

Final Report

Towards a Unified Impact Measurement: Leveraging AI to Standardise Metrics for Enterprises, Investors & Governments

Authors:

Abdullah Nazeri, Luiza Argenta, Morgane Bascle, Payel Barua

*In collaboration with Impact Investing Solutions AG, with the support of Sveta Banerjee,
Puru Rohilla, and David Martinez*

Faculty Lead: Rémi Viné

12 May 2026

Table of contents

I. Executive summary	4
II. Introduction	5
III. Literature review	7
3.1 Impact Investing	7
3.1.1 <i>Impact Investing as an evolving but conceptually ambiguous field</i>	7
3.1.2 <i>Structural mismatches in incentives and capacity</i>	8
3.2 Impact Measurement tools and frameworks	9
3.2.1 <i>The rise of measurement as a source of legitimacy</i>	9
3.3 AI and Impact Measurement harmonisation	11
3.3.1 <i>Why harmonisation remains limited and contested</i>	11
3.3.2 <i>AI and harmonisation: opportunities and limits</i>	12
IV. Empirical Contributions and Interview Methodology	14
4.1 Empirical contributions	14
4.2 Interview Methodology	14
V. Discussion	17
5.1 Comparison between standards	17
5.2 Visualising the landscape	18
5.3 The key barrier: moving from the technical to the motivational	20
5.3.1 <i>Impact measurement's role in practice</i>	20
5.3.2 <i>The stability of the equilibrium</i>	21
5.3.3 <i>What this means for the harmonisation agenda</i>	22
5.4 The role of AI in the Impact Investment ecosystem	23
5.4.1 <i>AI and the limits of harmonisation</i>	23
5.4.2 <i>Practitioner openness to AI</i>	23
5.4.3 <i>What AI cannot do - structural limits</i>	24
5.4.4 <i>What AI can realistically do</i>	25
VI. Recommendations	27
6.1 Strategic Proposal: AI as an operational support layer	27
6.1.1 <i>The necessary characteristics</i>	27
6.1.2 <i>Implementation roadmap</i>	28
VII. Conclusion	30
VIII. Bibliography	31
IX. Annexes	37
Annex 1: Main impact and sustainability standards and their characteristics	37
Annex 2: Methodology	39
Annex 3: Interview guiding questions	45
Annex 4: Interview Reporting	50
Annex 5: Interview thematic results	52

List of abbreviations

AI	Artificial Intelligence
DFIs	Development Finance Institutions
ESG	Environmental, Social, and Governance
GIIN	Global Impact Investing Network
GIIRS	Global Impact Investing Ratings System
GRI	Global Reporting Initiative
HIPSO	Harmonized Indicators for Private Sector Operations
IFC	International Finance Corporation
IFRS	International Financial Reporting Standards
IIS	Impact Investing Solutions AG
IMM	Impact Measurement and Management
IMP	Impact Management Project
IRIS+	Impact Reporting and Investment Standards
ITU	International Telecommunication Union
JII	Joint Impact Indicators
KG	Knowledge Graph
KPI	Key Performance Indicator
LPs	Limited Partners
MVP	Minimum Viable Product
NLP	Natural Language Processing
OECD	Organisation for Economic Co-operation and Development
OPIM	Operating Principles for Impact Management
PRI	Principles for Responsible Investment
RCT	Randomised Control Trial
SASB	Sustainability Accounting Standards Board
SDGs	Sustainable Development Goals
SRI	Socially Responsible Investment
TCFD	Task Force on Climate-related Financial Disclosures
TNFD	Taskforce on Nature-related Financial Disclosures
ToC	Theory of Change
UNDP	United Nations Development Programme

I. Executive summary

This report investigates whether artificial intelligence (AI) can support the harmonisation of major impact measurement frameworks in the impact investing ecosystem. The study, conducted in partnership with Impact Investing Solutions AG (IIS), combines a literature review of major frameworks with 15 semi-structured interviews involving investors, founders, researchers, and ecosystem actors.

The research started with the assumption that technical fragmentation is the primary barrier to better impact measurement. The impact investing field contains numerous standards, frameworks, and tools. While this fragmentation creates significant reporting burdens, the findings show that the root problem is not a lack of tools, but the lack of incentives to use them rigorously.

In practice, impact measurement is frequently used as a credibility signal rather than a genuine input into investment decisions. Investors remain mainly driven by financial performance, whereas founders, particularly those running early-stage and resource-constrained enterprises, tend to produce simply the minimum amount of reporting required to satisfy expectations. This results in what the paper refers to as a “*credibility-sufficiency equilibrium*”, in which enough measurement is produced to appear credible, but no actor has a strong incentive to go further.

The study therefore reframes the role of AI, it cannot solve the institutional and incentive problems behind harmonisation. However, AI can minimise operational friction by automating reporting, mapping indicators across frameworks, supporting metric selection, and processing qualitative or beneficiary-level data.

The report recommends that actors in the field, and especially IIS, first help convene investors around a shared minimalist standard that defines a credible baseline for impact reporting. AI should then be developed as an operational support layer, rather than a new framework, to help actors navigate fragmentation more efficiently. Ultimately, better impact measurement depends less on technology than on coordinated investor demand, institutional leadership, and stronger incentives.

II. Introduction

The Impact Investing sector has spent years debating which framework to standardise around. As this research shows, this might be the wrong question entirely.

The field, broadly defined as investments that integrate financial returns with intentional social and environmental outcomes (Clarkin & Cangioni, 2015; O'Donohoe, Leijonhufvud, & Saltuk, 2010), faces a well-documented proliferation of measurement methodologies with no dominant standard (Clarkin & Cangioni, 2015; Ogunfowora, 2019; Godeke & Briaud, 2020; Wilhelmi et al., 2024; Ferreira & Köbel, 2025). The academic literature and practitioners alike treat this fragmentation as the primary barrier to transparency and comparability (Clarkin & Cangioni, 2015; Gibson et al., 2025). The conventional prescription includes improving frameworks, increasing their interoperability and potentially even exploring AI-assisted tools to help map the heterogeneous standards.

This project, conducted in partnership with Impact Investing Solutions AG (IIS), a Swiss advisory and technology firm that develops AI-powered platforms to channel capital towards businesses with meaningful social and environmental impact (Impact Investing Solutions, 2026), set out to evaluate that prescription. It investigated how major frameworks differ and overlap, how stakeholders use or adapt them in their daily work, what major barriers prevent standardisation, and whether AI can support alignment without necessarily adding another framework to an already crowded landscape. To do so, the study combined a comprehensive literature review with 15 semi-structured interviews across investor, founder, and ecosystem stakeholder groups, analysed through hybrid thematic analysis.

The empirical findings complicated the starting assumption. Practitioners do not lack adequate measurement tools. Rather, they lack the motivation and incentives to use them rigorously. Investors treat impact reporting primarily as a trust and legitimacy signal rather than a genuine decision-making input. Resource-constrained founders rationally invest in the cheapest methodology that satisfies investor expectations rather than the most accurate one. This study identifies the result as a "*credibility-sufficiency equilibrium*": a structural stable-state in which just enough measurement occurs to maintain credibility, but no actor has sufficient incentive to push past that threshold. The barrier to harmonisation is motivational before it is technical, and demand-side before it is supply-side.

This reframing has direct consequences when it comes to potential solutions or the role of AI. Technologies such as Natural Language Processing (NLP), a branch of AI enabling machines to process and interpret human language, and cross-framework mapping can meaningfully reduce reporting burden and help practitioners navigate existing standards more efficiently. However, they cannot resolve the incentive problem that underlies fragmentation.

Our core recommendation is therefore structured in two elements. The first and necessary intervention is investor coordination around a shared minimalist standard that raises the standard for credibility. The second is an AI-driven operational support platform that makes working within a still-fragmented landscape less costly for both investors and founders.

The remainder of this report is structured as follows: a literature review mapping the framework landscape and existing debates around AI and harmonisation, a methodology and empirical contributions section, an analysis and discussion of the findings and what they mean for the field, and finally concrete recommendations for the ecosystem.

III. Literature review

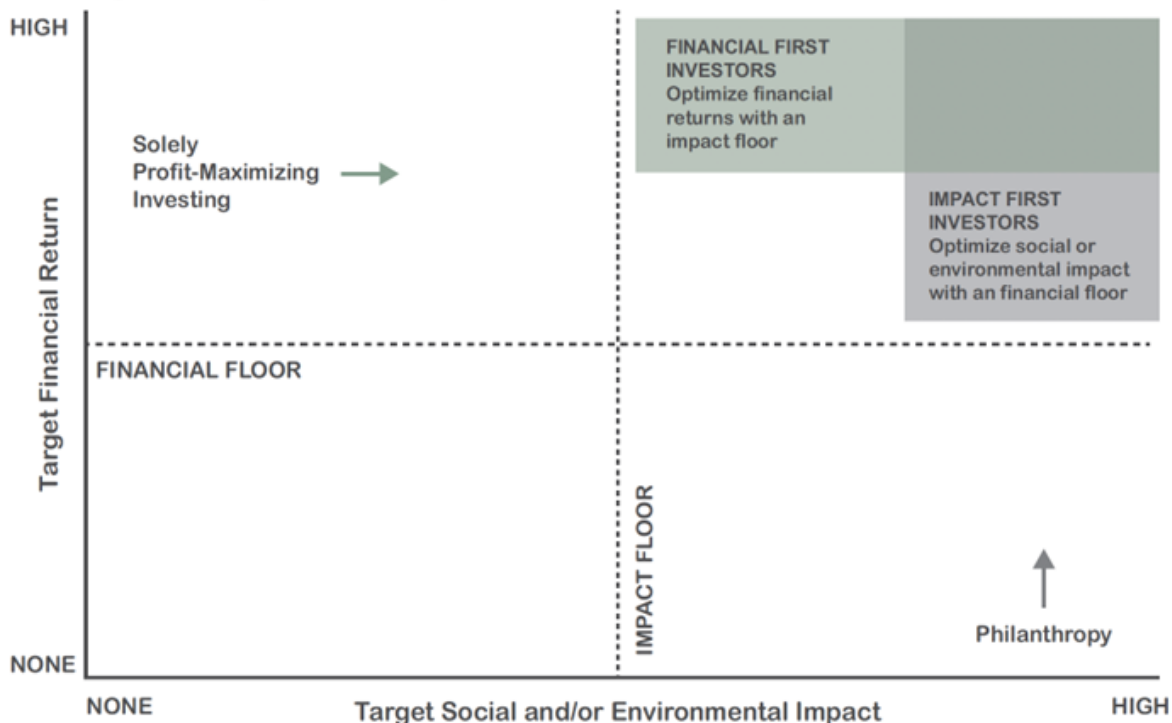
3.1 Impact Investing

3.1.1 Impact Investing as an evolving but conceptually ambiguous field

Impact investing mobilises global capital to address complex socio-environmental challenges, fostering a strategic convergence between philanthropy and traditional financial investment (Clarkin & Cangioni, 2015; O'Donohoe, Leijonhufvud, & Saltuk, 2010). The concept of impact investing emerged from the field of Socially Responsible Investment (SRI) and corporate responsibility movements in the late 20th century, but it differs from SRI by prioritising the active generation of positive and measurable outcomes, rather than focusing strictly on mitigating negative externalities (Ormiston et al., 2015; Gutterman, 2024).

The field has been spurred by the millennial generation (those born between 1980 and 1995), who are more inclined to align their investments with their personal values, a trend that is expected to grow among younger generations (Ogunfowora, 2019; Godeke & Briaud, 2020). Although the sector has demonstrated robust growth, surpassing the USD 1 trillion mark (Yau, 2024; Gibson et al., 2025), the literature notes a persistent heterogeneity between the impact-first and finance-first approaches (Arosio, 2011; Hebb, 2013; Busch et al., 2021), as shown in [Figure 1](#). This conceptual fluidity, however, poses challenges to the rigour of academic research and the comparability of data, highlighting the need for safeguards against impact washing practices (Agrawal & Hockerts, 2021; Brest & Born, 2013; Gibson et al., 2025).

It is therefore suggested that impact investing would benefit from more standardised and comparable impact assessment metrics (Vallejo & Wehn, 2016; Singhania & Swami, 2024; Gibson et al., 2025). Although organisations such as the Global Impact Investing Network (GIIN), Rockefeller Foundation, and International Finance Corporation (IFC) have been working to identify a common language for impact accounting, no widely accepted framework yet exists (Boiardi & Stout, 2021), especially for integrating financial and non-financial dimensions (Singhania & Swami, 2024). This conceptual instability is reflected not only in the standards landscape, but also in the relationship between actors in the field.



Source: Monitor Institute

Figure 1 - Impact investing spectrum: The figure shows how investment approaches vary according to expected financial return and intended social or environmental impact. Source: Arosio, 2011

3.1.2 Structural mismatches in incentives and capacity

Impact Measurement and Management (IMM) refers to the process through which impact investors assess the effects of their investments on people and the planet and use this information to adapt practices and improve outcomes. It combines measurement with management and relies on tools such as Key Performance Indicators (KPIs), ratings, disclosures, and reporting frameworks (Godeke & Briaud, 2020; Gutterman, 2024).

There is a notable misalignment of incentives and capabilities within IMM:

- **Investors:** While requiring comparable and robust impact data for accountability towards Limited Partners (LPs) and governments, they often encounter challenges in isolating the additionality (i.e. incremental net gain generated by the investor) of their capital from a company's intrinsic impact (IFC, 2021; Brest & Born, 2013).
- **Investees or Social Enterprises:** As these organisations are frequently resource-constrained startups, they may lack the structural capacity to implement complex IMM systems, even though they are often the ones expected to produce the data investors need. They therefore tend to

lean instead towards narrative-driven and lightweight models (Ebrahim & Rangan, 2014; Arena et al., 2015; Eckerle & Terzidis, 2025).

While this reflects a broader problem in finance, in impact investing this tension is intensified because the information requested is not only financial but also social and environmental.

This mismatch also includes a geographical dimension. Much of the impact investing capital and many impact measurement frameworks originate in developed economies, particularly within US and European institutional contexts, while investment demand is often concentrated in emerging markets where social and environmental challenges are more pressing (Arosio, 2011; Ormiston et al., 2015). This adds another capacity gap since organisations operating in these contexts are likely to face weaker data infrastructure and fewer resources to meet complex reporting expectations, which can potentially compromise the continuity of investments (McCallum & Viviers, 2020; Singhania & Swami, 2024).

3.2 Impact Measurement tools and frameworks

3.2.1 *The rise of measurement as a source of legitimacy*

Prior to the formal emergence of IMM, established methodologies for evaluating social and environmental programmes were already widespread, most notably the Theory of Change (ToC) and the Logic Model (Ebrahim & Rangan, 2014; Clarkin & Cangioni, 2015; Ogunfowora, 2019). The Theory of Change articulates both the “how” and the “why” of an intervention, outlining the pathways required to achieve the intended impact (Ogunfowora, 2019; Godeke & Briaud, 2020), while the Logic Model focuses on the operational “what”, a linear sequence of inputs, outputs, and outcomes (Ebrahim & Rangan, 2014; Ogunfowora, 2019). Within impact investing, both primarily serve as guides for structuring impact strategies but do not, in themselves, constitute measurement tools (Ogunfowora, 2019; Godeke & Briaud, 2020; Gibson et al., 2025).

Thus, measurement has become the key factor in distinguishing impact investing from other forms of sustainable investing (Gutterman, 2024; Gibson et al., 2025). The origins of current standards date back to a 2008 meeting in Bellagio, Italy, organised by the Rockefeller Foundation, where investors and philanthropists identified the need for rigorous metrics to legitimise the nascent sector. This gave rise to two fundamental tools: the Impact Reporting and Investment Standards (IRIS), conceived as a “universal language” for reporting social, environmental, and financial performance, and the Global Impact Investing Rating System (GIIRS), which introduced rating methodologies analogous to those used in traditional financial markets (GIIN, n.d.; Ogunfowora, 2019; Clarkin & Cangioni, 2015).

Since then, the landscape of IMM has expanded considerably, leading to the emergence of a wide range of standards, frameworks, and reporting approaches (Clarkin & Cangioni, 2015; Ferreiro & Kölbel, 2025;

Gibson et al., 2025). These standards can be classified into three broad categories: principles, which communicate the investor's intent and the integrity of the process; frameworks, which organise and guide the implementation of IMM; and tools, which provide the indicators and reporting systems through which measurement is conducted (Godeke & Briaud, 2020; Gutterman, 2024). [Table 1](#) illustrates these categories along with their reported adoption rates among impact investors, based on GIIN's 2023 GIINSight survey. The UN Sustainable Development Goals (SDGs) stand out as the most widely used framework (76%), while IRIS+ leads among the tools (78%), reflecting its role as the closest approximation to an industry-wide standard (Hand et al., 2023).

Institutional and academic literature mentions the existence of an even greater variety of standards (Clarkin & Cangioni, 2015; Ogunfowora, 2019; Godeke & Briaud, 2020; Ferreira & Kölb, 2025; Wilhelmi et al., 2024), in addition to those mentioned above. [Figure 2](#) presents a word cloud of the main examples, where the size of each acronym reflects the frequency with which it appeared in this research, and the colours represent each category (green for principles, blue for frameworks, and yellow/orange for tools). [Annex 1](#) provides a more comprehensive table of standards, including their descriptions and the organisations that developed them.

Table 1. IMM taxonomy and market reach: This table distinguishes between principles, frameworks, and tools used in impact measurement and shows their reported adoption among investors. Source: Based on Godeke & Briaud, 2020; Gutterman, 2024; Hand et al., 2023.

Category	Primary Function	Key examples	Market reach (GIINSight 2023)
Principles	Rules and best practices for communicating investor's intent and ensuring overall process integrity. They are not specific measurement techniques.	International Finance Corporation (IFC) principles and UN Principles for Responsible Investment (PRI)	IFC (43%), PRI (40%)
Frameworks	Methodologies used to organise IMM and implement the intentions expressed in the Principles.	UN Sustainable Development Goals (SDGs), Impact Management Norms	SDG (76%), Impact Management Norms (52%)
Tools	Systems, metrics catalogues, or reporting standards designed to execute IMM.	Impact Reporting and Investment Standards (IRIS+), Sustainability Accounting Standards Board (SASB), and Global Reporting Initiative (GRI) standards.	IRIS+ (78%), SASB (22%), GRI (19%)

3.3 AI and Impact Measurement harmonisation

3.3.1 Why harmonisation remains limited and contested

The literature suggests that progress has been made towards harmonising the impact measurement field, particularly through effort to align principle, metrics and reporting practices. A paper from OECD highlights that significant efforts have been made to harmonise impact measurement frameworks and metrics (Boiardi & Stout, 2021), notably through initiatives like IRIS+, HIPSO, Operating Principles for Impact Management (OPIM) metrics harmonisation to increase data utility and organisational reporting efficiencies (Narain et al., 2012). Narain et al. (2012) also underlined that harmonisation is feasible by having shared principles, common metrics and consistent management practices. However, it is argued that developing a complete unification is neither possible nor desirable due to the diversity across various regions and contexts in which impact investing operates.



Figure 2 - Main standards in the IMM ecosystem: Acronyms' size reflects citation frequency and colour denotes category. Source: Hand et al., 2023; Clarkin & Cangioni, 2015; Ogunfowora, 2019; Godeke & Briaud, 2020; Ferreiro & Kölbel, 2025; Wilhelmi et al., 2024.

Despite these efforts, the field has not yet found a solution to the whole problem of comparability. Ferreiro & Kölbel (2025) show that numerous investors in impact investing expressed concerns about the absence of standardisation in impact measurement, which leads to a decrease in their investments' ability due to various metrics' use. This issue is not isolated and it recurs across multiple reporting standards and assurance frameworks, including Global Reporting Initiative (GRI), Sustainability Accounting Standards Board (SASB), and Task Force on Climate-related Financial Disclosures (TCFD), undermining the comparability and reliability of Environmental, Social, Governance (ESG) disclosures

and ultimately generating greater confusion among stakeholders rather than clarity (Mayegun & Nwanevu, 2025). In this sense, the literature thus reveals a core tension: while harmonisation is growingly seen as necessary, the mechanisms currently existing still are limited to produce a coherent and broadly accepted reporting standard. While most metrics support impact reporting, they are still inadequate for performance measurement due to weak incrementality and comparability (Gibson et al., 2025). In direct response to this tension, Mayegun & Nwanevu (2025) started to call for the modernisation of current traditional reporting systems by using Big Data and AI to make them standardised, verifiable and delivered in a timely manner.

3.3.2 AI and harmonisation: opportunities and limits

Within impact investing, AI is increasingly presented in the literature as a potentially important tool for supporting how investors make decisions to achieve their intended outcomes and impacts (Ojonugwa et al., 2024). Ferreiro & Kölbel (2025) further emphasise that AI could be leveraged to find a pathway towards harmonisation of those practices and metrics. Indeed, they recommend AI-based tools to be harnessed in order to reduce the complexity and expenses of impact assessments across different projects (Ferreiro & Kölbel, 2025). Thus, AI is not introduced as a substitute for existing standards, but as a potential means for enabling fragmented systems to be more interpretable, comparable and operationally manageable.

Current debates around the use of AI underline the need for shared technical frameworks and interoperability. Several organisations like the International Telecommunication Union (ITU), the UN or OECD note that the future of AI governance's effectiveness will rely on achieving interoperability across different frameworks (International Telecommunication Union, 2024). Interoperability thus matters since it could potentially create more enforcement of AI regulations across different regional contexts. However, the literature is also clear that AI is not a standalone solution. Its success and usefulness are dependent on data quality, governance, transparency and human input because poorly designed AI systems are likely to reproduce bias, weaken trust rather than improve it (International Telecommunication Union, 2024; Ojonugwa et al., 2024; Mayegun & Nwanevu, 2025; Elhady & Shohieb, 2025).

More broadly, AI can support harmonisation by processing large volumes of structured and unstructured data and by improving the analysis and comparability of sustainability-related disclosures (Cojocaru et al., 2024; Mayegun & Nwanevu, 2025). In particular, AI-based techniques such as text analysis and data classification can help identify patterns across heterogeneous datasets, while ontology-based and knowledge graph (KG)-based models may assist in aligning concepts used and expressed differently across frameworks (Ehrlinger & Wöss, 2016). However, these tools are only reliable when data and governance structures are well established (De Villiers et al., 2024; Ojonugwa et al., 2024; Mayegun &

Nwanevu, 2025), since weak data quality and limited oversight may undermine the accuracy and trustworthiness of the outputs.

Although there is no existing literature addressing impact investing measurement harmonisation using AI, an analogous research exists in financial reporting where heterogeneous reporting standards were harmonised through the use of a combination of branches of AI (Badmus et al., 2025) such as Natural Language Processing (NLP) and KG-based approaches. This indicates a plausible avenue for inquiry, with AI having the capacity to support harmonisation in a field characterised by similar fragmentation.

However, research indicates that AI can create significant ethical and governance risks. Particularly when it is applied to reporting and investment decision-making, issues with algorithmic bias, data privacy, openness, and confidentiality are crucial (Ojonugwa et al., 2024; Alshi, 2025). These risks highlight the necessity of strong human oversight, data stewardship, and governance mechanisms to ensure that AI systems continue to be reliable, auditable, and compliant with ethical and legal requirements (Mayegun & Nwanevu, 2025; Ottun & Flores, 2025).

IV. Empirical Contributions and Interview Methodology

4.1 Empirical contributions

This study provides two main empirical contributions. First, it offers a synthesis of the current impact measurement landscape by sorting major standards, frameworks, principles and tools based on their function ([Annex 1](#)). Built on the literature reviewed, this classification clarifies that the field is not only fragmented because many instruments are coexisting, but also because they serve different purposes and operate in different ways. Second, this study contributes first-hand qualitative evidence through 15 semi-structured interviews conducted between March and April 2026 with practitioners from across the impact investing ecosystem. This primary data allowed this research to move beyond the literature and examine how impact measurement frameworks and reporting practices are understood and used in practice.

4.2 Interview Methodology

Semi-structured interviews were selected because the research questions concern practices, motivations, and barriers that are not visible in published sources. The sample was constructed using a purposive, non-probability strategy and included stakeholders across four different groups (startup founders, investors, researchers and ecosystem support agents) and five different world regions. [Figure 3](#) presents the distribution of interviewees across stakeholder groups.

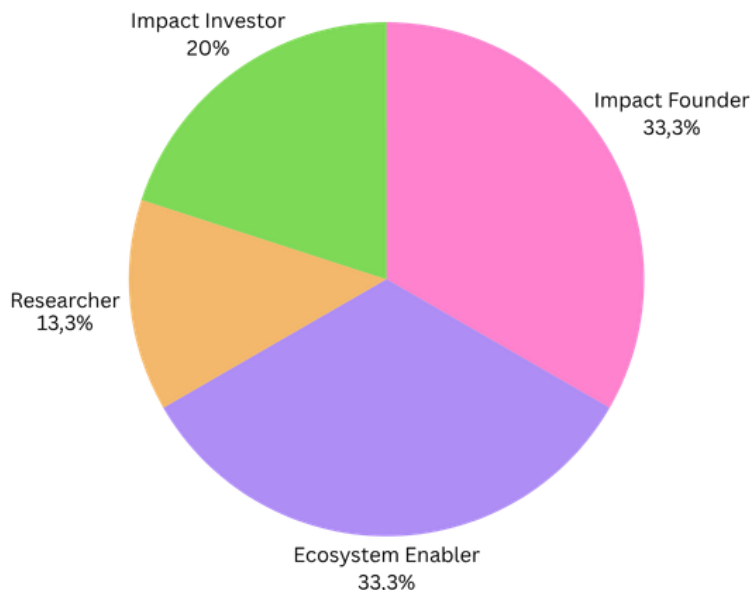


Figure 3 - Distribution of interviewees by stakeholder category: The figure shows the distribution of interviewees across the main stakeholder groups included in this study. Source: Created by the authors.

The interview data were analysed using a hybrid thematic analysis following the six-phase framework of Braun and Clarke (2006), illustrated in [Figure 4](#). The approach combined 25 deductive codes derived

from the research questions and the literature with 16 inductive codes that emerged directly from the data.



Figure 4 - Six phases of thematic analysis: The figure summarises the six-step thematic analysis process used to code and interpret the interview data. Source: Based on Braun & Clarke, 2006.

We began by transcribing the interviews and reading them to ensure full familiarity with the data. Next, we generated initial codes for interesting features and organised these into broader themes, as illustrated in [Figure 5](#). These themes were then reviewed and refined to ensure they accurately represented the interviews before being clearly defined and named. This systematic process formed the basis for the final analysis presented in the following discussion section.

These inductive codes, including the finding that impact measurement functions primarily as a credibility signal rather than a decision input, proved to be the most significant outputs of the process and carry the central arguments of the discussion section.

A full account of the research design, sampling strategy, analytical approach, and complete code list is provided in [Annex 2](#). The interview guiding questions are included in [Annex 3](#). Additional figures and breakdown about interviewees' profiles is provided in [Annex 4](#). Finally, the detailed thematic presentation of the interview material is provided in [Annex 5](#).



Figure 5 - Results themes: The figure presents the main themes that emerged from the interview analysis and structured the discussion section. Source: Created by the authors.

V. Discussion

5.1 Comparison between standards

Following the literature's proposal to divide IMM standards into three categories (principles, frameworks, and tools), we were able to analyse them in order to identify complementarities, differences, and both positive and negative aspects across acronyms.

Starting with the similarities: they all belong to the broader IMM ecosystem and therefore share the goal of increasing transparency and comparability of investments. The SDGs play a particularly important role in this landscape, functioning as a common reference framework across multiple tools such as in IRIS+ (e.g., SDG-IRIS+ Detailed Alignment) and in GRI (e.g., SDG Mapping Add-on). Even on the SASB website, the SDGs are referenced as relevant global goals connected to their respective standards, reflecting its broad legitimacy across the field. Regarding the principles (e.g., IFC and PRI), they also connect conceptually, though they may appear somewhat redundant. Still, their relevance is more symbolic and normative, serving as high-level commitments that guide organisations rather than providing operational measurement tools.

When looking at differences, standards can also vary in scope and function. For instance, GRI and SASB both focus on economic, environmental, and social impacts and include sector-specific disclosure standards. However, GRI is more universal, applicable to organisations of any type and geography, while SASB is narrower and financially focused, offering metrics for 77 specific industries. Because of this, they may appear to compete, but in practice they are widely used together, as they serve different stakeholder needs (Ferreiro & Kölbel, 2025; Gutterman, 2024).

Still, the proliferation of acronyms contributes to the perception of an "alphabet soup", illustrated in [Figure 2](#). This multiplication of frameworks and the lack of regulatory harmonisation result in sustainability reports that vary widely in size, format, and analytical depth (Wilhelmi et al., 2024; Gibson et al., 2025), requiring additional effort from experts for interpretation, assurance, and auditing. Therefore, what enables flexibility and adaptation across contexts comes at the cost of a lack of comparability and coherence in the reporting landscape.

Nevertheless, efforts are emerging to bridge these gaps, most notably the Joint Impact Indicators (JII) initiative. Launched in 2021 as a collaboration between the Global Impact Investing Network (GIIN) and the Harmonized Indicators for Private Sector Operations (HIPSO), the JII maps 27 IRIS+ metrics to 18 HIPSO indicators across common themes such as Climate, Gender, and Jobs (HIPSO, 2021). It creates a "common language" that allows impact investors and Development Finance Institutions (DFIs) to compare project performance. While JII can reduce the reporting burden and facilitates benchmarking, it operates primarily at the level of output-oriented metrics. Consequently, even with such promising alignments, the broader challenge of interpreting context-specific impact signals remains (Ferreiro & Kölbel, 2025).

5.2 Visualising the landscape

By combining the findings from the literature with the interviews conducted, it was possible to summarise the current state of the impact investing and IMM field, as illustrated in [Figure 6](#). The diagram organises the landscape into interdependent layers.

In the top layer, investors (i.e., Limited Partners (LPs), DFIs, funds, and family offices), concentrated predominantly in the Global North, allocate capital to the investees, mostly social enterprises and startups usually located in the Global South. In turn, the investees must provide impact data, creating a bidirectional flow mediated by the IMM layer.

This IMM layer brings together the field's key standards, organised into the three categories: principles (e.g., IFC/OPIM, PRI), frameworks (e.g., SDGs, Impact Management Norms), and tools (e.g., IRIS+, SASB, GRI). It is where impact measurement and management efforts are concentrated, but also where the field's main structural tensions emerge: conceptual ambiguity among stakeholders, a lack of incentives to measure impact rigorously, and the fragmentation of standards, which overburdens investees and often results in outputs that do not reflect actual impact.

At the frontier of this landscape, AI emerges as a potential layer of support. However, its application remains in its incipient stages, which is why it is represented by a dashed arrow, signalling a possible direction that is under development.

Impact Investing Landscape

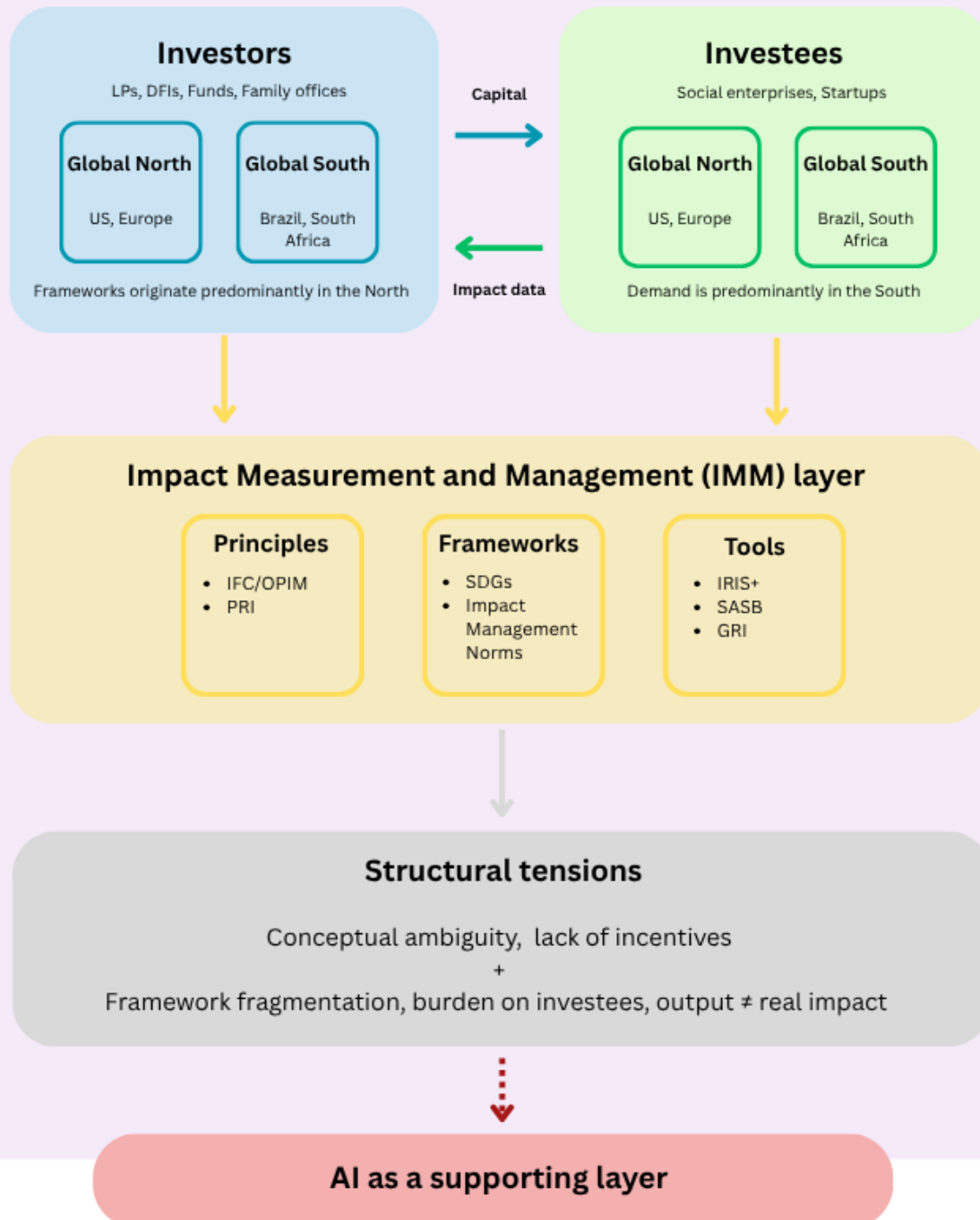


Figure 6 - Impact Investing Landscape: The figure maps the main actors, standards, tensions, and potential AI support layer within the impact investing ecosystem. Source: Created by the authors.

5.3 The key barrier: moving from the technical to the motivational

The literature analysed in the review part of this study framed the issue of impact measurement as a supply-side problem. As conventionally stated, the problem is the existence of a number of frameworks and metrics that is too large to make formal impact measurement and communication both easily feasible and popular. The most prominent critique is that frameworks like IRIS+, GRI, SASB, the SDGs, or Impact Management Project (IMP's five dimensions) create a structural confusion for practitioners: while each captures a legitimate dimension of impact and has a believable potential use case, none of them are interoperable or dominant enough to function as a common language across the ecosystem (Höchstädter and Scheck, 2015; Wilhelmi et al., 2024). The proliferation of standards creates the “alphabet soup” where these tools coexist but fail to create comparability of impact data. Based on this literature, we expected a potential solution to be technical such as a better standard alignment, framework design or AI-assisted interoperability to automate the mapping of these heterogeneous metrics.

Our original framing reflected this expectation: the inception report treated framework fragmentation as the primary barrier to harmonisation and proposed to investigate the feasibility of launching technical solutions in the space. The empirical work subsequently conducted, however, tells a very different story. Our interviews across investor, founder and ecosystem stakeholder groups consistently pointed to a more fundamental constraint. Actors do not lack adequate measurement tools, they lack the motivation and incentives to use them rigorously. The problem, therefore, is more on the demand-side than on the supply-side, more motivational than technical.

The “alphabet soup” is real and creates real operational frictions. However, it is not the primary obstacle to impact measurement becoming standard practice. The deeper problem is a structural lack of incentives for rigorous measurement on both sides of the investor-founder relationship, compounded by persistent conceptual ambiguity around what “impact” actually means in practice. The result is what this study identifies as the “*credibility-sufficiency equilibrium*”: a structurally stable state in which just enough measurement occurs to maintain credibility, but no actor has sufficient incentive to push past that threshold.

5.3.1 Impact measurement's role in practice

One of the most consistent findings across our interviews is that impact measurement often serves as a trust and legitimacy signal rather than as a primary decision-making input. The frameworks (e.g., IRIS+, GRI) are all designed on the assumption that measurement informs decisions, making this level of detail important for performance evaluation in both funds and individual companies. Our data suggests this is not how measurement is operationally used by most actors. The evidence for this comes from several different sides.

On the investor side, Joe Seddon (Founder of Zero Gravity), who runs Randomised Control Trials (RCTs) with UCAS, the centralised UK university admissions portal, and has built a proprietary counterfactual model for lifetime earnings at Zero Gravity, offered one of the most precise accounts of this gap: “*as long as you seem like you have a coherent methodology on how you measure it and communicate it, a*

lot of firms aren't going to get into the weeds of exactly what that looks like, because they'd almost just want the kind of stamp of approval." He added that even Oxbridge colleges with whom Zero Gravity has commercial relationships do not examine the RCT methodology: *"they love it from a messaging point of view because again it's just a trust signal."* His conclusion is unambiguous: *"do people make investment decisions by looking at a social impact model and evaluating it rationally? Absolutely not. I can't think of a single instance."* His rigour generates no incremental advantage in how partners evaluate him should not be interpreted as an anomaly but rather as a structural feature of how the market operates.

This account is corroborated by Tenke Zoltani (Founder of Better Finance), who described working for years on impact reporting for a family office before being advised that this depth of reporting was unnecessary: *"we know we're creating impact, you don't need to count the number of jobs that were created or the number of women that were hired. [...] The numbers don't matter to us as much as the financial performance."* When asked what had driven impact reporting in her experience, her answer was that it was peer pressure: *"I would just say peer pressure, being very honest, because the expectation from investors is that they have to have an impact report to prove what they did was impactful. I don't think there was a very strong driver otherwise."* The mechanism she describes is normative conformity, reporting as a response to peer expectations within the investment community, rather than any genuine demand for the underlying data.

On the founder side, interviewees primarily report their impact to satisfy investor expectations and sustain credibility in fundraising contexts. Given founders' limited resources, they focus on producing the minimum credible signal for investors rather than more rigorous tracking. Carmen Chan (Founder at Girls be Change) describes the reaction she often got from investors when explaining her impact. They would say something like *"Well, it's women empowerment. You're working with change makers, this, this, this — tick, tick, tick. And then now show me the money."* Measurement serves more as a threshold that needs to be crossed than as a guide.

A development finance specialist, with a cross-portfolio view of financial inclusion investing, was the most direct of all interviewees in naming what is actually being measured: *"what they're measuring is not impact. They're measuring really outputs. And I'm not saying it doesn't have a value, but it's not impact."* According to her, *"it suits everybody."* Investors can report metrics to their limited partners without exposing what is sometimes shallow or underprioritised due diligence. Founders can satisfy reporting requirements without undertaking costly outcome-level research their resources cannot sustain.

5.3.2 The stability of the equilibrium

If measurement is performing a signalling function, the question we can ask is: why doesn't a better signal create competitive pressure for more rigorous measurement? In other markets with signalling dynamics, we would expect at least some actors to differentiate by offering higher-quality data. Joe Seddon is the closest example in our data: he runs RCTs and has a proprietary counterfactual model. He observes that this rigour does not give him any incremental advantage in investment decisions. The market, as currently structured, does not reward measurement quality above a minimum credibility threshold.

To make sense of this, one can again look at the incentives on the different sides of the ecosystem.

On the investor side, no individual investor has the incentive to ask for more rigorous data unilaterally. Doing so increases the compliance cost and friction for founders, potentially pushing them to seek other investment sources, without improving the investor's competitive position. There is just no incentive to go further. Daniel Nowack (Head of Social Innovation at World Economic Forum), named this dynamic when describing the incentives of compliance-oriented investors: *"they care as much as they need to have it for compliance, right? [...] as long as I comply with that and as long as I'm able to qualify for Article 9, I'm fine."* Beyond the minimum regulatory threshold, he claims, *"they don't have a whole lot of incentives to go deeper into it."* The incentive is to satisfy the credibility floor, which is set by peer norms within the investment community, not by any external standard, rather than exceed it. Daniel Nowack adds another layer to this: because impact investors tend to impose more rigorous reporting requirements than conventional investors, they effectively raise the cost of capital for the very founders they aim to support. In his words, measurement is *"in 95% of all cases really a cost center and doesn't directly yield additional revenues"* for the investee. There is no financial offset for the burden it creates.

On the founder side, the rational response to an investor community that rewards the presence of methodology and impact narrative rather than its quality is to invest into the cheapest credible methodology available. This might be output counting, SDG alignment statements or narrative-based theories of change. Audrey Selian (Director of Artha Impact) gave the most candid account of how using proxy measurement as a tool for impact accounting: *"they sold 10,000 Fox Light torches to smallholder farmers in rural, unelectrified areas [...] people use multipliers, there are all kinds of standard multipliers, like the average household size being five. So if you sold 10,000 torches, in an impact report people will multiply that by five and say 50,000 people have benefited from light."* She acknowledged directly that this proxy approach is the industry norm and that going beyond it, to third-party verification, beneficiary feedback loops, is prohibitively expensive: *"most investors won't do it, including us. We haven't been able to afford it."*

The output/outcome conflation named by a development finance specialist and discussed in our findings is therefore not a lack of knowledge but rather a rational way of adapting to incentive structures.

The result of these dynamics is what we are calling the **"credibility-sufficiency equilibrium"**. There is a stable state where just enough measurement happens to maintain the appearance of rigorous methodology, thereby maintaining credibility, but without any actor having more incentive to push past that threshold. This is not a problem of bad intentions, ignorance, or a lack of care for impact. It is merely the outcome of rational behaviour when rigorous measurement is not rewarded for most.

5.3.3 What this means for the harmonisation agenda

If the equilibrium in the impact measurement state is indeed provoked by bad incentives and correspondingly low motivation, then the policy interventions that are suggested in the literature (i.e. better frameworks, AI-assisted interoperability or common standards) address the wrong constraint. They assume that the bottleneck is the technical quality of tools. We find that the real constraint is

demand: no one in the current ecosystem is incentivised or positioned to insist on rigorous and comparable measurement.

The practical implication for IIS specifically is that its leverage lies not at the norm-setting layer. Given its position as a platform between investors and founders, IIS has the ability to convene a coalition that collectively changes what credibility requires. If investors in the IIS network agree to require a shared minimalist standard as a condition of consideration, the founder incentive changes.

5.4 The role of AI in the Impact Investment ecosystem

5.4.1 AI and the limits of harmonisation

The original framing of this research saw AI as a potential tool for harmonising fragmented impact measurement frameworks. The expectation, considering most recent literature, was in fact that AI could act as a unifying layer. For example, processing heterogeneous data, aligning indicators across standards, and producing more comparable, interoperable reporting (Ferreiro & Kölbl, 2025; Mayegun & Nwanevu, 2025; Wilhelmi et al., 2024; Ojonugwa et al., 2024). The interviews suggest a deeper issue: the literature implicitly frames fragmentation as a technical problem, whereas practitioners experience it primarily as a governance and incentive problem.

No interviewee expressed their interest in AI in terms of harmonisation. Across the interviews, the term “harmonisation” was used and applied to a governance and institutional problem, not really a technological one. Daniel Nowack captured this point particularly well. Rather than dismissing AI altogether, he situated it within its proper scope. AI can be powerful at processing, organising, and aligning data across fragmented systems, but it cannot create the institutional coordination, shared commitment, or incentive structures required for genuine harmonisation. This distinction between data alignment and genuine standardisation is drawn much more clearly by practitioners than the literature which tends to blur the two.

The findings do not suggest that AI is irrelevant for the impact measurement field. It rather indicates that the focus of the question needs to shift. There should not be a question about whether AI can harmonise the field as such, but which specific problems AI can realistically solve within this fragmented landscape. This suggests that AI is often applied to a problem framed as technical, whereas practitioners understand it primarily as institutional and incentive-based.

5.4.2 Practitioner openness to AI

The interviews show that practitioners’ openness to AI is present, but under several conditions. Rather than viewing AI as a route to full harmonisation, most interviewees saw it as a practical tool for reducing the burden of working across fragmented systems. This nuance contrasts with the literature that frames AI in terms of its potential to enhance interoperability, comparability, and harmonisation (Ferreiro & Kölbl, 2025; Mayegun & Nwanevu, 2025; Cojocaru et al., 2024).

What practitioners appear to want is not a new standard, but a more efficient way of navigating existing ones. AI is attractive mainly because it can simplify repetitive processes, support alignment across

frameworks, and reduce reliance on labour-intensive, individualised work. Its appeal therefore lies not in solving these structural problems of impact measurement, but in making the existing reporting environment more manageable in practice. This shifts the question from whether AI can unify impact measurement to whether it can meaningfully reduce the operational costs created by its current lack of unification.

5.4.3 What AI cannot do - structural limits

Alongside their openness, several interviewees shared specific limits, not as objections to AI, but as conditions marking where its contribution ends. The main limits of AI in impact measurement are not really technical, but more social, ethical and structural. The literature often emphasised AI's potential for interoperability, data processing, and harmonisation but gave less attention to the conditions under which those functions remain meaningful in practice.

An important distinction that the literature tends to overlook concerns the nature of the data itself: between output-level data which AI can process, and real impact data which is self-reported, context-specific, and tied to people's lived experience. This shows not only a technical limit, but a deeper one in how impact can be captured. It creates a paradox where the most meaningful data for evaluating actual impact is also the least standardisable and least suited to automation. Human judgment and contextual knowledge remain core elements even when AI is used.

Another limit concerns legitimacy. Interviews highlighted the risk of using AI on beneficiary data where individuals may not have full knowledge of, or explicit control over, how their data and information is processed. Although the literature discusses governance and privacy in general terms, the interviews show that consent is not a secondary issue but a core condition for the legitimate use of AI in this field.

The literature tends to present AI as a tool to improve weak reporting systems and conditions (Cojocaru et al., 2024; Ferreira & Köbel, 2025), while often treating data quality as a secondary condition. Although several studies acknowledge that AI depends on robust data and governance structures (De Villiers et al., 2024; Ojonugwa et al., 2024; Mayegun & Nwanevu, 2025; International Telecommunication Union, 2024), the interviews suggest that this dependency is in fact decisive. In practice, data quality is not an outcome of AI adoption, but a precondition for it to generate meaningful value.

Connectivity and language access also emerged in the interviews as structural limits to the applicability of AI, especially in the contexts where impact investing is most active. This matters because, as the literature already shows, impact investing is characterised by a North-South asymmetry in which frameworks and reporting infrastructures are largely designed in high-resource Northern countries, while many target contexts are located in lower-resource environments (Arosio, 2011; Ormiston et al., 2015; McCallum & Viviers, 2020). AI therefore does not enter a neutral field. Rather than reducing fragmentation, it may reproduce existing disparities by favouring actors already better off in terms of data, infrastructure, and language. [Figure 7](#) visually summarises the main concerns regarding the use of AI as reported in the interviews, where larger circles indicate more frequently mentioned issues.

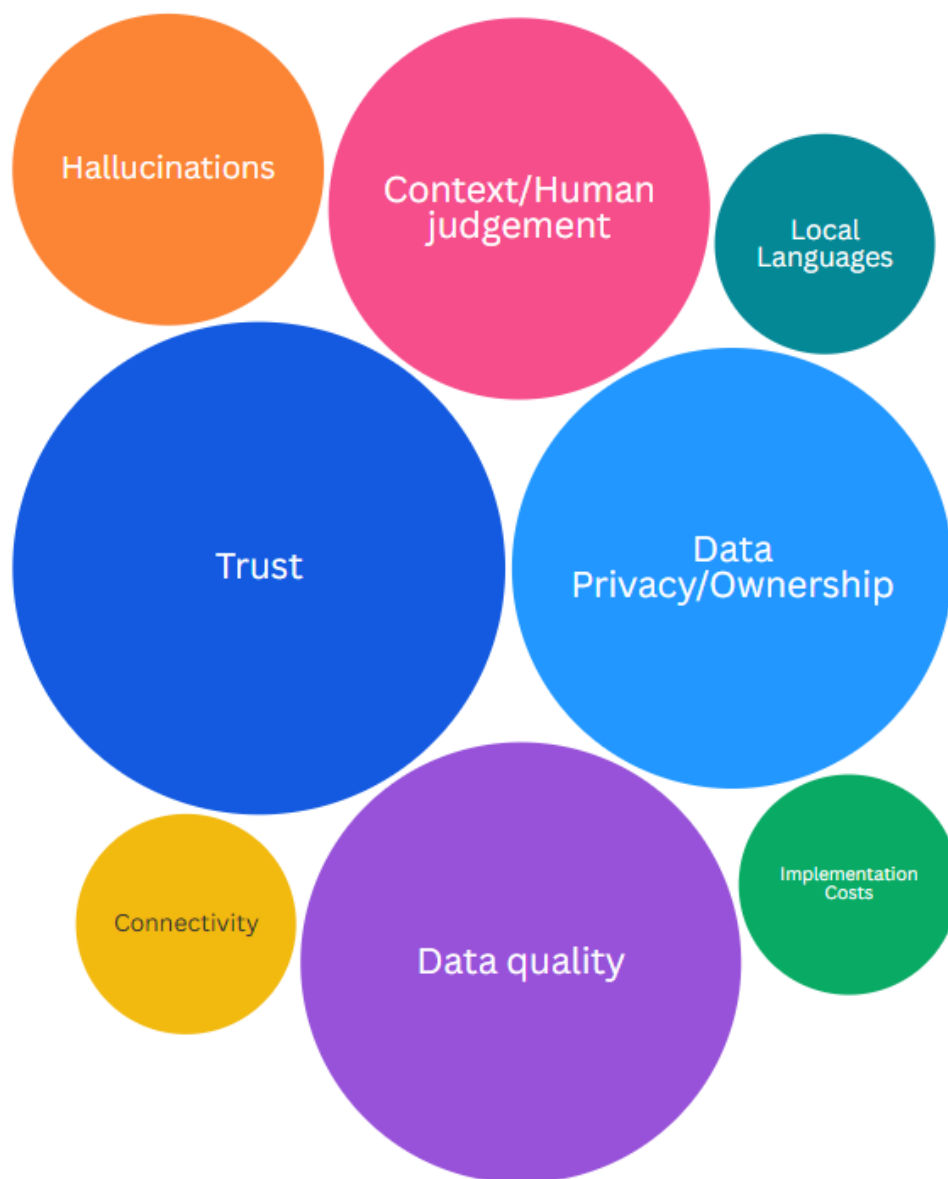


Figure 7 - Main concerns for the use of AI in Impact Measurement: The figure highlights the main concerns raised by interviewees regarding the use of AI in impact measurement. Source: Created by the authors.

5.4.4 What AI can realistically do

A first plausible use case is the reduction of reporting burden, particularly on the enterprise side, through the automation of data extraction, formatting, and translation across multiple templates and reporting systems. A second is framework navigation, where AI could help practitioners identify relevant indicators across existing standards without requiring those standards to be unified. Since many interviewees

already work by combining elements from different frameworks, AI could significantly create value by reducing the time and effort of this manual process. A third concerns qualitative data, where AI may help organise and process narrative or beneficiary-level evidence that is difficult to manage at scale. In that sense, any platform that also includes beneficiaries' perspectives could be highly valuable, given how strongly some interviewees identified the absence of beneficiary voice as one major gap in current measurement practices.

AI should therefore be understood not as a tool for harmonisation, but as an infrastructural support layer that reduces the operational costs of a fundamentally fragmented system.

VI. Recommendations

Given that the findings showed fragmentation in impact measurement is not a technical issue but a governance and incentive one, AI should thus be positioned as an infrastructural support tool reducing operational burden created by fragmentation, rather than a solution to harmonisation itself. Addressing the structural lack of incentives for rigorous measurement requires external pressure and coordination from regulators, institutional investors, and multilateral bodies to raise the baseline level of credibility beyond what any platform alone can achieve.

Before any platform is built, however, the most important intervention IIS can make is structural. Our research consistently points to the conclusion that the “*credibility-sufficiency equilibrium*” will not be shifted by a better tool. It can only be shifted by changing what investors collectively require, something which no investor can do unilaterally. As an institution between investors and founders, IIS has the ability to convene a coalition around a shared minimalist standard as a genuine condition of consideration for founders. If investors in the IIS network agree to require the same core set of metrics the founder incentive changes. These could be six to ten indicators across scale of reach, depth of outcome, and SDG alignment, which should all be achievable for founders to measure. This investor coordination is Tier 1, and it is the precondition that makes Tier 2 — the platform — worth building.

6.1 Strategic Proposal: AI as an operational support layer

Should IIS, and other stakeholders, choose to move forward with an AI initiative, we recommend that the platform should be treated as a second-order coordination mechanism and avoid developing a new standalone framework or attempting to impose standardisation through technology. The initiative should be used as a translation and support layer operating across existing frameworks, to reduce duplication, improve consistency and support more efficient workflow.

6.1.1 The necessary characteristics

Several points must be considered based on the empirical findings:

- The platform should be designed to support and not seek standardisation. It should help users navigate and connect frameworks more efficiently.
- The platform should contain different modules built around three defined use cases instead of one large and complex platform. These tools should directly reflect the operational needs identified in the interviews.
 - The first is reporting burden reduction, particularly on the investee side, through the automation of data extraction, formatting, and translation across multiple templates and reporting systems.
 - The second is framework navigation and metric curation, where AI could help practitioners identify relevant indicators across existing standards without requiring those standards to be unified.

- The third is qualitative data support, where AI can help organise and process narrative or beneficiary-level data that is currently difficult to manage at scale. Since the absence of beneficiary voice was identified as a major gap, integrating such a mechanism should be treated as a core design priority.
- The platform should include a benchmark dashboard that shows how any given founder's reported metrics compare to peers in the same sector, geography, and investment stage. This can be used to directly address the verification problem that was identified across interviews. Making relative performance visible creates a new incentive for founders to report more carefully and gives investors an analytical reference point they currently lack entirely.
- Data governance and legitimacy must be central features. This includes transparent data processing, clear data ownership, and explicit consent mechanisms, particularly when dealing with beneficiary-level data.
- The platform should also be designed to serve actors who already have genuine demand for rigorous measurement (e.g. Development Finance Institutions (DFIs), governments, and large institutional investors with compliance obligations) because these actors are more likely to drive adoption and create reputational pressure on others to follow.
- The platform should be designed to work in contexts with limited resources and imperfect data. This means it should have tolerance for incomplete information, remain simple to use, and support multiple languages. This would help avoid reproducing the North-South asymmetries existing in the field.
- Finally, the platform should provide direct practical value by saving time and fitting into the way users carry out reporting tasks. Its adoption will depend less on technical sophistication than on whether it genuinely reduces reporting effort.

6.1.2 Implementation roadmap

We propose four phases to ensure the feasibility and adoption of such a solution, summarised in [Figure 8](#).

In the first phase, IIS should begin convening a pilot coalition around the creation of a minimalist enforced standard. This should be done collaboratively with coalition members. In parallel, IIS should validate findings through interviews with founders in the network and identify the most acute friction points. This will also be the right time for IIS to initiate early contact with established actors such as GIIN and IFC, as institutional credibility must be built alongside the product, to avoid the platform being perceived as another standalone tool in an already crowded field.

In the second phase, with the coalition active and generating consistent data, IIS develops a minimum viable product (MVP) that incorporates the two most important use cases for its tool: the founder-first narrative extraction tool and the investor facing benchmark dashboard. This MVP should remain tight in scope to ensure that it performs well on what it does.

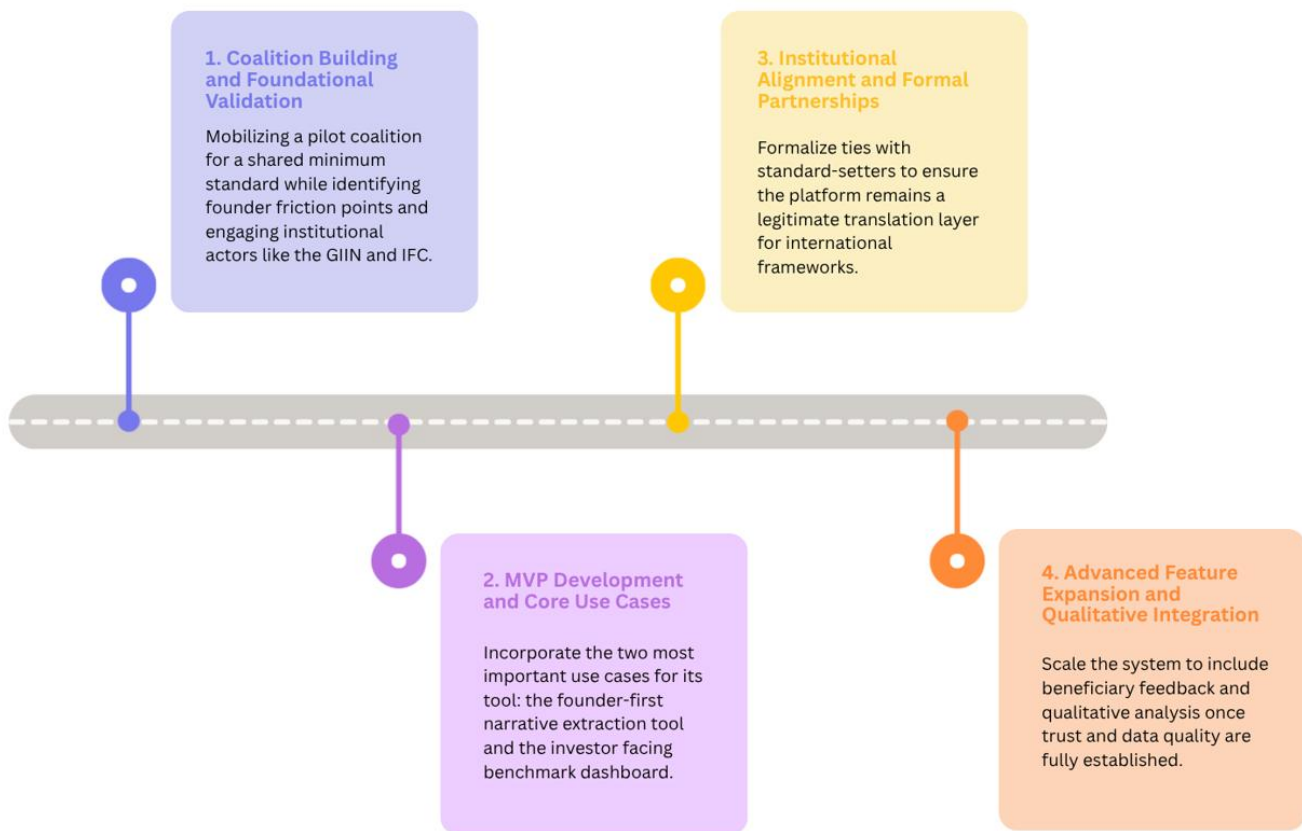


Figure 8 - Implementation roadmap: The figure outlines the proposed phased roadmap for developing AI as an operational support layer. Source: Created by the authors.

In the third phase, IIS should deepen and formalise partnerships with established supply-side actors to build credibility and align with major frameworks.

In the fourth and final phase, IIS could consider adding relevant features like qualitative data processing or beneficiary feedback integration. These features also have strong potential to be transformative (Audrey Selian described beneficiary feedback as “*completely game-changing*”). These can be expanded only once enough trust, data quality and user demand is established.

Taken together, these recommendations for IIS propose actions that have not yet been attempted by the field: leading from the demand side rather than the supply side. The field is not short on frameworks, standards or other technical tools. What is lacking is an actor that has the network and capacity to change what investors collectively ask of startups.

VII. Conclusion

This study aimed at assessing whether AI could support the harmonisation of impact measurement frameworks. The literature often treats harmonisation as a matter of aligning frameworks and improving interoperability. However, what the research ultimately found is that the question itself required reframing. Indeed, in addition to a persistent conceptual ambiguity around the term “impact”, findings also show that the barriers to harmonisation are not primarily technical, but institutional and incentive-based.

The current major frameworks share broad goals and common references, yet they differ in scope and function and are rarely applied as one single unified system in practice. Stakeholders’ priorities are indeed diverging sharply. On one side, investors remain primarily finance-first, using impact reporting as a trust signal rather than an input into decision-making. On the other side, social enterprises prioritise lighter measurement approaches due to resource constraints. These dynamics are further reinforced by a persistent North-South asymmetry in data availability and framework design.

As mentioned, the main barrier to standardisation lies less in the absence of technical tools than in the limited incentives to use them rigorously, and the lack of a shared understanding of impact. Another core obstacle is the persistence of what this study identifies as a “*credibility-sufficiency equilibrium*”, in which just enough measurement happens to maintain credibility, but no actor is incentivised to go further.

Furthermore, this also reframes the AI question. Methods such as NLP, data classification, and cross-framework mapping can support reporting automation and framework navigation, but only where data quality and governance structures are already sufficiently robust. An AI-based platform should therefore not be designed as a harmonisation system, but as a support layer which will reduce reporting burden, help users navigate existing frameworks, and process qualitative data more efficiently.

To conclude, this study emphasises that better tools alone will not transform the field. The more urgent need is leadership from key institutional actors capable of creating demand for a shared minimum standard and convening investors around it in order to raise the credibility floor. That coordination is the precondition, and an AI-based platform comes as the second step. More rigorous and comparable impact measurement will ultimately depend not on better technology, but on stronger incentives from regulators, institutional investors, and multilateral actors.

VIII. Bibliography

- Agrawal, Anirudh, and Kai Hockerts. (2021). 'Impact Investing: Review and Research Agenda'. *Journal of Small Business & Entrepreneurship* 33, no. 2: 153–81. <https://doi.org/10.1080/08276331.2018.1551457>.
- Alshi, Dr CA Shilpa. (2025). 'Artificial Intelligence, ESG, And the Future of Financial Reporting'. *Journal of Marketing & Social Research* 2 (July): 442–46. <https://doi.org/10.61336/jmsr/25-05-49>.
- Arena, Marika, Giovanni Azzone, and Irene Bengo. (2015). 'Performance Measurement for Social Enterprises on JSTOR'. Accessed 4 April 2026. <https://www.jstor.org/stable/43654699?seq=1>.
- Arosio, M. (2011). 'Impact investing in emerging markets'. *Global Impact Investing Network*. <https://thegiin.org/publication/research/impact-investing-in-emerging-markets/>
- Badmus, Oriyomi, Olumide Johnson Ikumapayi, Rebecca Olubunmi Toromade, and Abiodun Sunday Adebayo. (2025). 'Integrating AI-Powered Knowledge Graphs and NLP for Intelligent Interpretation, Summarization, and Cross-Border Financial Reporting Harmonization'. *World Journal of Advanced Research and Reviews* 27 (1): 042–062. <https://doi.org/10.30574/wjarr.2025.27.1.2517>.
- Boiardi, Priscilla, and Esme Stout. (2021). 'Towards Harmonised Management and Measurement of Impact: The Experience of Development Finance Institutions'. *OECD Development Perspectives*, ahead of print, December 21. <https://doi.org/10.1787/6c9891e9-en>.
- Boyatzis, Richard E. (1998). *Transforming Qualitative Information: Thematic Analysis and Code Development*. Thousand Oaks, CA: Sage Publications.
- Brest, P., and Born, K. (2013). When can impact investing create real impact? [https://sdg.tic.go.tz/wp-content/uploads/2021/12/When-Can-Impact-Investing-Create-Real-Impact .pdf](https://sdg.tic.go.tz/wp-content/uploads/2021/12/When-Can-Impact-Investing-Create-Real-Impact.pdf)
- Braun, V., and Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Busch, T., Bruce-Clark, P., Derwall, J., Eccles, R., Hebb, T., Hoepner, A., Klein, C., Krueger, P., Paetzold, F., Scholtens, B., & Weber, O. (2021). Impact investments: a call for (re)orientation. *SN Business & Economics*, 1(2), 33. <https://doi.org/10.1007/s43546-020-00033-6>
- Chen, Michael. (2025). "What Is an API (Application Programming Interface)?" <https://www.oracle.com/europe/cloud/cloud-native/api-management/what-is-api/>
- Clarkin, John E., and Carole L. Cangioni. (2015). "Impact Investing: A Primer and Review of the Literature". *Entrepreneurship Research Journal* 6, n. 2. <https://doi.org/10.1515/erj-2014-0011>.

- Cojocaru, Ana-Carolina, Svetlana Mihăilă, Veronica Grosu, and Ludmila Frumusachi. (2024). 'Enhancing Integrated Reporting Processes through the Integration of Artificial Intelligence into Management Systems'. *Human versus Machine: Accounting, Auditing and Education in the Era of Artificial Intelligence*, 102.
- Crucke, Saskia, and Adelién Decramer. (2016). 'The Development of a Measurement Instrument for the Organizational Performance of Social Enterprises'. *Sustainability* 8 (2): 161. <https://doi.org/10.3390/su8020161>.
- Davenport, Thomas H., and Jeanne G. Harris. (2007). 'Competing on Analytics: The New Science of Winning'. *Harvard Business Review Press, Language* 15 (217): 24.
- De Villiers, Charl de, Ruth Dimes, and Matteo Molinari. (2024). 'How Will AI Text Generation and Processing Impact Sustainability Reporting? Critical Analysis, a Conceptual Framework and Avenues for Future Research'. *Sustainability Accounting, Management and Policy Journal* 15 (1): 96–118. <https://doi.org/10.1108/SAMPJ-02-2023-0097>.
- Ebrahim, Alnoor, and V. Kasturi Rangan. (2014). 'What Impact? A Framework for Measuring the Scale and Scope of Social Performance'. *California Management Review* 56, no. 3: 118–41. <https://doi.org/10.1525/cmr.2014.56.3.118>.
- Ehrlinger, Lisa, and Wolfram Wöss. (2016). 'Towards a Definition of Knowledge Graphs'. *Paper presented at International Conference on Semantic Systems*. <https://www.semanticscholar.org/paper/Towards-a-Definition-of-Knowledge-Graphs-Ehrlinger-W%C3%B6ss/b18e4272a7b9fa2e1c970d258ab5ea99ed5e2284>.
- Eckerle, Christin, and Orestis Terzidis. (2025). 'From Ambition to Evidence: A Practical Tool for Startup Impact Assessment'. *Small Business Economics*, ahead of print, August 4. <https://doi.org/10.1007/s11187-025-01101-7>.
- Elhady, Abdelghafar M., and Samaa Shohieb. (2025). 'AI-Driven Sustainable Finance: Computational Tools, ESG Metrics, and Global Implementation'. *Future Business Journal* 11 (1): 209. <https://doi.org/10.1186/s43093-025-00610-x>.
- Enterprise Knowledge Team. (2020). 'What's the Difference Between an Ontology and a Knowledge Graph?' *Enterprise Knowledge*, <https://enterprise-knowledge.com/whats-the-difference-between-an-ontology-and-a-knowledge-graph/>.
- Feor, Leah, Amelia Clarke, and Ilona Dougherty. (2023). 'Social Impact Measurement: A Systematic Literature Review and Future Research Directions'. *World* 4, no. 4: 816–37. <https://doi.org/10.3390/world4040051>.
- Ferreiro, Micaela, and Julian Kölbel. (2025). 'Impact Measurement Harmonization: Challenges and Opportunities'.

https://fsi.unisg.ch/fileadmin/user_upload/HSG_ROOT/Institut_FSI/Dokumente/Sustainable_Finance/Final_Impact_Measurement_Harmonization_report.pdf.

Gibson Brandon, Rajna Nicole, Melita Leousi, and Camilo Mondragon-Velez. (2025). 'Towards an Impact Performance Measurement Approach for Impact Investing: Results from a Benchmarking Study for Credit Finance'. *Sustainability* 17, no. 23: 10431. <https://doi.org/10.3390/su172310431>.

Global Impact Investing Network (GIIN). n.d. *IRIS+ | Standards*. Accessed November 30, 2025. <https://iris.thegiin.org/standards/>.

Global Reporting Initiative (GRI). (2022). *About GRI – Brochure 2022*. Accessed November 30, 2025. <https://www.globalreporting.org/media/wmxlklms/about-gri-brochure-2022.pdf>

Godeke, S., & Briaud, P. (2020). *Impact Investing Handbook: An Implementation Guide for Practitioners*. Rockefeller Philanthropy Advisors.

Guterman, Alan S. (2024). 'Impact Measurement and Reporting'. *SSRN Scholarly Paper No. 3834003*. *Social Science Research Network*, August 24. <https://doi.org/10.2139/ssrn.3834003>.

Hand, D., Sunderji, S., Pardo, N. (2023) *GIINSight 2023: Impact Measurement & Management Practice*. The Global Impact Investing Network (GIIN). New York.

Hebb, T. (2013). 'Impact investing and responsible investing: What does it mean?' *Journal of Sustainable Finance and Investment*, 3(2), 71–74. <https://doi.org/10.1080/20430795.2013.776255>

HIPSO. (2021). *Joint Impact Indicators (JII)*. <https://indicators.ifipartnership.org/indicators-back/joint-impact-indicators-jii/>

Höchstädter, A.K., & Scheck, B. (2015). 'What's in a name: An analysis of impact investing understandings by academics and practitioners'. *Journal of Business Ethics*, 132(2), 449–475. <https://doi.org/10.1007/s10551-014-2327-0>

Hussien, Lina Fuad, Khaled Alshaketheep, Hind Al-Ahmed, et al. (2025). 'The Impact of Big Data Analytics on the Sustainability Reports Quality'. *Discover Sustainability* 6 (1): 776. <https://doi.org/10.1007/s43621-025-01678-9>.

IFRS. International Financial Reporting Standards. (2025). 'IFRS - SASB Standards'. Accessed 30 November 2025. <https://www.ifrs.org/issued-standards/sasb-standards/>.

IFC. International Finance Corporation. (2019). 'Investing for Impact: Operating Principles for Impact Management'. Washington, D.C.: International Finance Corporation / The World Bank.

Impact Frontiers. (2022). 'Impact Management Norms'. Accessed November 30, 2025. <https://impactfrontiers.org/norms/>.

- Impact Frontiers. (2024). *Impact Performance Reporting Norms, Version 1*. 12 April 2024. <https://impactfrontiers.org/work/impact-performance-reporting-norms-version-1/>.
- International Finance Corporation (IFC). (2021). 'Investing for Impact: The Global Impact Investing Market 2020'. Accessed 8 April 2026. <https://www.ifc.org/en/insights-reports/2021/impact-investing-market-2020>.
- Impact Investing Solutions AG (IIS). (2026). About. <https://www.iisolutions.ch/about>.
- International Telecommunication Union (ITU). (2024). 'AI for Good: Impact Report'. Geneva: ITU. <https://veri-media.io/wp-content/uploads/2024/12/ai-for-good-deloitte.pdf>.
- McCallum, S., and Viviers, S. (2020). 'Exploring key barriers and opportunities in impact investing in an emerging market setting'. South African Journal of Business Management. https://www.researchgate.net/publication/343286215_Exploring_key_barriers_and_opportunities_in_impact_investing_in_an_emerging_market_setting
- Malhotra, Naresh K., and David F. Birks. (2003). 'Marketing Research: An Applied Approach'. *Pearson Education*.
- Mayegun, Kabirat Olamide, and Chinonso Nwanevu. (2025). 'Harnessing Big Data and AI to Revolutionize Sustainability Accounting and Integrated Corporate Financial Reporting'. https://www.researchgate.net/profile/Chinonso-Nwanevu-2/publication/392314120_Harnessing_Big_Data_and_AI_to_Revolutionize_Sustainability_Accounting_and_Integrated_Corporate_Financial_Reporting/links/683ceef1054b0207f8d6ca/Harnessing-Big-Data-and-AI-to-Revolutionize-Sustainability-Accounting-and-Integrated-Corporate-Financial-Reporting.pdf.
- Mosley, Layna. (2013). 'Introduction. "Just Talk to People"? Interviews in Contemporary Political Science'. *Interview Research in Political Science*, 1–30.
- Narain, S., Schmidt, J., Geglio, A., Gelfand, S., & Pease, M. (2012). 'Collaborating to harmonize standardized metrics for impact investors'. *National Community Investment Fund (NCIF)/Global Impact Investing Network (GIIN), Chicago, IL/New York, NY*.
- O'Donohoe, N., Leijonhufvud, C., & Saltuk, Y. (2010). *Impact Investments: An Emerging Asset Class*. J.P. Morgan Global Research.
- Ogunfowora, O. A. (2019). *Social impact measurement and impact investing for social enterprises (Doctoral dissertation)*. University of Manchester, Manchester, UK.
- Ojonugwa, Benjamin Monday, Oluwasanmi Segun Adanigbo, Bisi Ogunwale, and Akoche Ochefu. (2024). 'Impact Investing and AI: Advancing Developmental Goals through Data-Driven Investment Strategies'. https://www.researchgate.net/profile/Oluwasanmi-Adanigbo/publication/392368054_Impact_Investing_and_AI_Advancing_Developmental_Goals

[Is through Data-Driven Investment Strategies/links/6840dc09df0e3f544f5cf90c/Impact-Investing-and-AI-Advancing-Developmental-Goals-through-Data-Driven-Investment-Strategies.pdf.](https://www.researchgate.net/publication/281181122)

Ormiston, Jarrod; Kylie Charlton; M. Scott Donald; and Richard G. Seymour. (2015). 'Overcoming the Challenges of Impact Investing: Insights from Leading Investors.' *Journal of Social Entrepreneurship*

<https://www.researchgate.net/publication/281181122> Overcoming the Challenges of Impact Investing Insights from Leading Investors

Ottun, Abdul-Rasheed Olatunji, and Huber Flores. (2025). 'Trustworthy AI in Practice: A Comprehensive Review of Human Oversight and Human-in-the-Loop Approaches'. *Authorea Preprints*. <https://www.techrxiv.org/doi/full/10.36227/techrxiv.176118749.93102582>.

Popov, Peio. (2024). 'The Power of Ontologies and Knowledge Graphs: Practical Examples from the Financial Industry'. *Graphwise*. Accessed 30 November 2025. <https://graphwise.ai/blog/the-power-of-ontologies-and-knowledge-graphs-practical-examples-from-the-financial-industry/>.

PRI. n.d. 'About Us | The PRI'. n.d. Accessed 30 November 2025. <https://www.unpri.org/about-PRI>.

Ram, Nitin. (2025). 'Integrating Blockchain and Agentic AI for Sustainability Reporting in Port Operations: PISA-Based Conceptual Framework'. *SSRN Scholarly Paper No. 5511996*. *Social Science Research Network*. <https://doi.org/10.2139/ssrn.5511996>.

SDGs. (2025). United Nations Department of Economic and Social Affairs. n.d. *About | Department of Economic and Social Affairs*. Accessed November 30, 2025. <https://sdgs.un.org/about>

Singhania, Monica, and Deepika Swami. (2024). 'Impact Investing: Scientometric Review and Research Agenda'. *Business Ethics, the Environment & Responsibility* 33 (3): 251–86. <https://doi.org/10.1111/beer.12599>.

Sustainability Accounting Standards Board (SASB). n.d. *About Us*. Accessed November 30, 2025. <https://sasb.ifrs.org/about/>

Tansey, Oisin. (2007). 'Process Tracing and Elite Interviewing: A Case for Non-Probability Sampling'. *PS: Political Science and Politics*, ahead of print. https://doi.org/10.1007/978-3-531-91826-6_24.

The Government Outcomes Lab. n.d. 'GIIRS'. Accessed 30 November 2025. <https://golab.bsg.ox.ac.uk/toolkit/wayfinder-assessment-resources/INDIGO-ARES-0064/>.

TNFD. (2025). 'Disclosure Recommendations – TNFD'. Accessed 30 November 2025. <https://tnfd.global/recommendations/>.

- TNFD. (2025). 'The Taskforce on Nature-Related Financial Disclosures'. Accessed 30 November 2025. <https://tnfd.global/>.
- UNDP. United Nations Development Programme. (2023). 'About the SDG Impact Standards'. Updated May 2023. <https://sdgimpact.undp.org/>.
- Vallejo, Bertha, and Wehn Uta. (2016). 'Capacity Development Evaluation: The Challenge of the Results Agenda and Measuring Return on Investment in the Global South'. *World Development* 79: 1–13. <https://doi.org/10.1016/j.worlddev.2015.10.044>.
- Wilhelmi, Lars, Robert Nickerson, and Matthias Schumann. (2024). 'Insights from the Alphabet Soup: Characterizing AI-Based Methods for the Analysis of Sustainability Reports'. *ECIS 2024 Proceedings*. https://aisel.aisnet.org/ecis2024/track17_greenis/track17_greenis/32.
- Wood, Elizabeth. (2007). "Field Research" in Carles Boix and Susan Stokes (eds.), *The Oxford Handbook of Comparative Politics*. Oxford: Oxford University Press. pp. 123-146.
- Yau, Alban. (2024). 'How Can Impact Investors Enable Systems Change? Exploring the Theory and Practice of an Emerging Field'. *SSRN Scholarly Paper No. 4707922. Social Science Research Network*. <https://doi.org/10.2139/ssrn.4707922>.

IX. Annexes

Annex 1: Main impact and sustainability standards and their characteristics

Standard	Focus	Description	Institution
IFC Operating Principles for Impact Management (IFC Principles) or Operating Principles for Impact Management (OPIM)	Principles/rules on integrity and governance	A set of nine principles that define good practices for the impact management cycle. It guides intention, management, monitoring, and exit.	International Finance Corporation (IFC), World Bank Group
UN Principles for Responsible Investment (PRI)	Principles/rules on responsible investment processes	It is composed of six voluntary principles aimed at guiding investors to integrate ESG factors into their decisions.	United Nations with institutional investors
UN Sustainable Development Goals (SDGs)	Global Sustainability Goals	A set of 17 global sustainable development goals that are part of the 2030 Agenda for Sustainable Development, adopted by all United Nations Member States in 2015. They serve as a guide for setting impact priorities.	United Nations
SDG Impact Standards	Operational Management Standard	A compilation of management standards to align investments with the SDGs. They help integrate the SDGs into organisational governance and strategy.	United Nations Development Programme (UNDP)
Impact Management Norms	Convention/ Management Principles	Norms showing how to measure, improve, and disclose the impacts of companies and investors	Impact Management Project (IMP) and Impact Frontiers.
Impact Reporting and Investment Standards (IRIS+)	Standardised Language/Metrics	Catalogue of metrics for measuring impact. Provides standardised metrics, core metric sets, and alignments with SDGs.	Global Impact Investing Network (GIIN)
Sustainability Accounting Standards Board (SASB)	Sustainability Accounting	A set of sector-specific disclosure standards (77	International Financial Reporting Standards

	Standards	sectors). Focus on financial materiality (what affects financial value).	(IFRS) Foundation
GRI Standards	Standard for disclosure of sustainability impacts	Sustainability reporting standards focused on the impact that companies have on society, the environment, and the economy.	Global Reporting Initiative (GRI)
Global Impact Investing Ratings System (GIIRS)	Fund/Company Rating System	An impact assessment and rating system for companies and funds	B Lab
Harmonized Indicators for Private Sector Operations (HIPSO)	Harmonisation of output-level metrics for Development Finance Institutions (DFIs)	It is a catalogue that aims to reduce the administrative burden on investee companies that need to report data under different metrics. Currently, the catalogue has more than 150 indicators divided into 21 sectoral and thematic categories.	Coalition of Development Finance Institutions (DFIs)
Task Force on Climate-related Financial Disclosures (TCFD)	Disclosure of climate-related financial risks	Reporting framework for companies and investors to disclose how climate change financially impacts the organisation (financial materiality)	International Financial Reporting Standards (IFRS) Foundation
Taskforce on Nature-related Financial Disclosures (TNFD)	Nature-related risk framework	Recommendations for companies to assess, report, and act on nature-related impacts. Aims to steer financial flows towards results that are positive for nature.	TNFD

Source: Based on IFC, 2019; HIPSO, 2021; Impact Frontiers, 2022; Global Impact Investing Network; SASB; GRI, 2022; UNDP, 2023; PRI; SDGs, 2025; The Government Outcomes Lab; TNFD; TCFD, 2025.

Annex 2: Methodology

Research Methodology

Research design

This research adopts an exploratory qualitative design to investigate whether and under which conditions the currently fragmented landscape of impact measurement frameworks could be better harmonised throughout the impact investing field. Knowing the complexity and variety of actors involved in this field, the study combines a structured literature review with primary qualitative interviews. This design aims not only at mapping the existing frameworks and debates around them, but also to comprehend how impact measurement is perceived, adapted and applied in practice by the various actors.

The first stage of the study includes a comprehensive review of the relevant academic and grey literature to grasp the fundamental aspects, tensions and similarities from both perspectives. This thorough review was necessary to establish the conceptual and empirical foundations of the study before engaging in primary data collection. More specifically, it sought to clarify the evolution of impact investing, pinpoint the key actors involved in the field, compare the major impact measurement frameworks currently in use, and evaluate the main critiques in the literature regarding fragmentation, comparability, and standardisation. In parallel, another body of literature was reviewed on artificial intelligence and data-processing tools, with particular attention to the types of AI capabilities that may be relevant for metric alignment, validation, interoperability, and reporting.

The literature review was not treated as a separate descriptive exercise, but as a foundational component of the research design and it fulfilled several purposes. Firstly, it helped to map the main frameworks and their underlying logics. Second, it allowed us to identify persistent tensions in the field, including the disparity between frameworks developed and actual organisational practices, the diversity in stakeholders' interests, and the geographical and sectoral variation that makes standardisation efforts more arduous. Third, it helped provide the basis for refining the research questions and elaborating the interview guide for primary data collection.

Data collection

Since the research questions are primarily exploratory and related with practices, perceptions, motivations, and barriers, qualitative semi-structured interviews were selected as the most appropriate method for collecting primary data. Rather than relying only on literature and formal documentation, this research adopts a field research logic and includes key stakeholders in the field. This allows the study to capture perspectives and practices that are not always visible in published or official sources (Wood, 2007; Mosley, 2013). Interviews are especially useful in such contexts because they can reveal how actors understand their actions, how decisions are actually made, and what informal processes may not be visible in other sources (Mosley, 2013).

The primary data collection consists of 15 semi-structured interviews with stakeholders from across the impact investing ecosystem. The sample includes actors such as investors, social enterprises,

intermediaries and advisory actors involved in impact measurement practices. The objective is not to obtain a statistically representative sample of the entire field, but to capture a diversity of informed perspectives from actors who engage with impact measurement in different ways and from different institutional positions.

Sampling strategy

Interviewee selection has therefore followed a purposive, non-probability sampling strategy. This is appropriate because the aim of the study is not to estimate the distribution of opinions across a larger population, but to gather insightful and relevant evidence from actors who are particularly well placed to speak to the research questions (Tansey, 2007). Initial participants were selected on the basis of their professional role, expertise, and relevance to the issues under study. This was also complemented by snowball sampling in order to identify additional interviewees with specialised knowledge or practical experience. However, there are limitations about the richness of perspective since initial actors we interviewed may recommend others who are similar to themselves, which ultimately may have narrowed the range of perspectives included in the sample (Tansey, 2007). Nevertheless, such a sampling approach enables the study to access informed actors across different positions and sectors.

The interviews explored several core dimensions of the project. These include:

- how organisations currently measure and report impact;
- which frameworks or metrics they use, adapt, or ignore
- how impact is perceived among them;
- what technical, organisational, or practical barriers they encounter;
- how their priorities vary by sector, geography or institutional position;
- and how they evaluate the potential contribution, risks, and limits of AI-based solutions for harmonisation, although this dimension was explored only in a limited and non-technical way.

This methodological approach nevertheless has important limitations. First, the study does not develop or test a concrete AI tool, and the AI dimension is still exploratory rather than technical. Second, the interview sample remains relatively small, mainly due to time and access constraints. As a result, the findings are not intended to be statistically generalisable to the entire impact investing ecosystem. Rather, they provide analytically useful insights into recurring patterns, contrasting perspectives, and key challenges related to impact measurement and standardisation.

Data Analysis Methodology

Analytic framework and overall approach

The interview data we collected were analysed through a hybrid thematic analysis that follows the six-phase framework set up in Braun & Clarke (2006). These phases include: familiarisation with the data, generation of initial codes, search for themes, review of themes, definition and naming of themes, and production of the report. We chose thematic analysis as our primary analytical method given its

recognition as rigorous qualitative research method and its alignment with an exploratory approach like ours where the aim is to surface key themes across a diverse range of survey responses rather than testing pre-specified hypotheses (Braun & Clarke, 2006; Boyatzis, 1998). Given that our research hones in on a small, intentionally selected sample of practitioners from diverse stakeholder groups, our priority was analytical depth and precision in our interpretation rather than statistical inference or sentiment analysis.

The qualitative analysis chosen was deliberately hybrid, incorporating both deductive and inductive codes. From our research questions and existing literature, we derived an initial set of 25 codes across 6 broad areas of interest (knowledge and use of frameworks, engagement with and methodologies for measurement, investor attitudes, barriers faced, opinions on an openness to AI use, and appetite for standardisation). At the same time, the coding process remained open to patterns emerging directly from the engagement with the data, with inductive codes generated when the interview material pointed to new dynamics or points of contention that we had previously ignored. These included the tension between measuring for internal operational learning and measuring for external accountability/marketing, or issues around impact attribution when impact was indirect, cascading or co-created. This combination was essential for two reasons. First, a purely deductive approach would have risked only confirming what the academic literature already suggested rather than leaving room for the perspectives of practitioners. Second, a purely inductive approach might have risked a lack of coverage and analysis around certain research questions, especially given the diversity of our sample and the breadth of the topics covered in interviews.

Consistent with guidance around exploratory qualitative analysis, we were conscious to avoid our analytical frame becoming too enumerative or inferential. Our guiding questions at each stage of the coding stage was what a certain passage means and reveals about how impact measurement functions in practice rather than how often certain statements or sentiments appear.

Another methodological decision that shaped the overall approach was treating divergent and unexpected findings as significant from the outset, rather than trying to explain them away. This was particularly important, given that several interviewees very early on described practices that differed from what the literature predicts and were even often actively critical or challenging of the premises of the research questions themselves. The inductive codes generated as a result — including, for example, the finding that impact measurement is perceived primarily as a legitimacy and trust signal rather than as a key crux in decision making, or that the AI-for-harmonisation framing may be answering the wrong question entirely — ultimately turned out to be some of the most significant and productive elements of our analysis and form a central part of our arguments in the discussion section.

Phase One: Familiarisation with the data

All fifteen interviews were conducted as semi-structured conversations, audio-recorded with participant consent and transcribed verbatim before analysis. We reviewed transcripts against the original recordings to correct transcription errors before any coding began. Each team member read the transcripts in full to ensure familiarisation before any coding began.

Phase Two: Developing the coding framework

After familiarisation, we constructed a coding framework of inductive and deductive codes. We established six deductive categories a priori from the research questions and the literature:

- Framework awareness and actual usage
- Measurement practices
- Perceived investor expectations
- Barriers to standardisation
- Attitudes toward AI
- Attitudes toward a common standard

For each of these categories, we used AI to come up with sub-codes which were all reviewed in detail prior to coding the interviews.

The barriers category, for instance, was divided into cost constraints, incentive misalignment, data availability, verification limitations, reporting fragmentation, definitional ambiguity, and political obstacles to support more precise codings and nuanced findings.

To this came 16 inductive codes which were arrived at during the coding process that had not been previously anticipated.

Phase Three: Coding the transcripts

Transcripts were coded at the level of the meaning unit, which typically included one to five sentences that express a complete and self-contained claim. Transcripts were read and assigned codes for relevant themes and topics that emerged in the interview. For each code assigned, we extracted a verbatim extract of what was said by the interviewee and added a commentary or interpretive note to ensure deeper analysis. This allowed us to go beyond what the interviewee said and reveal something further around the structure of the field.

To support the coding process, we used AI as a first-pass tool for organising and surfacing relevant passages across the transcripts and interviews. All AI-generated outputs were reviewed, questioned and revised where necessary, cross-checking coded passages against transcripts. The inductive codes that emerged in particular were arrived at independently by members of the team through a close reading of the material.

The inductive codes that surfaced in this process were:

- Impact reporting functioning primarily as a legitimacy signal rather than a decision input
- A systematic perverse incentive to overstate rather than accurately measure impact
- The conflation of output metrics with genuine impact
- The absence of beneficiary voice from metric design and validation

- The asymmetry with which investees bear the cost of measurement while investors capture the reputational and compliance benefits
- The risk that rigid metrics trap organisations in inappropriate accountability frames and can actively harm effective work
- Finance-first optimisation as the core logic of investment decisions regardless of stated impact commitments
- The defence of qualitative evidence as methodologically appropriate rather than merely a default in certain social contexts
- Impact measurement being designed retrospectively rather than at the beginning of investment cycles
- Peer pressure and social conformity as the primary, and weakening, driver of impact reporting in the absence of stronger incentives
- The political economy of impact investing, in which external political shifts impact what impact is valued, funded, and measured
- The attribution and ripple-effect problem where indirect and cascading impact that cannot be assigned to any single actor
- The deliberate substitution of observable business metrics for unobservable impact as a rational response to verification constraints
- The absence of shared baseline data, making it structurally impossible to assess whether any intervention produced change
- The tension between measuring for internal operational learning versus external accountability
- Platform and reporting fatigue from managing multiple incompatible investor reporting templates

Each passage was also assigned a flag that reflects its impact on our thesis, marking something either as supporting, challenging or adding nuances to our initial assumptions as well as highlighting generally surprising evidence.

Phase Four: Development and review

After all transcripts were coded, the coding matrix was reviewed to identify key convergences, divergences, and any notable exceptions. We then developed themes by clustering codes that described related concepts within a single broader category. For instance, the codes capturing framework non-use, performative engagement with measurement, and peer pressure as the primary reporting driver were grouped under the single overarching pattern of formal frameworks not being adopted in operations. This cluster served as the base of a deeper empirical discussion around the assumptions underlying our research questions.

For themes that were brought up in only a few instances, we gave them particular analytical attention to understand how they fit into the dominant pattern rather than automatically sidelining the findings

as irrelevant.

Phase Five: Connecting empirical findings to the literature

After developing and reviewing themes, we mapped these themes against our existing findings in the literature to understand whether the existing literature had fully anticipated it, partially anticipated it, or not anticipated it at all. Findings confirmed by multiple interviewees and consistent with the literature, such as fragmentation of reporting infrastructure, resource constraints, and the dominance of financial considerations in investment decisions, were treated as contextualisation and replication of the literature. Findings that departed from or complicated the literature were treated as the most important contributions of our primary data. In particular, the risk that standardisation may lock organisations into harmful accountability frames, and the argument that AI-assisted harmonisation may be addressing the wrong problem entirely, were findings that diverged from the existing literature and therefore received further attention in the Discussion.

Limitations of the approach

Several limitations should be taken into account when interpreting these findings.

First, the sample size, fifteen interviews, was appropriate for an exploratory qualitative study but constrains the range of perspectives captured. Despite the intentional sampling design, certain actor types are underrepresented: founders only represent three of the interviews, while no interviewee had a dedicated background to technology or AI. Findings therefore might be treated as generalisable analytically, especially around the mechanisms and dynamics of the field but should not be understood as statistically representative of the impact investing ecosystem as a whole.

Second, while the initial pool of potential interviewees was identified through a mapping of relevant actors across the impact investing ecosystem, access to participants ultimately often relied on the personal and professional network of members of the research team. This means that the sample, while diverse, was not arrived at through a systematic or replicable procedure. Actors outside these networks may be underrepresented.

Third, the use of AI-assisted tools to support the initial organisation of coded material represents a risk that pre-processing influenced the framing of certain passages. While the AI-generated outputs were reviewed and revised in detail by the research team, that possibility cannot be ruled out.

Annex 3: Interview guiding questions

List of questions for investors

1) Understand who they are and what they invest in

Could you tell me a bit about your organisation/role and what your organisation invests in?

- What kind of tickets / stage / sector do you focus on?
- Who do you usually report impact results to (i.e. LPs, a board, regulators, someone else?)

2) A real example

Can you think of a recent deal where impact information actually mattered and influenced the outcome? Like it influenced the decision in some way.

- What specific impact information did you look at?
- What was missing or unclear?
 - When it's missing, what do you do? Do you ask again, use proxies, rely on judgement, or whatever else?

3) Why they measure impact

What are the main reasons you measure impact?

- If they give several reasons: If you had to rank them, which one or ones are really the most important?

4) What they ask companies for

When you ask an investee for impact information, what are the top 2-3 things you want to know?

- Do you have minimum requirements (mandatory indicators or evidence)?
- How do you decide if results/data are good enough? Is it thresholds, trends, benchmark,...?

5) Current workflow and where it becomes difficult

Can you walk me through how impact data flows today? From the company, to you, to reporting and decision-making?

- Where do things get painful or slow?
- What's the hardest part (i.e. data quality, comparability, verification, time, cost...)?
- If you could fix one step in that chain, which step would you fix first?

6) Understand the gap

Where do you see the biggest gap between what companies can realistically report, and what you need as an investor?

- What do you ask for that you know is a heavy burden for them?
- If an early-stage company could only report a small "minimum pack", what should be in it? What is the most valuable to you?

7) Which frameworks and standards they use

Do you use any standards or frameworks for impact reporting? If yes, what do you use each one for?

- Do they overlap or create duplication?
- Which one/ones is/are really helpful, and which parts feel like box-ticking?

8) Standardisation

Imagine there was a small common core of metrics that many investors accept with optional modules depending on sector or stage you invest in.

Would that help you or would it create more problems?

- What absolutely must be in a “core”?
- What should never be standardised because it’s too context-specific?
- What would make you actually adopt something like this?

9) AI/technology

Where would you feel comfortable using tech or AI to support impact measurement?

- What use cases feel acceptable to you? (i.e. extracting indicators from reports, mapping frameworks, basic quality checks,...)
- Which use cases are an absolute “no”? And why?
- If you’re sceptical, what would you need to trust it (i.e. transparency, human input, data privacy, etc.)?

Optional if needed

- Is there anyone else you think we should speak to?

List of questions for entrepreneurs / social enterprises / investees

1) Understand who they are and what they do

Could you tell me a bit about your organisation and your role?

- What problem are you trying to solve, and who are your main customers?
- What stage are you at, and in what geographies?

2) Funding context and investors’ requirements

Have you recently raised funding? What type and from whom?

- What impact reporting do investors typically ask you for?
- Is it the same across investors, or does it change a lot?
- What happens if you don’t have the data they ask for?

3) A real example

Can you think of a recent moment when impact information really mattered? For example in fundraising, due diligence, renewal of a grant, a board discussion, or a partnership?

- What did they ask you for specifically?
- What was easy to provide, and what was hard or unclear?
- Did it change anything (the investment decision, conditions, timeline, your strategy)?

4) Why they measure impact

Beyond external reporting, why do you measure impact?

- Is it mainly for fundraising/reporting, or for other reasons? Which ones?
- If you had to rank your top reasons, what are the top 1–2?

5) What they currently measure

What are the top 2-3 impact indicators you track today?

- How did you choose them? (i.e. investor requirements, SDGs, industry norms,...)
- What do you wish you could measure but currently can't? What are the reasons? (i.e. too expensive, too complex, no data access, etc.)

6) Current workflow and where it becomes difficult

Can you walk me through your workflow, from collecting data, to storing it, analysing it, and reporting it to investors?

- Where does it get painful or slow (time, cost, staff capacity, tools, data quality, etc.)?
- What part is the most stressful or resource-intensive?
- If you could fix one step in the chain, what would you fix first?

7) Understand the gap

Where do you see the biggest mismatch between what investors ask for and what you can realistically report?

- What requests feel reasonable versus unrealistic?
- What requests feel like “box-ticking” rather than useful?
- If you could propose a “minimum pack” you can report consistently, what would it include?

8) Which frameworks and standards they use

Do you use any standards or frameworks for impact reporting? Or are you asked/pressured to align with them?

- Which ones come up most often (i.e. IRIS+, SDGs, GRI, SASB, etc.)?
- Do they overlap or create duplication?
- What is really helpful, and what feels like useless?

9) Standardisation

Imagine there was a small common core of metrics that many investors accept with optional modules depending on sector or stage they invest in.

Would that make your work easier or harder?

- What absolutely must be in the “core” for it to be useful and realistic?
- What should never be standardised because it's too context-specific?
- What would make you adopt it? (i.e. saving time, clearer definitions, acceptance by most investors, simple and user-friendly templates/tools, etc.)?

10) AI/technology

Where would you feel comfortable using technology or AI to support impact measurement and

reporting?

- What use cases feel acceptable (e.g., turning reports into structured data, auto-filling templates, mapping one framework to another, spotting inconsistencies, dashboards,...)?
- Which use cases are an absolute “no”? And why?
- If you’re sceptical, what would you need to trust it (i.e. transparency, human input, data privacy, etc.)?

Optional if needed

- Is there anyone else you think we should speak to?

List of questions for researchers / academics

1) Understand who they are and their perspective

Could you briefly introduce your research area and your current work related to impact investing or impact measurement?

How would you describe your position within the broader impact investing ecosystem (e.g. academic, policy-oriented, practitioner-oriented, interdisciplinary)?

2) General perspective on impact measurement

How do you currently assess the state of impact measurement and management (IMM)?

What do you see as the main strengths and weaknesses of the field today?

3) The definition problem

One issue that repeatedly appears in the literature is the lack of a unified definition of “impact.”

How do you personally define impact in the context of impact investing?

Do you think conceptual fragmentation is:

- a major problem,
- an inevitable feature of the field,
- or potentially even beneficial in some ways?

4) Theory vs practice

To what extent do you see a gap between academic approaches to impact measurement and what practitioners actually do?

Where do you think the disconnect is strongest?

- conceptual clarity,
- methodological rigor,
- data availability,
- reporting pressures,
- or something else?

5) Multiple standards

What is your view on the current ecosystem of frameworks and standards (e.g. IRIS+, SDGs, GRI, IMP,

HIPSO, SASB, etc.)?

Do these standards:

- improve comparability and accountability,
- or create duplication and reporting burden?
- Why do you think so many frameworks continue to coexist?

6) Harmonisation

Do you believe a shared “core set” of impact metrics across stakeholders would be realistic or desirable?

- What should absolutely be standardised?
- What should remain context-specific?
- Who do you think would need to lead such a harmonisation effort? (e.g. governments, DFIs, multilateral organisations, large investors, industry coalitions)

7) Measurement challenges

In your view, what are the biggest methodological or practical challenges in impact measurement today?

8) Regional and sectoral differences

Do you observe important regional differences in how impact is understood or measured?

For example between:

- Global North vs Global South,
- emerging vs developed markets,
- or across sectors?

9) AI/technology

Do you believe technology or AI could help reduce fragmentation in impact measurement?

- Which applications seem most promising? (e.g. mapping frameworks, extracting indicators, automating reporting, quality checks)
- Which limitations or risks concern you the most? (e.g. transparency, bias, governance, over-standardisation)

10) Looking ahead

Looking forward, what changes do you consider most necessary for impact measurement to become more effective and credible?

Optional if needed

- Is there anyone else you think we should speak to?

Annex 4: Interview Reporting

List of the interviewees

- Daniel Nowack - Head of Social Innovation at World Economic Forum (WEF)
- Audrey Selian - Director of Artha Impact (Rianta Capital)
- Juan J. Alarcón - Founder of Swissocial
- Tenke Zoltani - Founder of Better Finance
- Monica Blatytá - Founder of Eqwow Ventures
- Laís Rocha - Head of impact at Vox Capital
- Letícia Vargas - Head of Research, LATAM at the University of St. Gallen
- Carmen Chan - Founder at Girls be Change
- Balca Yilmaz - Founder of WeRover
- Joe Seddon - Founder of Zero Gravity
- Jessica Villanueva - Senior manager at World Wide Fund for Nature (WWF)
- Nathan Sussman - Professor of Economics at the Geneva Graduate Institute
- Evelyn Castle - CEO and Founder of EHA Impact Ventures

Additionally, we interviewed two practitioners that preferred to stay anonymous. One of them is a development finance specialist and the other one is a director of a social impact foundation.

Interviewees profiles

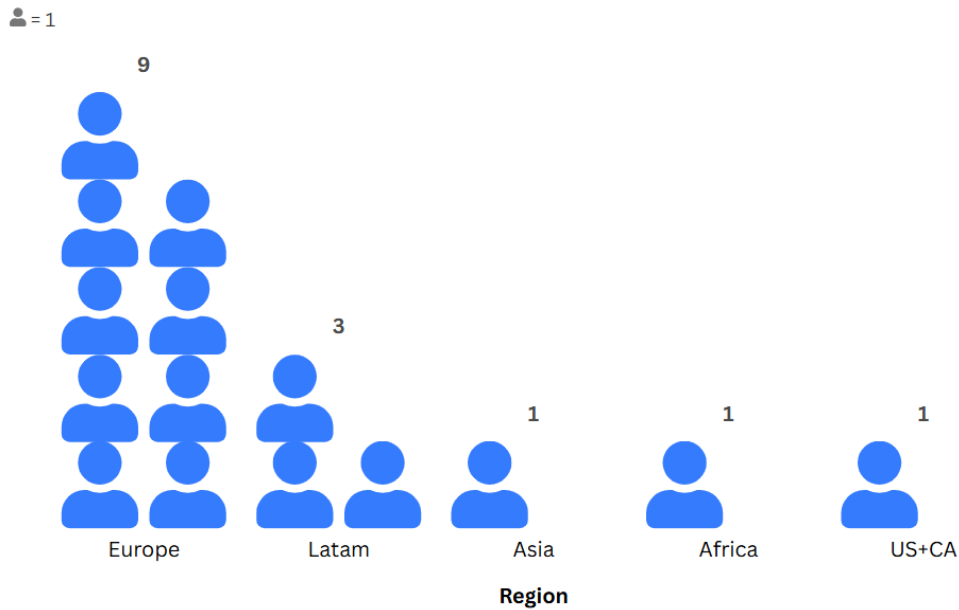


Figure 9 - Interviewees by geographic region. *Source: Created by the authors.*



Figure 10 - Interviewees' gender distribution. *Source: Created by the authors.*

Annex 5: Interview thematic results

The results are organised thematically into four main sections and several subthemes, based on the recurring topics identified in the interview material. For each theme, the findings are presented in a synthesised and structured way, while still reflecting the diversity of perspectives expressed by the interviewees.

Theme 1 - Frameworks in theory and in practice

1.1 Framework awareness is high across the ecosystem

One recurring theme and where there is convergence among interviewees is that major frameworks are widely known across the ecosystem. Interviewees from varied positions in the ecosystem referred to tools or frameworks like the SDGs, IRIS+, Impact Management Project (IMP), ESG standards, and Theory of Change. Tenke Zoltani (Founder of Better Finance), for instance, described mapping investments against SDG indicators before cross-referencing with IRIS+. Some even described other less-known tools like the Five Dimensions of Impact, and B Impact Assessment.

Awareness is not the problem.

1.2 Framework use in practice remains selective, adapted, and fragmented

Despite this awareness, however, frameworks are rarely used as the unified and rigorous systems they were designed to be. Instead, interviewees described a much more selective and adapted use in practice. Artha Impact uses the logic of the Impact Management Project (IMP) and draws on IRIS+ as a large database of possible indicators, but does not follow one single dominant framework strictly. Tenke Zoltani similarly stated that she takes bits and pieces from different references which are ultimately translated into an Excel file. Monica Blatyta (Founder of Eqwow Ventures) also relies on recognised frameworks but she reduces actual reporting to one impact KPI and a few ESG KPIs that startups can realistically manage, citing the practical limitations of early-stage startups: *“I don't believe in long lists, 10 or 15 KPIs won't work for startups”*.

A similar pattern with an additional emphasis was identified in interviews with stakeholders in Latin America. Indeed, Laís Rocha (Head of impact at Vox Capital) said that her organisation has developed internal methodologies since most tools are developed largely in the Global North which does not reflect Brazilian or Latin American contexts. Jessica Villanueva (Senior manager at WWF) also describes WWF as operating with parallel systems, one tied to conservation metrics and the other linked to development-oriented indicators, with ongoing internal work to make the relationship between them clearer. Letícia Vargas (Head of Research, LATAM at the University of St. Gallen) gives a more critical view noting that, in Brazil, Theory of Change is often treated as if it were a measurement standard. Joe Seddon and a development finance specialist also pointed to an asymmetry between environmental and social metrics because, in their view, indicators like CO2 reduction or hectares restored are generally more tractable than social ones.

Theme 2 - Perception of impact in practice

2.1 Confusion between output and impact

A core theme concerns the gap in understanding between what is labelled “impact” and what is actually being measured. A development finance specialist explicitly argues this aspect by stating that many actors claim to measure impact but in reality they mainly track output or outreach indicators such as the number of people served, number of women reached, or the volume of services provided. This is echoed by Daniel Nowack (Head of Social Innovation at World Economic Forum) who rhetorically deconstructed the field’s most common metrics: *“If you count the number of beneficiaries [...] What does that mean? Does that mean people have paid for your service? Does it mean that these people have actually increased their own income? Is it that these people actually have a healthier life? You can't judge anything by just a number of beneficiaries.”* He noted that highly aggregated metrics such as number of beneficiaries are easy to measure and communicate but they lack depth. Tenke Zoltani further added that those output indicators may be useful in practice but provide no direct evidence of how lives of people changed. Audrey Selian (Director of Artha Impact) moved closer to an outcome logic, including income improvement, access and working conditions but, at the same time, she acknowledges that it is still inferred through proxies like turnover or sales volume.

Juan J. Alarcón (Founder of Swissocial), whose organisation has conducted over 80 impact evaluations across 15 countries, said: *“Every time I say, “We, as an organisation, have done this or that,” all those things are results or outputs. When I start my sentence by saying, “They have improved their income” or “They have improved their business,” that is impact.”* He further argued that real impact can only be assessed by looking at how much beneficiaries’ living conditions have changed over time.

This confusion between outputs and impact outcomes is not just a problem of conceptual clarity. For many founders, it is much easier to track outputs such as how many people were reached or how much cheaper it has become for customers to repair wind turbines, over impact outcomes which can be time-intensive and difficult to collect.

2.2 Beneficiary voice is still absent

Several interviews reinforce this same tension from different perspectives. Audrey Selian reiterated several times that the field has not yet been able to integrate a direct feedback loop mechanism that includes beneficiaries’ input, and she described it as *“the big missing link”*. Then, she further argued that such a tool including this feature would be *“completely game-changing”*. Daniel Nowack also expressed the importance of being able to involve them in defining which outcomes matter to them. Carmen Chan (Founder at Girls be Change) rejected quantitative proxies for her work entirely, arguing that metrics such as number of gatherings or participants *“usually cannot really tell the impact,”* and instead focused on stories, feedback, and ripple effects as more meaningful type of evidence.

2.3 The term “impact” carries different meanings across stakeholders

Another finding is the absence of a shared definition of impact in the ecosystem. As highlighted by Joe Seddon (Founder of Zero Gravity), unlike profit and loss, social impact has no accepted pricing

mechanism and no universally trusted basis for comparison. For some interviewees, impact means an organisation that addresses a recognised problem, like Monica Blatyta's use of SDGs to determine investment eligibility. For others, it refers to changes in people's living conditions like Audrey Selian's focus on income improvement and access to essential services. At the same time, others see impact is closer to intention or positive orientation, which is the ambiguity that Letícia Vargas mentions in her doctoral research on the Brazilian impact investing field.

2.4 Different measurement cultures across environmental and social impact

One pattern identified across stakeholder groups is the difference between the maturity of shared standards and investor expectation between the environmental and social impact sectors.

On the environmental side, several interviewees pointed to a more developed measurement culture. Joe Seddon notes that the environmental sector *"has done the best job in terms of moving towards agreed standards"* and cites the progression towards common emission accounting across Tier 1, 2 and 3 categories. Daniel Nowack echoed this, stating that *"on the environmental side, it's probably a bit easier to harmonise,"* grounded in the observation that metrics like carbon reduction or hectares restored are less context-dependent and more amenable to standardised quantification.

Balca Yilmaz (Founder of WeRover) corroborates this with her personal experience as a cleantech founder. Her investors consistently ask for carbon reduction metrics which she believes they do really care about. This is a contrast to the sometimes more performative investor interest in measurement that some social impact founders describe.

This progress in the environmental sector is due to having more quantifiable environmental outcomes. Carbon emissions and biodiversity are measurable and comparable phenomena while improvement in resilience or empowerment are not. These sectors face different challenges, meaning that interventions around the environmental model might not transfer easily.

Theme 3 - Incentives, burdens and barriers

3.1 Financial return remains the core incentive

One strong convergence point is the primacy of financial return even within the impact investing field. Tenke Zoltani described investors she worked with as finance-first, and noted that in one family office case, impact reporting became less important once investors were satisfied with the financial performance of their capital. A development finance specialist stated that *"everything that influences decision in investments is risk and financial considerations"*. Monica Blatyta also said: *"Impact cannot reduce returns, it has to complement them"*. Even in organisations like Artha Impact where impact measurement is at the centre of its investment rationale, financial viability is a binding constraint. Joe Seddon explicitly stated how impact information works in investment decisions: *"Do people make investment decisions by looking at a social impact model and evaluating it rationally versus other organisations? Absolutely not. I can't think of a single instance..."*. In his view, impact reporting works as a trust signal of an organisation rather than the basis for decision-making. Carmen Chan described the same dynamic from her experience with capital providers who tend to see impact as secondary (*"tick, tick, tick [...] now show me the money"*).

3.2 Measurement burden falls on those with less capacity to carry it

One major finding is that impact measurement often creates a practical burden, particularly on the social enterprises side. Daniel Nowack framed measurement as a “*cost center*” and insisted on a paradox within this field. As he suggests, by requiring more reporting than conventional investors, impact investors are effectively raising the cost of capital for the very organisations they aim to support in the first place. Monica Blatyta refers to the same issue by saying: “*If a large company today already spends a fortune hiring [...] analysts to enter data into different platforms*”, or also “*then imagine having to create different spreadsheets and reports for every investor who comes on board, sometimes there are 30 of them*”.

This dynamic is confirmed on the social enterprise side as well. Balca Yilmaz said being requested to provide reporting without templates or clear methodologies (“*They just ask, like, do something. And I do something. I don't know if it is proper or not*”). Carmen Chan described a conservation programme that lost its funding because the metrics were based on an unrealistic timeline, making a problem in measurement design look like a failure of the programme itself. Letícia Vargas found that in her interviews in Brazil, both investors and investees said impact measurement was not worth the time or cost, mainly because there were no external incentives. Finally, Juan J. Alarcón added a temporal dimension to this dynamic by stating that real impact is only visible one to two years after a project ends, but often donors ask for results right at the end, so in practice impact is never actually measured.

3.3 Data quality, verification and distance from beneficiaries

Several interviewees flagged the problem with the underlying quality and credibility of data, independent of which frameworks are at use. Audrey Selian identified the absence of a direct beneficiary feedback loop as the most important gap. Tenke Zoltani noted that even when companies are complying with reporting requirements, data still is available late or inconsistently formatted. Monica Blatyta further added that the data she receives from social enterprises cannot be fully audited which is why she puts disclaimers in all impact reports.

On the social enterprise side, Balca Yilmaz explained that calculating the environmental impact of her wind turbine sensing technology depends heavily on customers who are not motivated to provide such information. Daniel Nowack highlighted one strong barrier of data which is about collecting sensitive social indicators, such as those related to survivors of trafficking or sexual exploitation, that can be ethically impossible to collect at scale regardless of technical capabilities.

Audrey Selian mentioned the blockchain-based verification systems which were proposed several years ago as “*garbage in, garbage out*”. So, the data may be fixed but if the input is wrong then the results are still unreliable, and that problem has not been solved.

Theme 4 - Standardisation and Artificial Intelligence

4.1 Fragmentation is recognised but solutions are contested

Although all interviews showed there was acknowledgement of the fragmentation of impact measurement frameworks, no convergence exists on what the best solution should be like. When asked if a universal core metrics set would be feasible, Daniel Nowack was sceptical about it by suggesting that sector-level approaches would be more achievable. Similarly, Letícia Vargas showed less interest in technical solutions, arguing that the main problem is not a lack of frameworks but a lack of real demand for rigorous measurement. She asked a basic but important question: *“The question is: who wants this measurement? What is this measurement for?”*. A development finance specialist made a similar point, arguing that better frameworks will have little effect unless impact data is truly used in decision-making and backed by the right incentives and culture.

Audrey Selian clearly emphasised who should drive solutions, arguing that *“this has got to come from institutional leadership in the space. It's got to come from the likes of major banks or coalition of banks or the stuff that the Global Impact Investing Network hasn't done, which is to create the information architecture to bridge all these asymmetries”*. She still acknowledged that previous attempts were done but failed. Tenke Zoltani, however, was optimistic about how SDGs are a sufficient common denominator and suggested that a small set of meaningful indicators could provide practical value.

Laís Rocha argued that collaborative, sector-level theories of change would be more valuable than each investor using fragmented versions independently. But she still noted that business confidentiality is a major barrier to this type of coordination.

4.2 AI is seen as a practical enabler, not as a distinct solution

Interviewees' responses on AI as a potential solution were more nuanced than simply supporting or rejecting it. Most of them expressed some openness in principle but under certain conditions.

Monica Blatyta was the most motivated, describing a fragmented and manual data environment across financial, ESG, and impact reporting, and noting that better integration and automation would significantly improve her work. Tenke Zoltani seemed also open to any AI tool that could save her time, but she noted data privacy must be adhered to. Laís Rocha responded that she saw AI as a means of making measurement “less artisanal”, particularly for processing large columns of mixed qualitative and quantitative data.

By contrast, several interviewees thought about constraints that AI cannot address. For example, Daniel Nowack drew a distinction between the technical dimension of harmonisation from its political dimension (*“AI could provide the thematic and the scientific basis for it, but it doesn't solve and bring people to the table”*). He further raised concerns about connectivity issues and language access in low- and middle-income contexts, and highlighted a key consent challenge which is that AI system training on beneficiary data may be doing so without the knowledge or approval of those individuals. Likewise, a development finance specialist also distinguished between output-level aggregate data which AI may potentially process, and real impact data involving people's lives which is sensitive, self-reported and

structurally different in kind. She further added that even among investors who already experiment AI tools, human judgment and contextual knowledge are still necessary in the process.

Juan J. Alarcón added that AI tools will only become useful once the field has built up a strong enough base of reliable impact evaluations and track records, meaning that better data quality and larger volumes of data are still a necessary first step.

Nathan Sussman (Professor of Economics at the Geneva Graduate Institute) mentioned the Impaakt platform as a practical example of the use of technology (and probably AI) in the context of Impact Investing. The platform uses a crowdsourcing approach and data analysis to measure the impact of companies on ESG pillars, relying on a network of many analysts and citizens. This methodology can be relevant because it offers a way to integrate the voice of beneficiaries and civil society, a gap identified in previous sections. This allows the validation of impact to go beyond the data self-reported by companies.

AI Statement

AI-assisted tools were used during the preparation of this report in a limited editorial capacity. As detailed in [Annex 2](#), AI was used as a first-pass support tool during the thematic analysis process to help organise interview material and identify relevant passages across transcripts. All coding decisions, interpretation of themes, selection of evidence, and analytical conclusions were made by the research team.

AI was also used to support language quality, including grammar correction, sentence shortening, rephrasing, translation from French, Portuguese and Bengali into English, and checks for clarity and coherence. It was not used to generate the research design, research questions, findings, recommendations, or original ideas. All content and analytical decisions remain the responsibility of the authors.