2021, 125-138.

Gates, E. F., Teasdale, R. M., Shim, C., & Hubacz, H. (2024). Whose and what values? Advancing and illustrating explicit specification of evaluative criteria in education. Studies in Educational Evaluation, 81, 101335. https://doi.org/10.1016/j.stueduc.2024.101335

Gates, E., & Schwandt, T. (2021). An Emerging Alternative Frame for Evaluating. In Evaluating and Valuing in Social Research (pp. 113-139).

Guba, E. G., & Lincoln, Y. S. (1989). Fourth generation evaluation. Newbury Park, CA: Sage Publications.

Hargreaves, M. B., & Podems, D. (2012). Advancing systems thinking in evaluation: A review of four publications. American Journal of Evaluation, 33(3), 462–470. https://doi.org/10.1177/1098214011435409

Hofstetter, D. (2020). Transformation capital: Systemic investing for sustainability. EIT Climate-KIC. https://transformation.capital

House, E. R. (2015). Evaluating with validity. Sage.

Impact Management Project. (2018). A guide to classifying the impact of an investment. Impact Management Project.

Ivashina, V., & Lerner, J. (2019). Patient Capital: The Challenges and Promises of Long-Term Investing. Princeton, NJ: Princeton University Press.

Jay, J. (2024). Presentation at Katapult Future Fest.

Kania, J., & Kramer, M. (2011). Collective impact. Stanford Social Innovation Review. https://ssir.org/articles/entry/social_progress_through_collective_impact

Kania, J., Kramer, M., & Senge, P. (2018). The water of systems change. https://www.fsg.org/publications/water_of_systems_change

Korijn, A., & Fort, A. (2024). Emergence and co-creation: A collective's perspective on investing for systems change. TWIST: Together We Invest for Systems Transformation. https://www.socialfinance.org.uk/assets/documents/Navigating-system-change-evaluation.pdf

Kshetri, N. (2018). Blockchain's roles in meeting key supply chain management objectives. International Journal of Information Management, 39, 80-89. https://doi.org/10.1016/j.ijinfomgt.2017.12.005

Meadows, D. H. (1999). Leverage Points: Places to Intervene in a System. Hartland, VT: The Sustainability Institute. Retrieved from https://donellameadows.org/archives/leverage-points-places-to-intervene-in-a-system/

Meadows, D. H. (2008). Thinking in systems: A primer (D. Wright, Ed.). Chelsea Green Publishing.

Midgley G. (2000). Systemic Intervention: Philosophy, Methodology and Practice. Kluwer Academic/Plenum Publishers: New York.

Nicholls, J., Lawlor, E., Neitzert, E., & Goodspeed, T. (2012). A guide to social return on investment. The SROI Network.

OpenAI. (2025). *ChatGPT* (GPT-4o and GPT-4o Mini version) [Large language model]. Accessed throughout April and early May 2025. https://chat.openai.com. Used for outlining, language refinement, and developing AI scoring outputs in this thesis.

Patton, M. Q. (2011). Developmental evaluation: Applying complexity concepts to enhance innovation and use. The Guilford Press.

Patton, M. Q. (2016). What is essential in developmental evaluation? On integrity, fidelity, adultery, abstinence, impotence, long-term commitment, and sensitivity in implementing evaluation models. American Journal of Evaluation, 37(2), 250–265. https://doi.org/10.1177/1098214015611744

Patton, M. Q. (2020). Evaluation criteria for evaluating transformation: Implications for the coronavirus pandemic and the global climate emergency. American Journal of Evaluation, 42(1), 53–89. https://doi.org/10.1177/1098214020933689

Picciotto, R. (2020). Towards a Complexity Framework for Transformative Evaluation. Journal of MultiDisciplinary Evaluation, 16(35), 54-76.

Sachs, J. D., Schmidt-Traub, G., Mazzucato, M., Messner, D., Nakicenovic, N., & Rockström, J. (2019). Six transformations to achieve the Sustainable Development Goals. Nature Sustainability, 2(9), 805-814. https://doi.org/10.1038/s41893-019-0352-9

Schot, J. W., Benedetti del Rio, R., Steinmueller, W. E., & Keesman, S. J. (2022). Transformative investment in sustainability: An investment philosophy for the second deep transition. Deep Transitions. https://www.transformativeinvestment.net

Scriven, M. (1991). Evaluation thesaurus (4th ed.). Sage.

Stubbs, W. (2017). Sustainable entrepreneurship and B corps. Business Strategy and the Environment, 26(3), 331-344. https://doi.org/10.1002/bse.1920

Teasdale, R. M. (2021). Evaluative criteria: An integrated model of domains and sources. American Journal of Evaluation, 42(3), 354–376. https://doi.org/10.1177/1098214020955226

The Investment Integration Project. (2024). System-level investing: Case studies of investors leading the way. https://tiiproject.com

TransCap Initiative. (2024). Definition and hallmarks of systemic investing: Version 1.0 (before public consultation). https://transformation.capital

Trujillo, D. (2018). Cross-sector partnerships for systemic change: A systematic literature review and agenda for further research. Journal of Business Ethics, 150(2), 303–313. https://doi.org/10.1007/s10551-018-3859-4

Walton, M. (2016). Setting the context for using complexity theory in evaluation: Boundaries, governance and utilisation. Evidence & Policy, 12(1), 73–89. https://doi.org/10.1332/174426415X14447331264518

Walton, M., Gates, E. F., & Vidueira, P. (2021). Insights and future directions for systems and complexity-informed evaluation. New Directions for Evaluation, 2021, 159-171.

White, H. (2010). A contribution to current debates in impact evaluation. Evaluation, 16(2), 153-164. https://doi.org/10.1177/1356389010361562

Yau, A., & Jay, J. (2024). Systemic investing to tackle the US food waste challenge: The Fink Family and ReFED. MIT Sloan Sustainability Initiative. https://ssrn.com/abstract=4615351