Maastricht University

Master Thesis

Title:	A Deforestation-Free Future: Bridging Data Management and Social Governance for EUDR Compliance
Confidential:	*No*

By

Paul		Geerlings
Student number	I6406755	

Study	*International Business*
Specialisation for MSc IB students*	*Information Management and Business Intelligence*
Thesis supervisor *(1st specialisation)*	Amar, Sidi
Second reader *(if already known)*	Grigoriev, Alex

Date & Place	*20-6-2025, Maastricht*

School of Business and Economics

Master

Abstract

The European Union Deforestation Regulation (EUDR) mandates that companies prove the traceability and deforestation-free origin of high-risk commodities, such as cocoa, palm oil, and soy. This regulation presents substantial data management challenges, particularly in supply chains involving smallholder farmers who often operate in low-infrastructure settings. This thesis explores the question: What data management practices ensure EUDR compliance efficiency within companies?

To address this, a mixed-methods approach was employed, combining survey responses from 11 sustainability professionals with expert interviews from three industry practitioners. The research investigates four key data dimensions: collection, integration, validation, governance and reporting, using a conceptual framework that incorporates "social governance" as a cross-cutting factor, encompassing farmer participation, cooperative mediation, and data ownership. The findings reveal that data collection remains largely managed by companies internally (centralized) and dependent on third parties, with low levels of direct farmer involvement despite its perceived benefits for accuracy. Data integration is fragmented and technically difficult, with internal coordination emerging as a key barrier. Validation practices blend manual and automated systems, but governance remains predominantly centralized, excluding smallholders from meaningful oversight. And business intelligence (BI) tools are only effective when supported by structured, integrated data.

This thesis contributes to the growing body of literature on sustainability data governance by demonstrating that compliance is not merely a technical challenge, but a socio-technical one shaped by power asymmetries. It recommends fostering inclusive governance models, empowering cooperatives, and investing in interoperable systems to meet EUDR requirements while promoting equity and long-term sustainability.

Table of Contents

1. Introduction

The rapid expansion of global agricultural supply chains has contributed to widespread deforestation, biodiversity loss, and environmental degradation (IPCC, 2019; Curtis et al., 2018). In response, the European Union introduced the European Union Deforestation Regulation (EUDR), which mandates that companies placing high-risk commodities on the EU market, such as cocoa, palm oil, and soy must prove their products are deforestation-free and fully traceable. This landmark regulation requires companies to collect geolocation data, conduct due diligence risk assessments, and submit compliance statements, placing considerable pressure on the quality and structure of corporate data systems (European Union, 2023). For example, Nestlé's cocoa sourcing operations in West Africa face compliance pressure as they must trace beans back to individual farms while dealing with incomplete farmer data (Reuters, 2024)

Ensuring compliance with EUDR presents considerable data management challenges. Companies must now integrate data from diverse sources: farmer-level data, satellite imagery, and certification records, into cohesive and verifiable information streams. These data must be accurate, timely, and traceable across complex, multi-tiered supply chains (European Union, 2023). However, as highlighted by Achilles (n.d.) and TraceX Technologies, firms frequently face fragmented data systems that inhibit integration, while inconsistencies and data gaps undermine the accuracy of compliance reporting. Particularly problematic is the inclusion of smallholder farmers in data ecosystems. Many operate in low-infrastructure contexts, face literacy or technology access barriers, and are often reduced to passive data sources rather than active governance participants (Development Gateway, 2023). Moreover, while the role of farmers as data providers is acknowledged, their lack of control over how data is governed and shared, often framed as limited "data ownership", and the absence of structured cooperation mediation through trusted intermediaries like cooperatives signal a deeper governance gap that affects both inclusivity and data quality (Development gateway, 2023).

The challenge is twofold: on one hand, companies must adopt internal systems capable of integrating and validating diverse data streams; on the other, they must foster inclusive governance frameworks that give smallholder farmers greater visibility and participation in compliance processes (Amar & Beranek, 2024). As noted by SupplyChainBrain (2023), data collection inefficiencies and weak integration are among the most cited obstacles to meeting

EUDR requirements. These limitations are exacerbated by inconsistent validation protocols, insufficient internal coordination, and heavy reliance on third-party technology providers whose tools may lack transparency or inclusiveness (Development Gateway, 2023).

Despite growing attention, academic literature has only begun to explore how companies operationalize EUDR compliance through data management (Gardner et al., 2019). In particular, there is little insight into the effectiveness of farmer-centric governance models, the integration of satellite and field-level data into business intelligence (BI) platforms, and the actual barriers companies face in aligning data workflows with EUDR criteria (Amar & Beranek, 2024; Fairtrade Foundation, 2021). Most research to date emphasizes technical solutions while under examining the institutional dynamics that determine who controls, validates, and ultimately benefits from data, especially in relation to smallholder participation, equity, and ownership (Ruder & Wittman, 2024). Moreover, while tools such as Satelligence or TraceX provide technical support, they often lack mechanisms to include farmers in decision-making or data validation processes (Developmental Gateway, 2023).

The study uses a mixed-methods design, combining survey data from industry professionals with expert interviews. It seeks to fill that gap by exploring how companies manage deforestation-related data in light of EUDR requirements, with a particular focus on the role of smallholder farmers, data integration practices, governance and validation models, and reporting and BI tools. The findings aim to offer actionable insights for companies, regulators, and technology providers seeking to analyse more inclusive, efficient, and transparent compliance frameworks with a focus on smallholder farmers.

To guide this research, the following central research question is posed: *What data management practices ensure EUDR compliance efficiency within companies?*

To answer this question, four sub-dimensions are explored:

1. **Data Collection**: What role do smallholder farmers play in data collection, and how does their involvement influence data quality?
2. **Data Integration**: How effectively do companies merge satellite, certification, and farmer-provided data into usable systems?

3. **Data Validation and Governance**: What governance structures are used to ensure data accuracy and transparency, and how do they influence error rates?
4. **Compliance and Reporting**: How are business intelligence tools used to generate EUDR reports, and what challenges or factors emerge?

While the sub-dimensions above are structured along operational data processes, the conceptual model introduced in Chapter 2 highlights "Social Governance", encompassing farmer participation, cooperative mediation, and data ownership, as a cross-cutting dimension. Rather than existing as a separate stage, social governance influences each phase of the data pipeline and is central to shaping the equity, quality, and inclusivity of EUDR compliance systems (Development Gateway, 2023; Ruder & Wittman, 2025). Based on the insights gathered in this study, the conceptual model is further refined into an annotated framework presented in Figure 9 (chapter 5), which highlights critical bottlenecks, governance tensions, and opportunities for participatory reform across the EUDR data pipeline.

This research holds both theoretical and practical relevance. Academically, it expands the understanding of data governance in sustainability regulations by integrating perspectives from the field, including farmer-centric governance theory and participatory data models. Practically, it provides guidance for companies navigating EUDR compliance and offers recommendations for building more robust data infrastructures that balance compliance efficiency with inclusivity and accountability.

Ultimately, this thesis argues that EUDR compliance is not merely a matter of technical optimization but a question of governance, equity, and participation. By placing smallholders at the center of data ownership and oversight, companies can meet regulatory demands while promoting inclusive, sustainable, and socially just supply chain practices.

The remainder of this thesis is organized as follows. First, the literature review situates this study within existing academic and practitioner debates around traceability, data governance, and smallholder inclusion. This is followed by the methodology chapter, which outlines the mixed-methods approach using a survey and expert interviews. The results chapter presents the key findings, while the discussion interprets their significance in light of the research questions and existing literature. Finally, the conclusion summarizes the main insights, discusses limitations, and offers recommendations for future research and practice.

2. Literature Review

This literature review synthesizes existing academic and practitioner knowledge on data management for compliance with the EUDR, with a particular focus on farmer-centric governance, data collection, data integration, traceability systems and reporting for compliance. The aim is to identify what is already known and where gaps persist, especially regarding smallholder inclusion and the operationalization of compliance mechanisms in deforestation-linked supply chains such as cocoa, palm oil, and soy (Development Gateway, 2023). To structure this review and guide the empirical analysis, a conceptual framework was developed, visible at the end of this chapter, that emphasizes both technical data practices and a cross-cutting dimension of social governance, encompassing farmer participation, cooperative mediation, and data ownership (Figure 1).

The review situates this study within broader academic debates on sustainability governance, digital traceability, and inclusive data ecosystems. While recent literature has examined traceability technologies and supply chain due diligence (Gardner et al., 2019; Lambin et al., 2018), fewer studies have explored the intersection between technical systems and participatory data governance models involving smallholder farmers (Ruder & Wittman, 2025; Development Gateway, 2023). This research addresses that gap by combining institutional and technological perspectives on EUDR compliance.

2.1 EUDR and Supply Chain Traceability

Definition and Regulatory Context

The EUDR aims to eliminate deforestation from EU supply chains by requiring that products placed on the market are proven to be deforestation-free (meaning not produced on land deforested after December 31, 2020) and legally produced in accordance with the laws of the country of origin. Companies must submit detailed due diligence statements to regulatory authorities and provide geolocation coordinates that trace commodities back to their origin (European Union, 2023). Covered commodities include cocoa, soy, palm oil, coffee, rubber, and cattle, as well as derived products such as chocolate and leather.

Traceability Requirements

The EUDR outlines three primary components of compliance (European Union, 2023):

1. **Geolocation**: Companies must provide the geographic coordinates of the plots of land where the commodities were produced.
2. **Risk assessment and mitigation**: Companies must assess the risk that products originated from land deforested after December 31, 2020, and take mitigation measures where risks are identified.
3. **Due diligence reporting**: A due diligence statement must be submitted to the competent national authorities prior to placing products on the EU market.

Operationalization Challenges

Despite its strong regulatory intent, EUDR implementation presents significant operational obstacles. Traceability is particularly challenging in multi-tier supply chains where companies must establish plot databases and manage data quality across smallholder farmers with limited infrastructure (SupplyChainBrain, 2023). Gathering accurate geolocation data is difficult, especially when plot boundaries are unregistered or informally managed (TraceX Technologies, 2023). Moreover, companies struggle with integrating diverse data streams, from satellite imagery to field-level records, into centralized BI systems (Ahoa et al., 2025; Fairtrade Foundation. 2021).

Cost and Technological Limitations

Compliance costs are expected to be highest for firms reliant on complex supply chains or legacy data systems (Gocsik et al., 2024). While technologies such as geospatial analytics offer promise, they often lack farmer-facing components or mechanisms for participatory data governance (Ruder & Wittman, 2025). Additionally, inconsistent validation protocols and low-quality field data remain key bottlenecks (Achilles, n.d.).

2.2 The Central Role of Smallholder Farmers in Deforestation-Linked Supply Chains

Smallholder farmers play a crucial role in the global supply of key agricultural commodities linked to deforestation, including cocoa, palm oil, and soy. It is estimated that over 90% of cocoa production in countries like Ghana and Côte d'Ivoire is carried out by small-scale producers (Fairtrade Foundation, 2021). Similarly, in Indonesia and Malaysia, more than 40% of palm oil is produced by smallholders, who often operate on less than two hectares of land (RSPO, 2022). These producers are often located in biodiversity-rich tropical regions, where agricultural expansion directly intersects with deforestation frontiers (FAO, 2020; Curtis et al., 2018). Consequently, smallholders are not only central to commodity supply but also to the success or failure of deforestation mitigation efforts under the EUDR.

Despite their importance, smallholders frequently face systemic exclusion from compliance frameworks. As Development Gateway (2023) notes: farmers are often treated as passive data providers rather than active participants in sustainability governance. This marginalization creates a disconnect between regulatory objectives and implementation on the ground. For example, in its Towards Forest Positive Cocoa report, Nestlé reveals it has mapped over 125,000 smallholder farms in West Africa, but this achievement represents a substantial logistical and financial effort, requiring significant field data collection and coordination with local cooperatives. Similarly, Fair Labor Association's assessment in Cameroon shows Nestlé traced only about 2,368 farmers across 12 cooperatives, highlighting the limited scale of current tracing efforts and the disproportionate cost per farm (FLA, 2023). Even Nestlé's 2017 sustainability statement admits only 42.9% of its cocoa was traceable, illustrating the steep resource investment needed to expand traceability further (Nestlé, 2023).

A key strategy to address these barriers involves intermediary support structures such as farmer cooperatives. The following section (2.3) further elaborates the operational and technical roles of cooperatives in facilitating data collection.
Empowering smallholders through inclusive governance, capacity building, and digital access is not merely a social imperative, it is a practical requirement for operationalizing the EUDR (Development Gateway, 2023). Their unique position at the origin of supply chains makes them key agents in achieving deforestation-free outcomes, provided the data systems are designed to be participatory, accessible, and responsive to local realities.

2.3 Data Collection in Smallholder Contexts

Smallholder farmers are critical actors in global supply chains, yet their involvement in data collection for regulatory compliance remains limited and problematic. Multiple studies have highlighted the technical, social, and economic barriers that hinder effective farmer data participation in traceability systems (Abubakari & Sarpong, 2022; Development Gateway, 2023).

Barriers to Farmer-Generated Data

Smallholder participation in data collection remains limited due to a combination of technical, educational, and structural barriers. Many farmers lack access to smartphones, internet connectivity, or geospatial tools, preventing them from contributing directly to digital compliance processes reliant on geolocation and recordkeeping (Development Gateway, 2023). Low literacy and limited digital skills further reduce engagement, with adoption of new technologies often hindered by usability issues and insufficient training (Abubakari & Sarpong, 2022; Ruder & Wittman, 2025). Moreover, the data that is collected tends to be highly fragmented. Informal sourcing practices, inconsistent data formats, and reliance on paper-based or rudimentary tools undermine standardization. For instance, while one cooperative might use GPS-enabled apps, another may submit hand-drawn maps or verbal plot descriptions (TraceX Technologies, 2023). In Côte d'Ivoire, more than 55% of cocoa remains untraced to the farm level, largely due to such inconsistencies and digital system weaknesses (Renier et al., 2023). This lack of standardization decreases trust, complicates traceability, and increases the validation burden for downstream actors. Overcoming these challenges will require interoperable data standards, stronger cooperative support, and sustained investment in infrastructure and training for frontline actors.

The Role of Intermediaries

To bridge the data collection and communication gap between smallholder farmers and downstream supply chain actors, many companies rely on farmer cooperatives, collectively organized groups of producers that aggregate, process, and market agricultural outputs on behalf of their members. These cooperatives act as intermediaries that coordinate training, facilitate data collection, and often manage certification and compliance processes. Their organizational structure allows for more standardized reporting than individual farmers and offers a critical access point for companies and certifiers engaging with fragmented

smallholder networks (Development Gateway, 2023; Rainforest alliance, 2022). Cooperatives are also positioned as potential mediators of data governance, capable of translating EUDR requirements into accessible processes for smallholders while balancing compliance demands from buyers (Rainforest Alliance, 2023; Development Gateway, 2023). However, this mediation role is not always formalized or supported, especially where cooperatives face resource constraints.

For example, the Rainforest Alliance often works through cooperatives to implement sustainability standards and traceability requirements. In West Africa's cocoa sector, cooperatives affiliated with Rainforest Alliance certification help farmers meet environmental and social criteria by providing training, digital recordkeeping tools, and guidance on deforestation-free practices. These cooperatives collect farm-level data (including geolocation, yields, and pesticide use) and share it with certifiers and buyers as part of the compliance documentation process (Rainforest Alliance, 2022). This model allows companies to scale compliance efforts more efficiently, but it also concentrates responsibility in the hands of intermediary organizations, which may lack sufficient digital capacity or face misaligned incentives.

While cooperatives can improve efficiency and coordination, their effectiveness depends heavily on local governance quality, trust, and access to technology. In some cases, cooperatives themselves are under-resourced or lack internal accountability mechanisms, which can compromise the quality of data provided. Therefore, while they remain vital to scaling EUDR compliance in smallholder contexts, relying exclusively on cooperatives without supporting capacity-building and transparency measures may still lead to inconsistent data and compliance gaps (Ruder & Wittman, 2025).

Identified Gap

Although cooperatives and other intermediaries help mitigate some of the digital and literacy barriers faced by smallholder farmers, they cannot substitute for meaningful farmer participation in data governance. Most current compliance systems continue to treat smallholders primarily as data subjects, entities from whom data is extracted, rather than as active contributors or co-owners of data (Development Gateway, 2023; Ruder & Wittman, 2025). This exposes a critical gap in both academic research and practice: despite increasing attention to farmer-centric governance (Amar & Beranek, 2024), systematic research on its operationalization within traceability systems remains limited (Ruder & Wittman, 2025;

Development Gateway, 2023). Further exploration is needed to understand how data systems can move beyond top-down extraction toward participatory and inclusive design.

2.4 Data Integration

Effective compliance with the EUDR relies not only on data collection but also on the seamless integration of diverse data sources into coherent reporting systems. These sources include farmer-level field data, certification documents, and satellite imagery, each existing in different formats and governed by distinct validation protocols. The challenge lies in transforming this heterogeneous information into standardized, interoperable datasets that meet EUDR's stringent geolocation and traceability requirements (Ahoa et al., 2025; Renier et al., 2023). While governance-oriented frameworks stress inclusivity and farmer participation, the technical infrastructure to support data integration remains fragmented and underdeveloped.

System Compatibility and Technical Fragmentation

One of the most commonly cited obstacles in achieving EUDR compliance is the incompatibility between data systems. Many companies operate with legacy enterprise resource planning (ERP) platforms or internal BI systems that were not originally designed to process geospatial data or integrate field-level inputs (Satelligence, 2023). As a result, these systems often lack the capability to interpret geographic coordinates or overlay them with satellite-derived imagery, making real-time land-use verification and monitoring difficult. This technical gap often necessitates custom-built interfaces or reliance on third-party service providers, increasing both operational costs and complexity.

In addition, while certification data (from Rainforest Alliance or Fair Trade) and satellite imagery are typically captured in structured and standardized formats, farmer-generated data remains highly variable.

Without shared formatting rules (such as harmonized data types, coordinate systems, and timestamp conventions) merging inputs from satellites, certifiers, and farmers can result in duplication, inconsistency, or exclusion. These interoperability issues are a central barrier to system integration in agri-food supply chains (Ahoa et al., 2025; TraceX Technologies, 2023). Without robust metadata harmonization protocols and data mapping standards, integrating these sources can lead to duplication, omissions, or inconsistencies that undermine traceability efforts.

Lack of Empirical Insights on Integration Rates

Despite the technical importance of data integration, empirical research quantifying integration success in supply chains remains limited. Most studies emphasize the potential of digital tools rather than evaluating their practical implementation. Industry evidence from Satelligence and TraceX suggests that only a minority of companies have achieved full interoperability across geospatial, certification, and farm-level datasets (Satelligence, 2023; TraceX Technologies, 2023). Firms operating across diverse sourcing regions often encounter persistent integration barriers due to heterogeneous digital capacities, further complicating compliance with deforestation-related regulations (Gardner et al., 2019).

Implications for EUDR Compliance

The lack of seamless data integration directly undermines companies' ability to perform accurate risk assessments and submit due diligence statements that reflect the full scope of production practices. This limitation is particularly acute for firms managing complex supply chains with hundreds or thousands of actors. Without automated linkages between satellite imagery, farmer-level data, and certification frameworks, companies must rely on manual validation processes, often time-consuming, costly, and error-prone (Satelligence, 2023; Ahoa et al., 2025).

In summary, integration challenges remain a core barrier to effective EUDR compliance. Addressing these issues will require the development of interoperable data standards, upgrades to enterprise system architectures, and a shift toward platforms capable of ingesting both structured and semi-structured data streams, from field to cloud (Gardner et al., 2019).

2.5 Data Validation and Governance Frameworks

Data validation and governance are critical components of effective compliance with sustainability regulations like the EUDR. These processes ensure that the data collected (from farmers, satellite providers, or certifiers) is accurate, verifiable, and trustworthy. However, challenges persist around scalability, inclusiveness, and effectiveness, especially when smallholder actors are excluded from data ownership or when validation relies on non-transparent third-party technologies (Ruder & Wittman, 2025; Development Gateway, 2023).

Manual vs. Automated Validation

In practice, companies use a mix of manual checks, third-party audits, and automated tools to validate deforestation-related data. Manual methods often include field agent verification or cross-checks against supplier declarations. While highly context-sensitive, manual validation is labor-intensive and error-prone, especially across geographically dispersed supply chains (Gardner et al., 2019). Automated remote sensing tools offer scalable, standardized detection of land-use changes. However, without ground verification, these systems can misinterpret seasonal changes or low-resolution imagery, leading to both false positives and negatives. As noted by Dimas Perceka, Remote Sensing & Climate Lead at Koltiva, "Users, regardless of GIS expertise, can verify land cover changes using high-resolution satellite imagery directly from their desktop. They visually compare 'before and after' images to confirm forest clearing, manually validate automated alerts, and leverage detailed imagery for precise assessments" (Koltiva, 2025). This hybrid approach, blending manual oversight with automated detection, is essential for reliable EUDR compliance. However, it also underscores the resource intensity and technical complexity involved in operationalizing robust validation frameworks across diverse sourcing landscapes.

Governance Models: Centralized vs. Farmer-Centric

Data governance refers to the rules, processes, and ownership structures that determine how data is collected, used, and shared across supply chain actors. Most companies today operate under centralized data governance models, where data flows through organizational hierarchies and is managed by compliance or IT teams, often without input from producers (Ruder & Wittman, 2025). While efficient for large-scale reporting, this structure often excludes farmers from participating in data decisions and limits their visibility into how their information is used.

In contrast to top-down compliance models, farmer-centric data governance emphasizes participatory practices, transparency, and shared ownership of agricultural data. This approach seeks to involve smallholder farmers not just as data providers but as co-governors, with meaningful input into how data is accessed, verified, and used. Such models are typically facilitated through cooperatives or digital platforms that allow for joint data management and localized validation. By redistributing power in data systems, farmer-centric governance can foster trust, improve data quality, and align more closely with sustainability and equity goals. Not just participation but also data ownership by smallholders (control over how data is used

and shared) is fundamental to building trust and accountability in sustainability systems (Ruder & Wittman, 2025; Development Gateway, 2023).

Trust, Transparency, and Accuracy

Trust is a central factor influencing the success of data validation systems. Where farmers and intermediaries perceive data collection as extractive, they are less likely to report accurate or complete information (Lambin et al., 2018). Transparency emphasizes that where data providers understand how their information is used, correlates positively with data quality and a willingness to cooperate in verification (Developmental Gateway, 2023). In other words, transparent governance builds trust; trust builds compliance.

2.6 EUDR Compliance and Reporting

As the EUDR raises the bar for traceability, verification, and reporting standards, many companies are turning to BI tools to meet compliance demands (Passionfruit, 2025). These platforms serve as digital infrastructures that facilitate the integration, analysis, and visualization of deforestation-related data, including satellite imagery, geolocation coordinates, certification documents, and field-level reports. Their core function is to convert diverse data sources into actionable compliance dashboards that streamline regulatory reporting and risk assessment (Orbify, 2024).

BI tools typically offer features such as remote sensing integration, automated deforestation alerts, and customizable traceability documentation tailored to meet EUDR reporting formats. By digitizing and centralizing data management processes, these tools can reduce reliance on manual validation and enhance the timeliness and consistency of compliance reporting. However, their effectiveness is highly dependent on the underlying data architecture of the implementing company. Firms with unstructured or siloed internal data often face challenges in fully leveraging BI functionalities, limiting scalability and increasing reliance on third-party data processors (Satelligence, 2024; TraceX Technologies, 2024).

Despite growing interest in these tools, empirical evaluations of their cost-effectiveness and real-world implementation success are limited. Most available evidence is based on promotional case studies or pilot projects, with little peer-reviewed research examining their long-term impact on compliance quality or operational efficiency. Furthermore, many existing platforms focus primarily on technical traceability and lack mechanisms for inclusive

governance, such as incorporating smallholder feedback or ensuring data ownership rights for farmers. This raises concerns about whether current BI solutions can support both regulatory performance and equitable participation across the supply chain (LiveEO, n.d.; BanQu, n.d.).

2.7 Theoretical Framework and Conceptual Model

This thesis adopts a conceptual framework grounded in farmer-centric data governance and sustainability data management to explore how companies comply with the EUDR. It provides structure to the study's four main themes: data collection, data integration, data validation and governance, and compliance reporting. Rather than aiming to test formal hypotheses, the framework serves as an analytical lens to organize and interpret the descriptive findings gathered through the survey and expert interviews. As such, this conceptual model or framework replaces a traditional hypothesis-testing approach by guiding thematic exploration rather than testing directional relationships.

At the core of the framework is the concept of farmer-centric data governance, a participatory approach that promotes inclusion, transparency, and shared ownership of data by smallholder farmers. This model challenges traditional centralized systems where data flows are often top-down and flat. By placing farmers at the center of compliance data ecosystems, it becomes possible to build trust, enhance data accuracy, and align compliance mechanisms with broader sustainability goals (Development Gateway, 2023).

The framework distinguishes several interlinked thematic components that guide the analysis of EUDR compliance practices in smallholder-dominated supply chains:

1. **Governance Models**: This dimension contrasts centralized, supplier-led, and third-party governance structures with farmer-led models that prioritize participation, data ownership, and local oversight. The analysis explores how these governance configurations shape control over data flows and influence the inclusivity of compliance strategies.
2. **Social Governance:** Rather than acting as a discrete stage, social governance represents a transversal influence on all other components. It encompasses farmer participation, cooperative mediation, and data ownership.

3. **Core Data Management**: This includes the operational processes of data collection (from farmers or satellite sources), integration (via ERP, geospatial, or BI systems), and validation (through manual, automated, or hybrid methods). These technical practices form the backbone of traceability and due diligence systems.
4. **Compliance and Sustainability Outcomes**: The framework qualitatively examines perceived improvements in EUDR compliance readiness, traceability efficiency, and farmer inclusion as outcome indicators. It emphasizes that regulatory performance and equitable sustainability must be pursued together.

Figure 1 illustrates the current structure of EUDR-related data flows and governance dynamics in supply chains according to the literature review. The model does not represent causal relationships but instead maps the functional flow of data from inputs (such as smallholder farmers and cooperatives) through governance models and core data management processes to final compliance reporting. Notably, the figure highlights how social governance, defined here as farmer participation, cooperative mediation, and data ownership, is presently limited in scope, influencing primarily the input and governance model stages. This reflects findings from the literature review showing that smallholders are largely excluded from downstream data processes such as validation and reporting. The figure thus captures both the technical sequencing of data activities and the institutional asymmetries that currently shape participation and control in EUDR compliance systems.

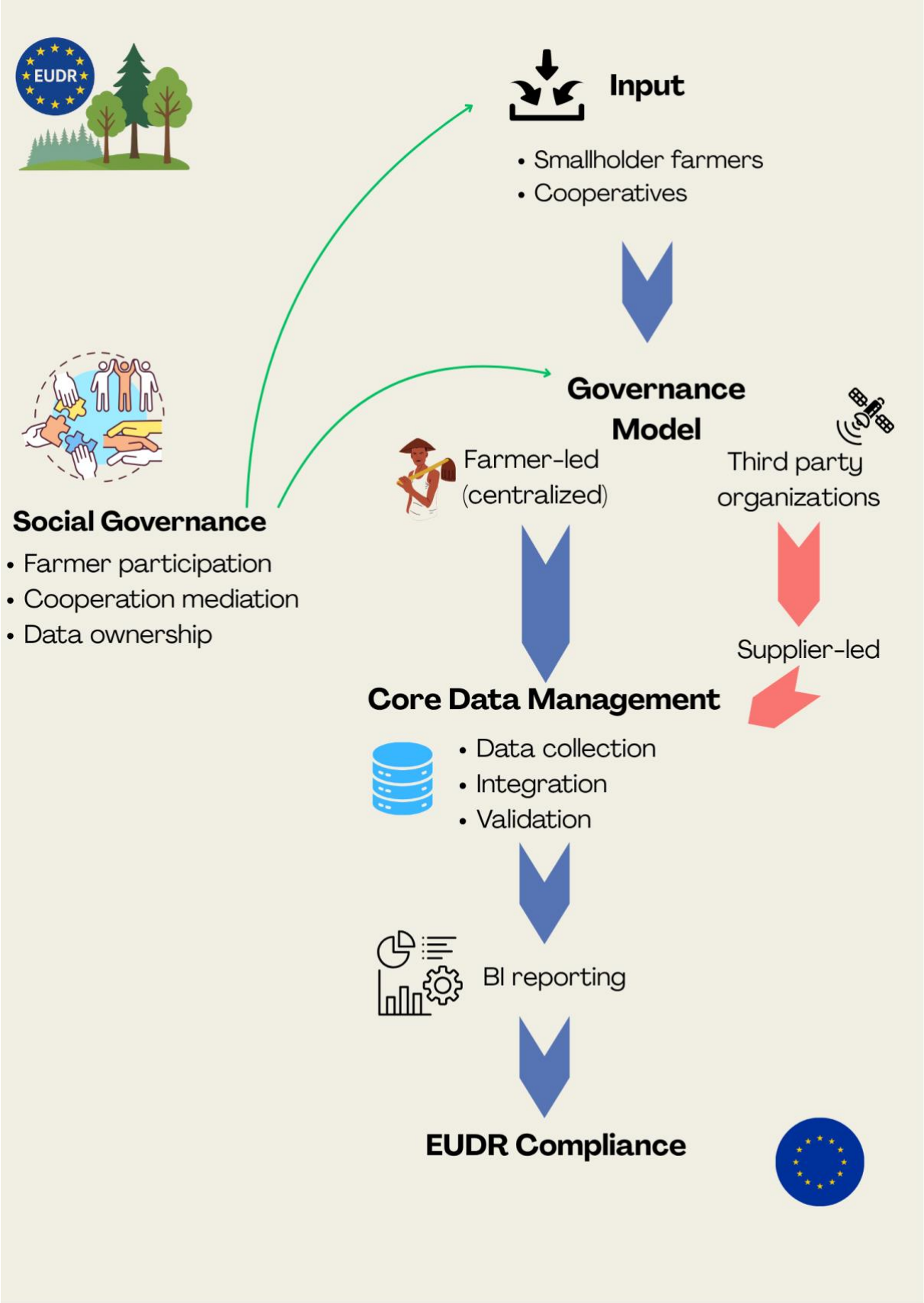

Figure 1: **Conceptual Framework for EUDR Compliance Through Farmer-Centric Data Governance.**

3. Methodology

This chapter outlines the methodological approach used to investigate how companies manage deforestation-related data in response to the EUDR. Given the exploratory nature of the research and the emerging nature of EUDR implementation, a mixed-methods design was chosen to capture both quantitative trends and qualitative insights. The study combines survey data from professionals involved in sustainability and compliance with expert interviews to provide a comprehensive view of current data practices, governance structures, and the role of smallholder farmers. The following sections detail the research design, sampling strategy, data collection procedures, measurement techniques, and analysis methods.

3.1 Research Design

This study follows a non-experimental, descriptive, and cross-sectional design, using a mixed-methods exploratory approach to investigate how companies manage deforestation-related data in response to the EUDR. Given the growing implementation of the EUDR and the lack of large-scale empirical data, a mixed-methods approach allows for both quantification of emerging practices and deeper contextual insights especially regarding smallholder governance, through interviews (Creswell & Creswell, 2018). Rather than testing hypotheses or establishing causal relationships, the goal is to explore current practices, challenges, and perspectives surrounding data collection, integration, validation, and compliance reporting, particularly regarding the inclusion of smallholder farmers.

This mixed-methods approach enables triangulation of findings and offers more depth to the survey answers (Saunders, 2019).

3.2 Sample and Setting

Survey Participants

The survey targeted professionals involved in EUDR implementation, sustainability, traceability, and supply chain management. A purposive sampling strategy was employed, primarily through LinkedIn and company websites, to reach practitioners working in sectors exposed to deforestation risks such as cocoa, palm oil, and soy (Saunders et al., 2019).

A total of 11 valid survey responses were collected. Respondents held roles such as sustainability officers, compliance analysts, and data managers. All participants worked for

European-based importers or traceability solution providers all exposed to deforestation related supply chains. To preserve anonymity, detailed demographic data was not collected. These roles were selected because they are typically responsible for operationalizing regulatory frameworks like the EUDR and possess detailed knowledge of traceability and compliance systems.

Interview Participants

Three semi-structured expert interviews were conducted with professionals from organizations actively working on EUDR-relevant compliance. Interviewees were selected either from the survey pool or via direct outreach based on their industry relevance and willingness to provide deeper insights. The interviews offered contextual understanding of the challenges and strategies that cannot be fully captured through quantitative methods alone.

Sampling Difficulties

Despite targeted outreach, sampling proved challenging due to the novelty of the EUDR and confidentiality concerns within companies. Approximately 50 professionals were contacted via LinkedIn and company websites, but many declined to participate or were restricted by internal non-disclosure policies. As a result, there is a risk of selection bias, with the sample potentially skewed toward more open or compliance-ready organizations. These limitations are acknowledged and should be considered when interpreting the findings.

3.3 Procedure

Survey

A structured online survey was developed and distributed via Qualtrics. The survey was developed specifically for this study based on themes identified in the literature review (Fink, 2017). The questionnaire consisted primarily of closed-ended multiple-choice and 5-point Likert scale questions, covering the four core themes:

- Data Collection
- Data Integration
- Data Validation and Governance
- Compliance and Reporting

Respondents were asked about their data sources, the extent of farmer involvement, integration difficulties, governance structures, and validation practices. The survey took approximately 5–10 minutes to complete and was available for four weeks. All participants were informed about the academic purpose of the research, and their responses were recorded anonymously. An overview of the survey framework is present in table 2 in chapter 4. The full survey instrument is provided in Appendix A.

Interviews

The interviews followed a semi-structured format based on a predefined interview guide (see Appendix B), which included open-ended questions aligned with the study's four analytical themes: data collection, integration, validation, and reporting. Interviews were conducted virtually using video conferencing platforms such as Zoom or Teams or via telephone. Each interview lasted between 15 and 25 minutes and followed a semi-structured guide to allow open-ended discussion on data practices and challenges. Questions focused on:

- Practical experiences with data collection
- Validation methods and error detection
- Integration of satellite/farmer/certification data
- The role of farmers and cooperatives in governance

Interviews were documented through field notes and partial transcripts, and verbal consent was obtained from all participants. Anonymity and confidentiality were guaranteed.

3.4 Measures

This study is descriptive and does not involve inferential statistics or hypothesis testing, given the limited sample size. Instead, the research uses an exploratory framework based on operational themes derived from the conceptual model.

Each variable was measured as follows in table 1. There is an extended version of the survey framework in table 2 in the results chapter.

Table 1: Short overview of the variables.

Variable	Operationalization	Measurement Type
Data collection	Sources used and focus on smallholder farmers	Multiple choice, Likert scale
Data integration	Level of success, perceived difficulty	Percentage estimate, Likert scale
Data validation & Governance	Governance structure, error rates, transparency	Multiple chioce, open-ended
Compliance Efficiency	Perceived reporting ease and audit preparedness	Likert scale

Linking Survey Themes to Literature

The survey was designed based on established themes from recent literature on supply chain traceability, sustainability compliance, and farmer-centric governance. For example, the inclusion of data source types (Q1, Q2) reflects operational distinctions found in Gardner et al. (2019) and TraceX Technologies (2023), which describe common data collection actors in EUDR-aligned systems. Questions on farmer involvement and data contribution (Q3, Q4) who emphasize the limited but potentially high-impact role of smallholders in compliance ecosystems (Development Gateway, 2023), The perceived link between farmer involvement and data quality (Q5) aligns with trust and transparency (Lambin et al., 2018).

The second section on data integration (Q6–Q8) draws on literature highlighting technical challenges in merging disparate datasets, such as geospatial, certification, and field-level records, into centralized platforms. These questions explore real-world integration performance and perceived linkages with EUDR readiness, building on calls for empirical insight into digital traceability system effectiveness (TraceX Technologies, 2024).

The third section (Q9–Q14) concerns data governance and validation practices. Governance models (Q10–Q11) were framed based on centralized, supplier-led, and farmer-led structures in sustainability governance (Pérez-Aleman, 2008). Validation frequency and transparency

items (Q12–Q14) were informed by best practices in sustainability data verification and risk reduction, as described by Lambin et al. (2018) and Skalkos (2023), and also reflect recent guidance from on real-world error detection challenges.
The final section on compliance and reporting (Q15–Q17) investigates the use and limitations of BI tools for EUDR reporting. While BI platforms are increasingly adopted, their effectiveness is often constrained by poor data interoperability, lack of farmer inclusion, and high implementation complexity. This section captures whether BI tools improve audit readiness, reduce manual workload, or face resistance due to internal data challenges.

3.5 Data Analysis

Survey Data

Survey responses were analysed using descriptive statistics, such as frequencies and percentages, to uncover patterns in current EUDR compliance practices. Due to the exploratory nature of the research and a small sample size (n=11), inferential statistics would lack sufficient power and were not appropriate. Descriptive analysis was chosen to highlight trends rather than test relationships. Visualizations were grouped by theme and are included in the results chapter.

Interview Data

Interview responses were analysed using manual thematic coding. Given the limited number of interviews, manual thematic analysis was sufficient for identifying core patterns and did not require the use of automated software or inter-coder testing (Merriam & Tisdell, 2016). Field notes and transcribed experts were reviewed to identify recurring themes and align them with the four analytical dimensions of the study. Codes focused on identifying:

- Perceived integration barriers
- Views on data ownership and governance
- Challenges in involving smallholders
- Organizational adaptations for EUDR compliance

The qualitative findings were used to complement and explain trends found in the survey data, not to make generalizable claims.

3.6 Validity and Reliability

To ensure the validity of this study, both the survey and interview instruments were grounded in existing academic literature on EUDR compliance and data governance. Survey questions were designed to reflect key dimensions identified in the literature, namely data collection, integration, validation, and reporting, and underwent iterative refinement. Although no formal pilot was conducted, feedback from peers and the academic supervisor helped clarify wording and structure.

Triangulation was used to enhance internal validity. Quantitative survey data was complemented by qualitative insights from expert interviews, allowing for comparison and confirmation across different data sources. This methodological triangulation helps improve the depth of the findings and reduces the risk of mono-method bias (Creswell & Creswell, 2018).

Reliability was ensured through standardized data collection procedures. All survey participants received the same online instrument via Qualtrics, and all interviews followed a consistent semi-structured format based on a predefined question guide. Although the interviews allowed for open dialogue, all participants were asked about the same core topics. Thematic coding of interview responses was done manually. While inter-coder reliability was not assessed due to the small number of interviews, care was taken to document recurring patterns transparently and link themes back to the research dimensions. While the small sample size limits generalizability, the exploratory nature of the study prioritizes depth and insight over representativeness.

Reflection on Methodological Limitations

The survey design, though informed by relevant literature, relied heavily on closed-ended questions and a small sample (n=11), which limited the scope for nuanced insights or statistical generalization. Similarly, the interviews (though rich in practical detail) were short in duration and limited in number (n=3), restricting the depth of contextual understanding that could be achieved. As such, while the findings offer valuable exploratory insights, they should be interpreted as indicative rather than conclusive. Future research could benefit from a broader survey sample and longer, more in-depth interviews to validate and expand upon these observations.

3.7 Ethical Considerations

All participants were informed about the academic nature of the study, its confidentiality safeguards, and their right to withdraw at any time. No personal or sensitive information was collected. Interviewees provided informed verbal consent, and all identifying information was anonymized in the reporting of findings. The study complies with ethical standards set by the university's research guidelines.

4. Results

This chapter presents the results of the study, based on descriptive analysis of survey responses (n=11) and qualitative input from three expert interviews. The findings are structured according to the four main data management themes identified in the research design. While the quantitative data highlights trends across respondents, the interview experts offer deeper context and nuance to interpret those trends.

Table 2: Survey Framework

Variable	Question	Survey question
Data collection		
Data source usage	Q1	Which sources does your company use to collect deforestation-related data?
Data source importance	Q2	Which sources contribute most significantly to your data collection?
Farmer involvement	Q3	To what extent are farmers involved in data collection?
Data Type (Farmer Input)	Q4	What type of data do they typically provide?
Farmer Impact on Accuracy	Q5	To what extent does farmer involvement improve data accuracy?
Data integration		
Data Integration Level	Q6	What % of deforestation-related data integrated into internal systems?
Integration–Compliance Link	Q7	Relationship between integration and EUDR compliance?
Integration Difficulty	Q8	Rate difficulty integrating data from different sources
Data quality and validation		
Governance Existence	Q9	Does your company have a formal governance system?
Governance Type	Q10	Which governance model do you use?
Error Rate Perception	Q11	How do error rates vary by governance model?
Transparency Improvement	Q12	Change in transparency since governance implementation?
Validation Frequency	Q13	How often is data validated?

Validation Tools Used	Q14	Which validation tools are used?
Compliance and reporting		
BI Tool Performance	Q15	Performance of BI tools for EUDR reporting?
BI Tool Limitations	Q16	Challenges experienced with current BI tools
BI Tool Value	Q17	Why are BI tools effective/ineffective (if applicable)?

4.1 Data Collection

This section outlines how companies collect deforestation-related data and explores the level of farmer involvement, types of data collected from smallholders, and how these elements relate to data quality under EUDR compliance.

4.1.1 Sources of Deforestation-Related Data

According to the survey, 7 respondents indicated they use satellite imagery providers, and 7 also reported relying on farmer cooperatives or producer organizations. Certification bodies were cited by 6 respondents, while internal field staff and smallholder farmers were mentioned by 4 and 3 respondents, respectively.

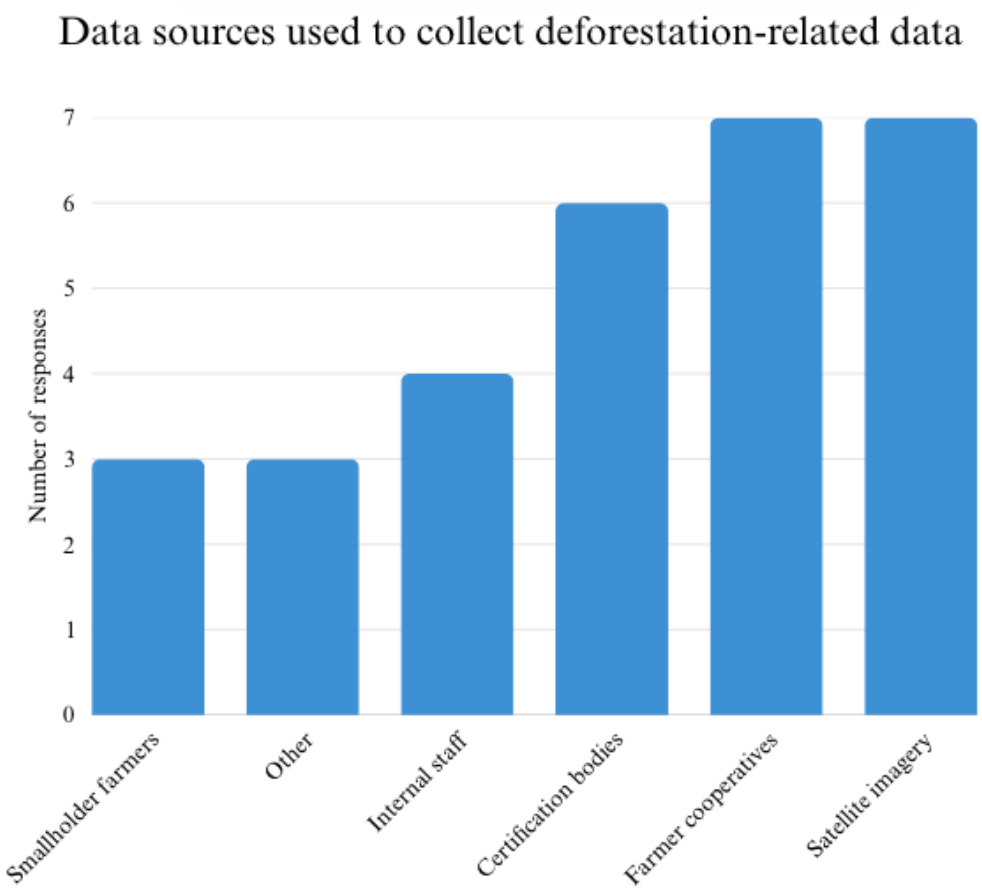

Figure 2: Data sources used to collect deforestation-related data.

These findings suggest that deforestation data collection remains largely dependent on centralized, third-party systems rather than direct farmer engagement. Three respondents who selected "Other" mentioned internal GIS systems, supplier records, trading company data, and middlemen who aggregate farm-level information.

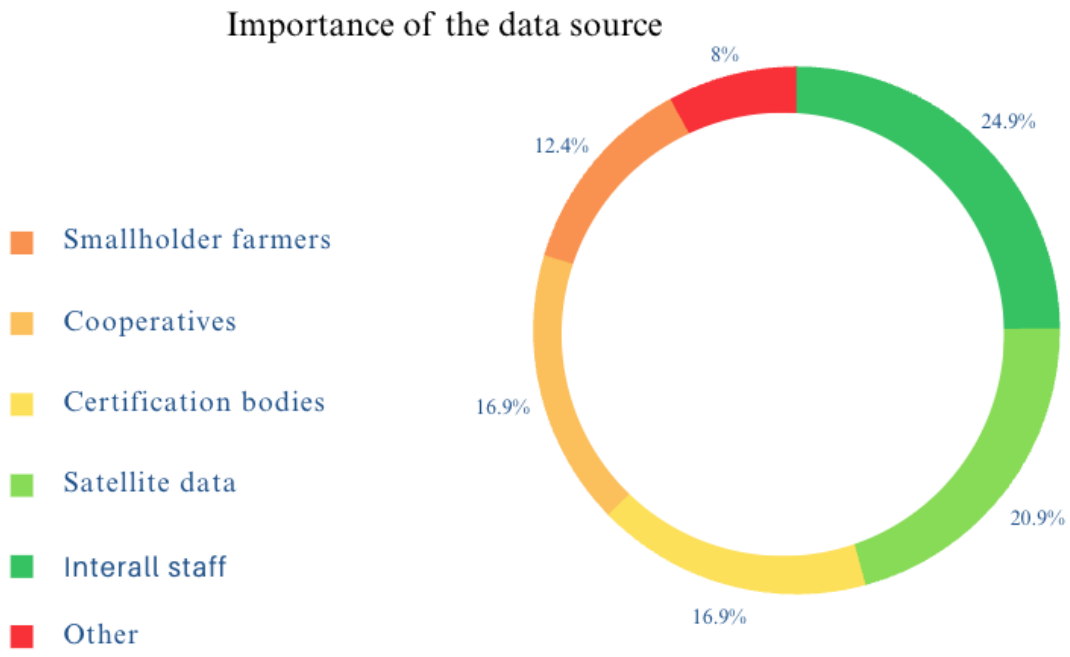

Figure 3: Importance of each data source in the compliance process, from high (green) to low (red) importance.

4.1.2 Farmer Involvement in Data Collection

When asked about the degree of farmer involvement, 5 respondents described it as minimal, and 2 reported no involvement at all. Only 3 respondents indicated moderate involvement, and just 1 reported active participation from farmers.

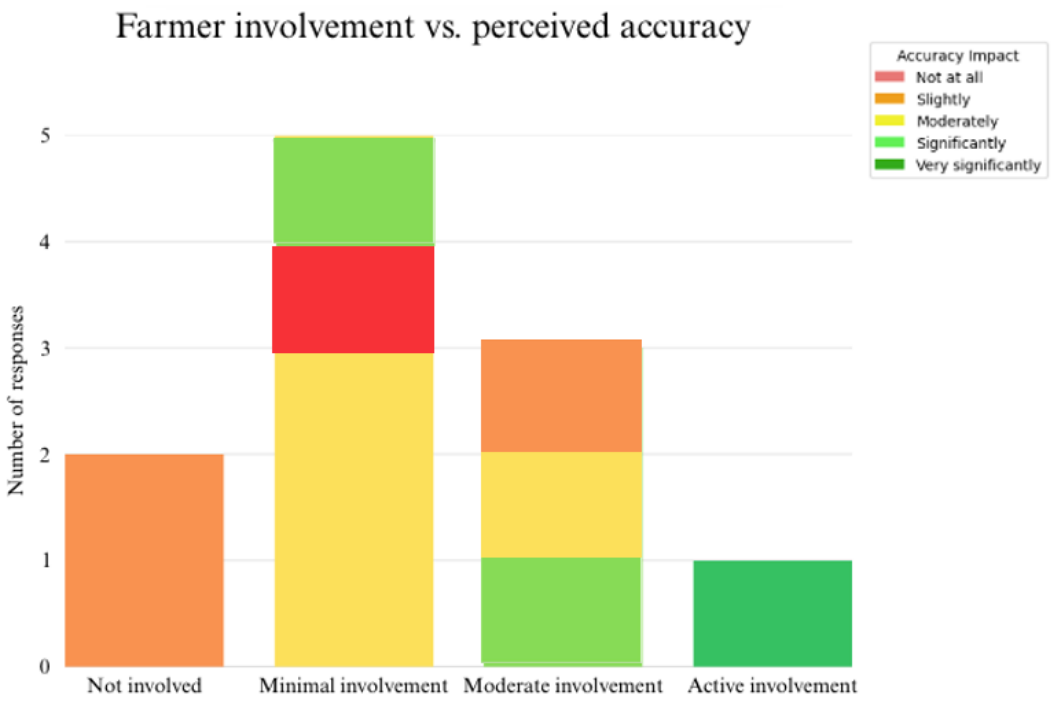

Figure 4: Farmer involvement vs. perceived impact on data accuracy

Nonetheless, 4 respondents reported that farmer engagement had a moderate or significant positive effect on data accuracy. This suggests that although farmer participation is relatively limited, it is perceived as beneficial when it does occur.

4.1.3 Types of Data Provided by Farmers

Respondents indicated that smallholder farmers most often provide GPS coordinates or farm location data (8 responses), followed by crop type (7 responses), harvest information (6 responses), and certification documents (6 responses). Land-use history and other data types were mentioned less frequently.

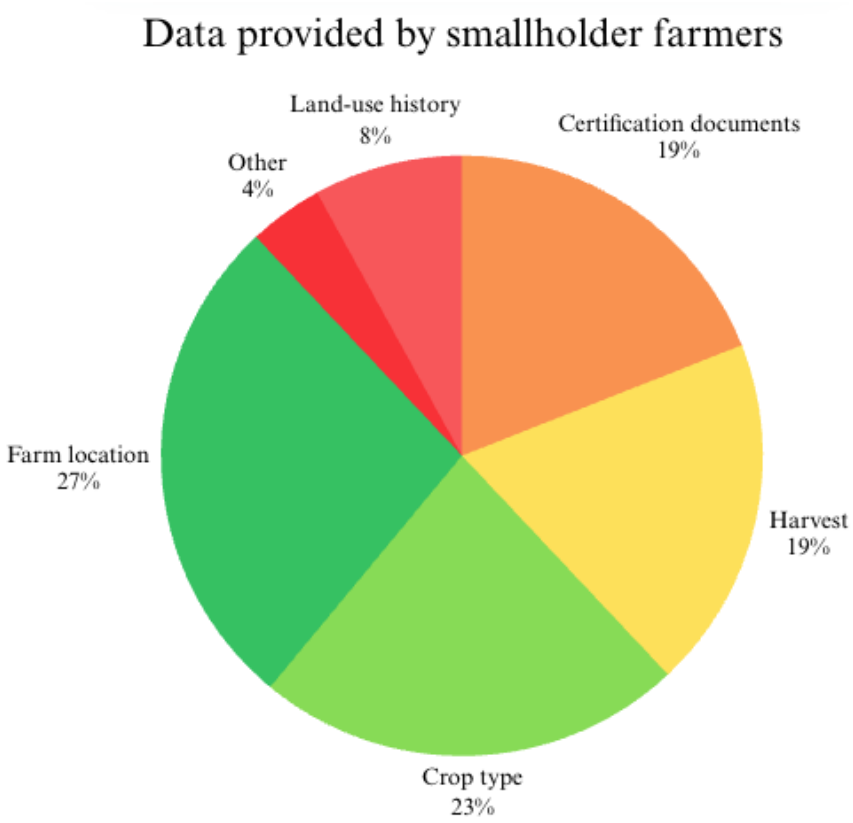

Figure 5: Types of data provided by smallholder farmers

These findings show that farmer contributions tend to be technical and administrative in nature. This reflects a functional rather than strategic role in data ecosystems, where smallholders provide inputs but are not directly involved in governance or integration processes.

4.1.4 Qualitative Insights (Interviews)

To complement the survey findings, one interviewee from a cocoa-focused organization explained that they directly engage smallholders via origin teams, offering data education and feedback mechanisms. This supports the survey finding that moderate farmer involvement often improves data quality. Another interviewee highlighted that in sectors dominated by plantations (palm oil), smallholder participation is minimal and often not prioritized. The third interviewee, from a downstream processor, confirmed that direct farmer engagement is rare in their operations, relying instead on batch-level traceability via suppliers. Interviewees highlighted the critical role of suppliers in the data collection process, particularly in palm oil and cocoa supply chains. Suppliers were described as central intermediaries for delivering geolocation and traceability data. However, multiple experts expressed concern about the

quality and reliability of supplier-provided information. Common issues included incomplete data, incorrect GPS coordinates, outdated records, and inconsistent formatting. One expert noted that “we often spend more time correcting supplier data than analysing it,” underscoring the operational burden posed by weak upstream validation.

To counteract these issues, a company of one of the interviewees has begun bypassing suppliers entirely, deploying in-house origin teams to collect field data directly from cooperatives or farmers on location. This approach was described as more reliable but also resource-intensive and less scalable.

Interviewees also emphasized the dual role of cooperatives as both enablers and potential bottlenecks. In the cocoa sector, cooperatives help organize smallholder data collection and facilitate certification. However, several interviewees noted that many cooperatives struggle with digital literacy, capacity, or standardization, which hampers their ability to serve as effective data stewards.

4.2 Data Integration

This section examines the extent to which companies integrate multiple sources of deforestation-related data, such as satellite imagery, farmer records, and certification documents, into internal reporting systems, and how this integration is linked to EUDR compliance.

4.2.1 Integration Success

When asked to estimate the percentage of deforestation-related data currently integrated into their systems, respondents reported a wide range, from as low as 10% to nearly 100%. The average level of integration across the sample was 43%. This variation highlights major differences in internal capacity and infrastructure readiness between companies.

Respondents were also asked to rate the level of difficulty they experienced when integrating data from different sources. As shown in Figure 6, 7 out of 11 respondents characterized the process as “difficult,” and 2 described it as “very difficult.” Only 2 respondents reported it was “moderate.”

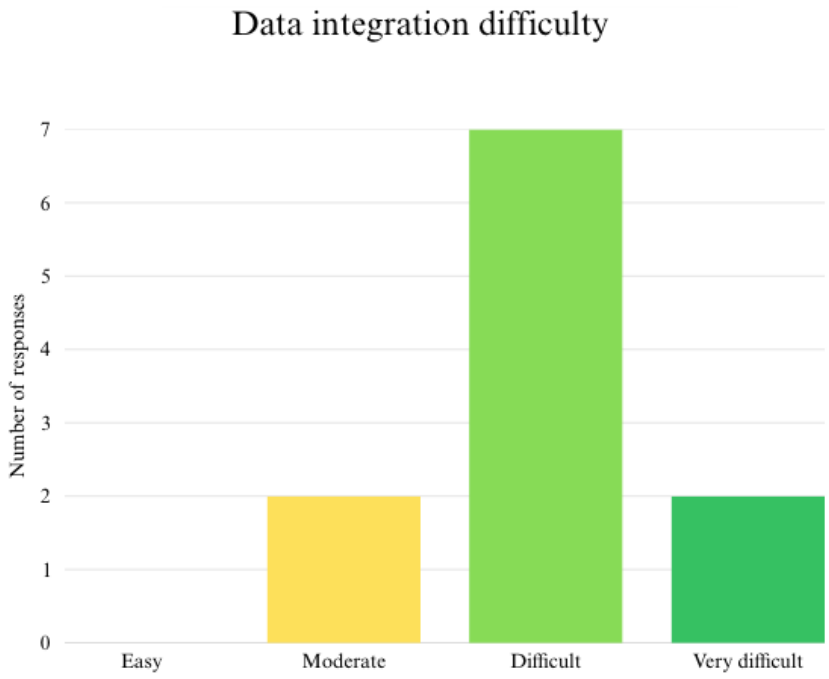

Figure 6: Bar chart of reported data integration difficulty.

These findings suggest that technical integration remains a persistent challenge, especially in environments with legacy systems or inconsistent data formats.

4.2.2 Link to EUDR Compliance

Respondents were also asked whether the degree of successful integration influenced their organization's ability to comply with EUDR requirements. A majority (6 out of 11) reported a "strong" or "very strong" connection between integration success and compliance effectiveness. Only 1 respondent indicated there was "no relationship."

4.2.3 Qualitative Insights (Interviews)

Interviewees consistently emphasized the operational burden of merging satellite, farmer, and certification data into internal platforms. One expert working in commodity traceability for a mid-sized firm noted that their team struggled to synchronize spatial data (GIS) with business systems like SAP and that these integration problems often delayed compliance reporting. Another interviewee, from a company in the cocoa sector, described similar issues and emphasized that "integration is not just a technical problem, it's a coordination problem across

departments." These reflections echo the survey results and highlight the importance of internal alignment and system compatibility in meeting regulatory expectations.

4.3 Data Quality and Validation

This section explores how companies manage the accuracy and transparency of deforestation-related data, focusing on governance structures, error rates, and the tools used to validate incoming data streams.

4.3.1 Governance Models and Error Rates

Most respondents (9) indicated that their company has implemented a formal data governance system. Among these, centralized models were most common, reported by over half of participants. However, in practice, many centralized models incorporate supplier-led elements, as companies often depend on external providers (such as satellite platforms or upstream suppliers) for geolocation and traceability data. One-third of respondents explicitly reported using hybrid models, combining centralized oversight with supplier-led data collection and validation (Pérez-Aleman & Sandilands, 2008). Only one respondent reported using a farmer-led governance model.

When asked whether governance structure influenced data quality, responses were divided. Two participants believed centralized governance led to more errors, while one other believed there was no significant difference. The remaining participants were unsure. This lack of consensus suggests that governance style alone may not determine accuracy; implementation quality and internal capacity likely play more important roles (Ruder & Wittman, 2025).

4.3.2 Transparency and Validation Practices

Most respondents (7) reported that implementing governance frameworks led to greater data transparency. More than half described the improvement as "significant," while others reported moderate gains, as seen in Figure 7. This trend reflects a broader shift toward enhanced internal oversight in preparation for EUDR enforcement.

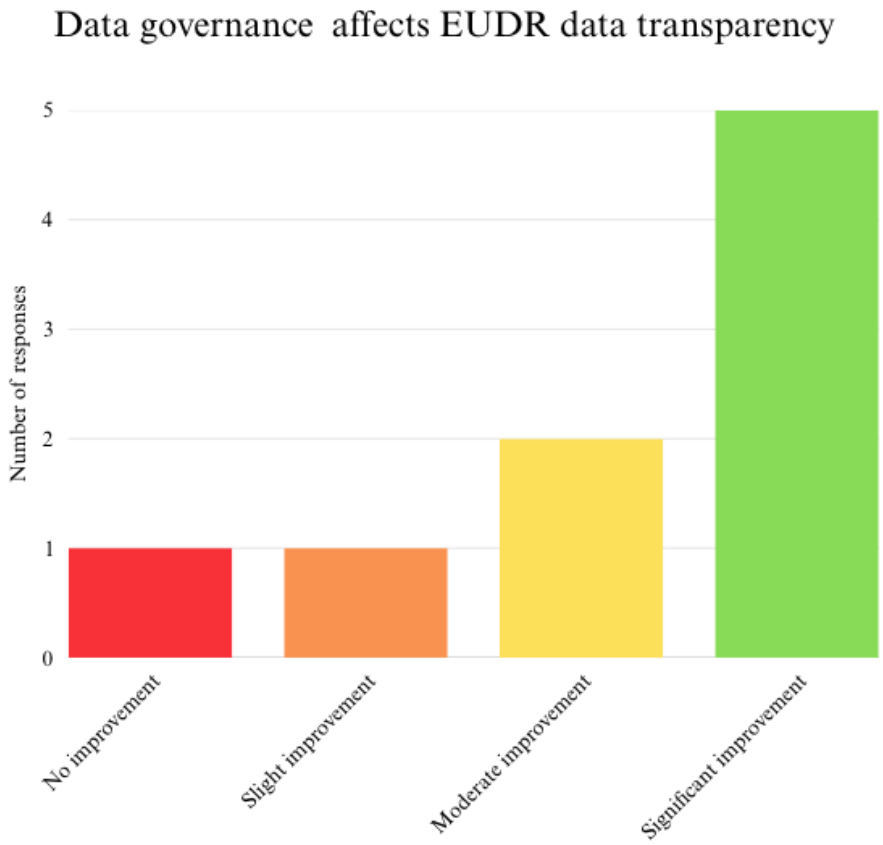

Figure 7: Perceived Impact of Data Governance on EUDR Transparency.

Data validation appears to be a mature practice among respondents. According to six respondents they validate more than 75% of their data before reporting for compliance. Other respondents validate for lower percentages (2 respondents validate 25% and 1 respondent 50% of the data). Seven respondents reported using manual validation techniques, while 6 respondents use third-party audits and another 6 respondents cross-check their data with satellite imagery. Automated validation systems, such as Meridia Verify, were used by five respondents. Some companies also reported internal custom tools, ad hoc audits or governmental data (Figure 8).

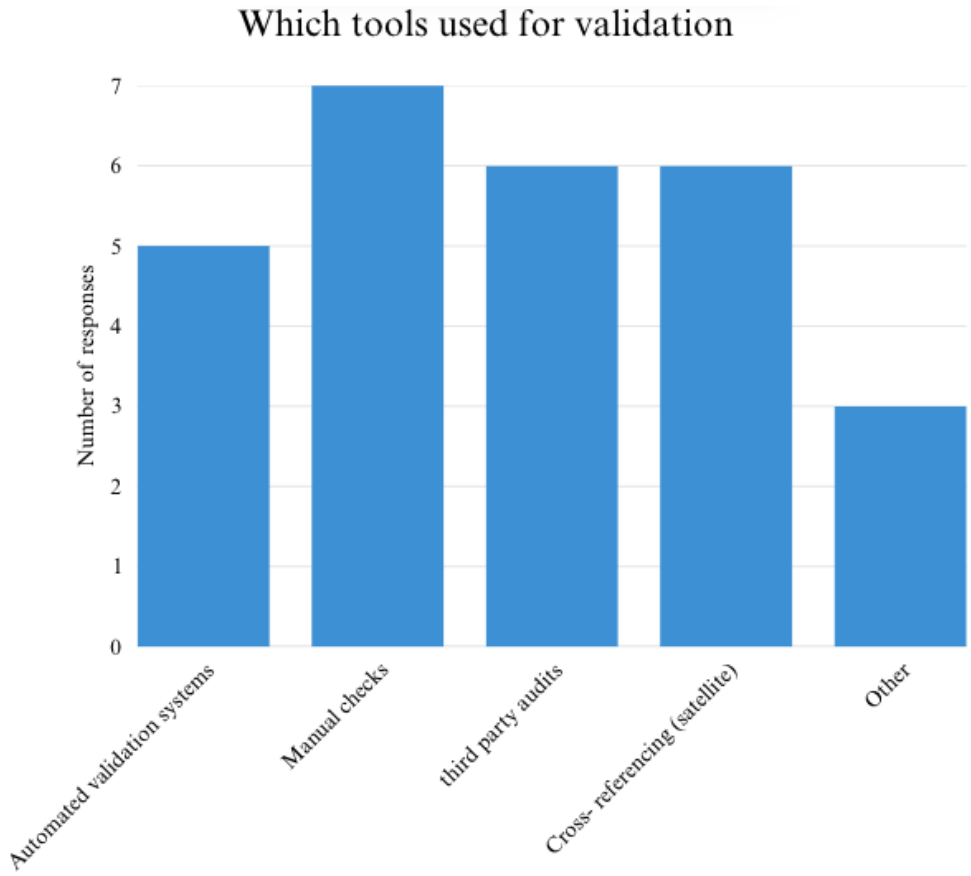

Figure 8: *Tools used for validation of deforestation related data.*

These responses indicate that while manual methods are still widely used, there is increasing adoption of more advanced and scalable technologies to enhance reliability.

4.3.3 Qualitative Insights (Interviews)

Interviewees emphasized that validation is not only about tools but also about internal discipline and the interpretation of inconsistencies. Several interviewees noted that upstream suppliers and cooperatives often lack the capacity or oversight needed to reliably validate field data. This puts pressure on downstream firms to develop additional safeguards or manual checks. In some cases, supplier-submitted data failed to meet EUDR traceability requirements, prompting downstream actors to intervene directly in data collection and verification. One interviewee noted that while automated checks help scale operations, they still rely on human review to flag context-specific anomalies such as recent land-use changes or informal plot expansions.

Another expert observed that more errors often emerge under centralized governance, not due to poorer data, but because internal checks are more rigorous. “It’s not that the data is worse,

it's that you're finally looking at it closely," they explained. This comment reinforces the idea that governance frameworks help reveal, rather than cause, data inconsistencies.

4.4 Compliance and Reporting

This section explores how companies utilize BI tools, such as TraceX, Satelligence, SAP, and Power BI, to support EUDR compliance reporting, and identifies common challenges associated with these tools.

4.4.1 Effectiveness of BI Tools for Reporting

Survey results suggest that BI tools are seen as useful for EUDR reporting. Four respondents rated them as effective, while another four considered them neutral. Three respondents found them ineffective. This distribution indicates that while these tools hold potential, their impact varies significantly based on implementation context.

4.4.2 Challenges in BI Implementation

The most frequently reported limitation was poorly structured internal data, noted by two respondents. Additionally, one respondent cited a lack of standardization across systems as a barrier to effective BI tool use.

Interview responses aligned with these observations. One interviewee explained that even though their company had adopted a well-known platform, the tool could not function optimally because foundational data was disorganized and inconsistently formatted. This supports the idea that BI tool performance is heavily dependent on internal data quality and integration capability.

4.4.3 Perceived Value Drivers of BI Tools

Among respondents who rated BI tools as effective, the most commonly cited reasons were: 1) smoother integration with satellite, certification, and field data (2 responses), 2) increased internal collaboration and transparency (1 response), and 3) reduced manual effort for report generation (1 response). These insights underscore that BI tools offer the most value when paired with interoperable data and strong internal workflows.

4.4.4 Qualitative Insights (Interviews)

One interviewee noted that while BI platforms provided useful dashboards and audit-ready documentation, the setup process was time-intensive and required dedicated internal resources. Another participant emphasized that effective use of BI tools depends less on the tool itself and more on upstream data structuring and ownership clarity.

5. Discussion & Conclusion

This thesis set out to answer the central research question: *What data management practices ensure EUDR compliance efficiency within companies?* To address this, the study explored how smallholder farmers contribute to data collection and whether their involvement enhances data quality; how companies integrate satellite, certification, and farmer-generated data; what governance structures are used to ensure data accuracy and transparency; and how business intelligence tools support EUDR reporting and what limitations they face.

The research was guided by a conceptual model (Figure 1) comprising four interlinked components: (1) governance models, (2) social governance, (3) core data management processes (collection, integration, validation), and (4) compliance and sustainability outcomes. Social governance, encompassing farmer participation, cooperative mediation, and data ownership, is treated as a cross-cutting force influencing every stage of the data pipeline. This framework highlights that compliance is not merely a technical exercise, but a socio-technical process shaped by power, participation, and inclusivity.
To align theory with findings, Figure 9 presents an adapted version of the model, mapping insights onto each stage of the data pipeline. It identifies key weaknesses, constraints, and opportunities, reinforcing the argument that both technical systems and governance dynamics shape the effectiveness and equity of EUDR compliance.

Insights

Data Collection

To address the first sub-question: *What role do smallholder farmers play in data collection, and how does their involvement influence data quality?* The findings reveal that deforestation-related data collection in EUDR-relevant supply chains remains predominantly centralized, relying heavily on third-party systems such as satellite providers and other companies. Direct involvement of smallholder farmers is minimal. When farmers do participate, their contributions are largely technical, limited to providing GPS coordinates, crop types, and harvest information, rather than strategic or governance-related. This reflects their continued role as data subjects rather than co-owners or decision-makers within compliance systems (Development Gateway, 2023). Despite this limited involvement, survey respondents consistently reported that farmer participation, even when modest, tends to improve data

accuracy. This suggests an underutilized opportunity: while smallholder input is often overlooked in the design of traceability frameworks, it has the potential to enhance data reliability and should therefore be more deliberately incorporated into future compliance strategies. These findings correspond to the "Input" and "Social Governance" elements of the conceptual model, highlighting how limited farmer participation weakens data quality and signals missed opportunities for cooperative mediation and ownership at the entry point of EUDR compliance systems.

Data Integration

The next sub-question explored is: *How effectively do companies merge satellite, certification, and farmer-provided data into usable systems?* The study finds that data integration remains a major challenge for companies seeking to comply with the EUDR. On average, only 43% of deforestation-related data, such as satellite imagery, certification documents, and farmer-provided inputs, is successfully integrated into internal systems. This limited capacity stems not only from technical issues like incompatible data formats but also from organizational barriers, including departmental silos and insufficient coordination between IT, sustainability, and procurement teams. Expert interviews reinforced that integration is "not just a technical problem, it's a coordination problem," highlighting the importance of internal alignment. Importantly, respondents that reported higher levels of integration also perceived themselves as better prepared for EUDR compliance, suggesting that integration is not merely a back-end concern but a core enabler of regulatory readiness. These findings emphasize the need for firms to invest in interoperable systems and cross-departmental collaboration to translate raw data into actionable compliance insights. This insight reinforces the "Core Data Management" component of the conceptual model, emphasizing that fragmented technical infrastructure and siloed coordination obstruct the integration processes that are foundational to effective compliance outcomes.

Data Validation and Governance

To answer the third sub-question: *What governance structures are used to ensure data accuracy and transparency, and how do they influence error rates?* The findings show that EUDR data validation remains predominantly centralized, with most companies managing accuracy and oversight internally. Only one company reported a farmer-led model, reflecting broader concerns that smallholders are excluded from data decision-making.

Hybrid governance structures are beginning to emerge, in which suppliers or cooperatives validate data under centralized oversight. These offer some flexibility but deliver inconsistent accuracy, often depending on the capacity of upstream actors. Validation methods are increasingly hybrid as well, blending manual checks, third-party audits, and satellite-based tools, highlighting the complexity and resource intensity of ensuring accuracy across fragmented supply chains. (Ruder & Wittman, 2025). These findings relate to the "Governance Model" dimension of the conceptual model, showing that centralized and hybrid structures shape how data quality is monitored and how inclusion or exclusion of upstream actors affects transparency and trust.

Compliance and Reporting

Finally, the last sub-question examined is: *How are business intelligence tools used to generate EUDR reports, and what challenges or factors emerge?* Another key finding relates to the use of BI tools in supporting EUDR compliance. Their effectiveness is heavily contingent upon the quality and structure of upstream data. As the findings show, companies with fragmented or unstructured internal datasets struggle to fully leverage these tools, often requiring costly customization and extensive internal coordination. Contrary to vendor claims, BI tools are not "plug-and-play" solutions; rather, their utility depends on prior investments in system integration, standardization, and data governance. Moreover, most BI platforms remain focused on compliance documentation rather than participatory features, lacking mechanisms to include smallholders in data validation or reporting workflows. This raises important questions about their long-term value and whether these tools can contribute not only to regulatory efficiency but also to inclusive and sustainable data ecosystems. Interestingly, despite the widely documented challenges in integrating satellite, certification, and field-level data, a small number of survey respondents reported smoother integration when using BI tools. This contradiction likely reflects differences in internal data maturity. Firms with more advanced digital infrastructures, such as pre-integrated traceability platforms or harmonized data formats, are better positioned to realize the benefits of BI systems. In these cases, tools function effectively not because integration is easy, but because foundational integration barriers have already been addressed. This nuance reinforces that BI tools are not standalone solutions but amplifiers of pre-existing data system quality.

Role of Suppliers

A key insight concerns the inevitable yet potential questionable role of suppliers in EUDR compliance. Interviews revealed that while suppliers often serve as the main data providers, particularly in upstream segments of the supply chain. They frequently deliver information that is incomplete, inconsistently formatted, or lacks proper validation. Several experts noted that suppliers often lack a full understanding of EUDR requirements, leading to data submissions that fail to meet regulatory standards, such as incorrect geolocation coordinates or missing timestamps. These shortcomings pose major obstacles for downstream companies, which must either perform costly corrections or find alternative data strategies. In response, some firms have begun bypassing suppliers altogether, deploying their own origin teams to directly collect farm-level data. While this approach significantly enhances data accuracy and traceability, it demands greater internal resources and limits scalability. These findings suggest that although suppliers are currently embedded in many compliance models, their effectiveness is highly variable and often undermines the reliability of EUDR-related data, reinforcing the importance of reevaluating supplier-led governance structures in favor of more participatory or cooperative-led alternatives.

Social Governance and the Role of Cooperatives

A key insight from this research is the disconnect between the conceptual ideal of farmer-centric, participatory governance and its limited implementation in practice. While the literature and the conceptual model emphasize the value of inclusion, trust, and shared data ownership, the findings reveal that smallholders remain largely excluded from governance processes. Data ownership and decision-making power continue to reside with downstream actors, such as buyers, suppliers, and third-party platforms.

Farmer cooperatives, in theory, could bridge this gap. Interviewees noted that cooperatives often serve as the primary intermediaries for collecting, verifying, and transmitting smallholder data, particularly in cocoa supply chains. However, many cooperatives face challenges including limited digital literacy, weak internal capacity, and inconsistent governance structures. These constraints undermine their ability to support participatory models effectively. Moreover, their influence varies by sector: cooperatives play a central role in cocoa, but are largely absent in palm oil, where plantation or supplier-led models dominate. Most BI and traceability tools used for EUDR compliance also lack structured feedback loops or governance mechanisms that would allow farmers or cooperatives to access or influence how data is used. This reinforces the asymmetry in current compliance systems and validates

the inclusion of “Social Governance” as a transversal axis in the conceptual model. Ultimately, without investments in cooperative capacity and the development of participatory frameworks, even the most advanced technical systems will fall short of achieving credible, inclusive, and equitable EUDR compliance.

These insights strongly support the inclusion of the “Social Governance” axis in the conceptual model, underscoring that effective EUDR compliance depends not just on technical infrastructure, but also on who controls, contributes to, and ultimately benefits from the data ecosystem. By re-centering data governance on inclusivity and farmer participation, companies can enhance the accuracy, legitimacy, and sustainability of their compliance efforts. This thesis contributes to the broader literature by illustrating that data management for regulatory compliance is not a neutral or purely technical exercise, it is a socio-technical process shaped by institutional design, digital inequality, and power asymmetries across global supply chains. As the EUDR enters its implementation phase, the findings emphasize that compliance systems built for reporting to systems should also be built for participation. Empowering cooperatives, creating mechanisms for shared data ownership, and ensuring that smallholders are not just data providers but decision-makers, will be crucial in aligning regulatory efficiency with sustainability and equity goals. This dynamic is visually captured in Figure 9, which maps the findings onto the conceptual framework and highlights critical weaknesses, persistent constraints, and potential levers for improvement across the EUDR data governance landscape.

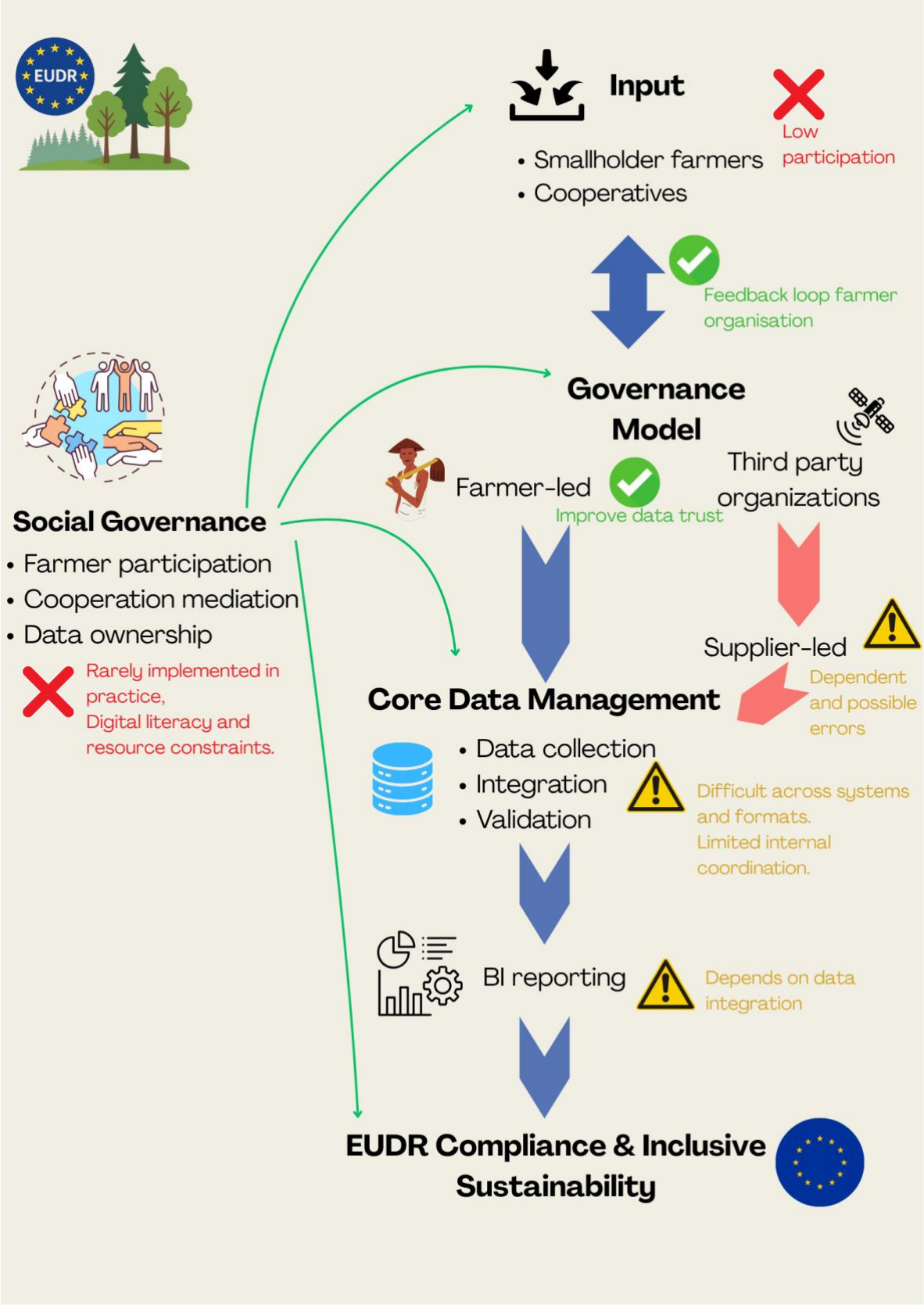

Figure 9: Adapted Conceptual Model Annotated with Data Management Findings under EUDR. This figure visualizes how the conceptual framework developed in Chapter 2 maps onto the findings. Annotations indicate areas of success (green check mark), constraint or

complexity (yellow danger exclamation mark), and critical weaknesses (red cross) across governance models, data systems, and farmer participation under EUDR compliance. Where social governance plays a role over the whole supply chain.

Implications, Limitations, and Future Research

This study offers several practical implications for companies, policymakers, and technology providers navigating EUDR compliance. First, enhancing data integration and regulatory readiness demands investment in interoperable systems and better coordination between IT, sustainability, and procurement departments, an insight supported by both survey responses and expert interviews (Ahoa et al., 2025; Satelligence, 2023). Second, in light of the inconsistent quality of supplier-provided data, firms may need to reassess the dominance of supplier-led governance models. As interviews revealed, some companies are now bypassing suppliers by deploying origin teams to directly collect geospatial and farm-level data, an approach that improves accuracy but increases resource demands. Third, smallholder inclusion remains a critical but underdeveloped dimension. While farmer participation is likely to improve data accuracy, few companies incorporate farmers into governance or feedback loops. This calls for participatory models that empower cooperatives and promote data ownership, consistent with recommendations by Development Gateway (2023) and Ruder & Wittman (2025).

However, the study is not without limitations. Its small survey sample size (n=11) and limited number of interviews (n=3) restrict generalizability, positioning the findings as exploratory rather than representative. In addition, while cross-sector insights were gathered, sector-specific patterns (cocoa vs. palm oil supply chains) warrant more focused examination. For instance, the role and capacity of cooperatives vary substantially between commodities, impacting data governance efficacy. Future research should expand on these variations by conducting larger, sector-specific studies and longitudinal assessments of farmer-led governance models. Additionally, given the reliance on third-party platforms like Satelligence or TraceX, future work could evaluate how these tools evolve to support inclusive data governance and long-term traceability performance (2024; LiveEO, n.d.). A deeper understanding of how social governance mechanisms, like participatory validation and cooperative-led data oversight, interact with technical systems will be essential to building transparent, equitable, and resilient compliance infrastructures.

References

- Abubakari, M. Y., & Sarpong, D. B. (2022). *What factors influence the likelihood of rural farmer participation in digital agricultural services? Experience from smallholder digitalization in Northern Ghana*. ResearchGate. https://www.researchgate.net/publication/366597491
- Achilles. (n.d.). Achilles white paper: Data management for sustainable supply chains. Retrieved from https://www.achilles.com
- Ahoa, E., Kassahun, A., Verdouw, C., & Tekinerdogan, B. (2025). Challenges and solution directions for the integration of smart information systems in the agri-food sector. Sensors, 25(8), 2362. https://doi.org/10.3390/s25082362
- Amar, S., & Beranek, N. (2024). Smallholder-oriented Data governance principles . Maastricht University.
- BanQu. (n.d.). Retrieved from https://banqu.co
- Creswell, J. W., & Creswell, J. D. (2018). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*.
- Curtis, P. G., Slay, C. M., Harris, N. L., Tyukavina, A., & Hansen, M. C. (2018). Classifying drivers of global forest loss. Science, 361(6407), 1108–1111. https://doi.org/10.1126/science.aau3445
- Development Gateway. (2023). Data inclusion in smallholder agriculture: Pathways to empowerment. Retrieved from https://www.developmentgateway.org
- Development Gateway. (2023). Farmer-Centric Data Governance: Towards a new paradigm. Retrieved from https://developmentgateway.org/wp-content/uploads/2023/02/Farmers_Report-Full_vFebFiNAL.pdf
- European Union. (2023). Regulation (EU) 2023/1115 on deforestation-free products. Official Journal of the European Union.
- Fair Labor Association. (2023). Mapping working conditions and child labor risks in Nestlé's cocoa supply chain in Cameroon. https://www.fairlabor.org/reports/mapping-working-conditions-and-child-labor-in-nestles-cocoa-supply-chain-in-cameroon/
- Fairtrade Foundation. (2021). Cocoa farmers and climate change. https://www.fairtrade.org.uk

- Fink, A. (2017). How to Conduct Surveys: A Step-by-Step Guide (6th ed.). Sage Publications.
- Food and Agriculture Organization. (2020). The state of the world's forests 2020: Forests, biodiversity and people. FAO. https://www.fao.org/documents/card/en/c/ca8642en
- Gardner, T. A., Benzie, M., Börner, J., Dawkins, E., Fick, S., Garrett, R., Godar, J., ... & Wiegand, K. (2019). Transparency and sustainability in global commodity supply chains. *World Development*, 121, 163–177. https://doi.org/10.1016/j.worlddev.2018.05.025
- Gaviota, M. M. (2025, May 28). Overcoming barriers in geolocation data: An interview with Koltiva's GIS expert for deforestation monitoring & supply chain transparency. Koltiva. https://www.koltiva.com/post/post-geolocation-data-gis-expert-for-deforestation-monitoring-and-supply-chain-transparency
- Gocsik, É., Bekamp, B., & Valeeva, N. (2024, April). The EU Deforestation Regulation is a complex and costly undertaking. Rabobank RaboResearch Food & Agribusiness. https://media.rabobank.com/m/49f1008d03f026ea/original/The-EU-Deforestation-Regulation-is-a-complex-and-costly-undertaking.pdf
- Intergovernmental Panel on Climate Change (IPCC). (2019). Climate Change and Land – Special Report. https://www.ipcc.ch/srccl/
- Jopke, P., & Schoneveld, G. C. (2020). Corporate promises in tropical forest agriculture: A review of zero deforestation commitments and their implementation (Occasional Paper 181). Center for International Forestry Research (CIFOR). https://www.cifor-icraf.org/publications/pdf_files/OccPapers/OP-181.pdf
- Khatri, V., & Brown, C. V. (2010). Designing data governance. Communications of the ACM, 53(1), 148–152. https://doi.org/10.1145/1629175.1629210
- Lambin, E. F., Gibbs, H. K., Heilmayr, R., et al. (2018). The role of supply-chain initiatives in reducing deforestation. Nature Climate Change, 8(2), 109–116. https://doi.org/10.1038/s41558-017-0061-1
- LiveEO. (n.d.). TradeAware – Deforestation detection and EUDR compliance. Retrieved June 1, 2025, from https://www.live-eo.com/product/tradeaware
- Merriam, S. B., & Tisdell, E. J. (2016). *Qualitative Research: A Guide to Design and Implementation*.
- Nestlé. (2023). Towards Forest Positive Cocoa: Nestlé Cocoa Plan Progress Report 2022. Nestlé Cocoa Plan. Retrieved from

https://www.nestlecocoaplan.com/sites/default/files/2023-08/NEST7399_22_NCP-Progress-Report-2022_V19_1.pdf

- Orbify. (2024, December 21). Navigating EUDR compliance: The best SaaS solutions for deforestation monitoring. https://orbify.com/blog/navigating-eudr-compliance-the-best-saas-solutions-for-deforestat ion-monitoring
- Passionfruit. (2025, March 7). Top EUDR compliance tools to prepare your supply chain. https://www.passionfruit.earth/blog/top-eudr-compliance-tools-supply-chain?
- Pérez-Aleman, P., & Sandilands, M. (2008). Building value at the top and the bottom of global supply chains: MNC–small producer collaboration in the Agri-food sector. Global Networks, 8(2), 125–154. https://www.researchgate.net/publication/270024009_Building_Value_at_the_Top_and_the_Bottom_of_the_Global_Supply_Chain_MNC-NGO_Partnerships
- Rainforest Alliance. (2022). Retrieved from https://www.rainforest-alliance.org/business
- Rainforest Alliance. (2023). How the Rainforest Alliance Supports EUDR Compliance from Farm to Retailer. Retrieved from https://www.rainforest-alliance.org/business/certification/how-the-rainforest-alliance-supports-eudr-compliance-from-farm-to-retailer/
- Renier, C., Vandromme, M., Meyfroidt, P., Ribeiro, V., Kalischek, N., & Zu Ermgassen, E. K. H. J. (2023). Transparency, traceability and deforestation in the Ivorian cocoa supply chain. Environmental Research Letters, 18(11), 114009. https://www.researchgate.net/publication/367653840_Transparency_traceability_and_deforestation_in_the_Ivorian_cocoa_supply_chain
- Reuters. (2024, October 1). *Cocoa traceability rates fail to improve as EU deforestation law looms*. Reuters. https://www.reuters.com/markets/commodities/cocoa-traceability-rates-fail-improve-eu-deforestation-law-looms-2024-10-01/
- RSPO. (2022). Certified Smallholders. Roundtable on Sustainable Palm Oil. https://rspo.org/nl/
- Ruder, S.-L., & Wittman, H. (2025). Agricultural data governance from the ground up: Exploring data justice with agri-food movements. Big Data & Society, 12(1). https://doi.org/10.1177/20539517251330182

- Satelligence. (2023). EUDR compliance solutions for coffee and cocoa supply chains. Retrieved June 1, 2025, from https://satelligence.com/eudr-compliance-for-the-coffee-sector/
- Saunders, M., Lewis, P., & Thornhill, A. (2019). *Research Methods for Business Students*.
- Skalkos T. (2023). Prospects, Challenges and Sustainability of the Agri-Food Supply Chain in the New Global Economy II. Retrieved from https://www.researchgate.net/publication/373228177_Prospects_Challenges_and_Sustainability_of_the_Agri-Food_Supply_Chain_in_the_New_Global_Economy_II
- SupplyChainBrain. (2023). How to Overcome the Top Four Challenges of EUDR Compliance. Retrieved from https://www.supplychainbrain.com
- TraceX Technologies. (2023). Unlocking traceability: The role of integrated systems in EUDR compliance. Retrieved from https://www.tracextech.com
- TraceX Technologies. (2024). 10 features in EUDR compliance solution. Retrieved June 1, 2025, from https://tracextech.com/10-features-in-eudr-compliance-solution

Appendix

A. Survey

Section 1: Data Collection

Q1
W*hich of the following sources does your company use to collect deforestation-related data? (Select all that apply)*
☐ Smallholder farmers (direct reporting)
☐ Farmer cooperatives or producer organizations
☐ Certification bodies (e.g., Rainforest Alliance, RSPO)
☐ Satellite imagery providers (e.g., Satelligence)
☐ Internal field staff or agronomists
☐ Other: _____

Q2
Which of these sources contributes most significantly to your data collection? (Select all that apply)
☐ Smallholder farmers (direct)
☐ Farmer cooperatives
☐ Farmer cooperatives or producer organizations

☐ Satellite data
☐ Internal staff
☐ Other: _____

Q3
To what extent are farmers involved in the data collection process (e.g., providing geolocation, yield data, or plot-level verification)?
☐ Not involved
☐ Minimal involvement (e.g., providing basic field details like farmer name, farm size, or crop type)
☐ Moderate involvement (e.g., confirming GPS boundaries or submitting paper-based records)
☐ Active involvement (e.g., submitting digital records regularly or providing feedback during audits)
☐ High involvement (e.g., co-designing data collection tools or participating in validation decisions)

Q4
What type of data do they typically provide? (Select all that apply)
☐ Crop type or yield estimates
☐ Harvest or planting dates
☐ Land-use history
☐ Certification or compliance documents
☐ Other: _____

Q5
To what extent does farmer involvement improve the accuracy of the deforestation-related data your company collects?
☐ Not at all
☐ Slightly
☐ Moderately
☐ Significantly
☐ Very significantly

Q1.3
How often does your company rely on farmer organizations (cooperatives or farmer groups) to collect and organize deforestation-related data from smallholder farmers?
☐ Never
☐ Rarely (less than 25% of interactions)
☐ Occasionally (25-50% of interactions)
☐ Frequently (50-75%of interactions
☐ Always (more than 75% of interactions)

Section 2: Data Integration

Q6
What percentage of your deforestation-related data is successfully integrated into your company's business intelligence or reporting system (e.g. TraceX, SAP, Power BI…)?
☐ 0–25%
☐ 26–50%
☐ 51–75%
☐ 76–100%

Q7
How would you rate the relationship between your ability to integrate data and the ability to meet EUDR compliance requirements?

☐ No relationship
☐ Weak relationship
☐ Moderate relationship
☐ Strong relationship
☐ Very strong relationship

Q8
Rate the difficulty your company faces when integrating diverse data sources (satellite imagery, certification records, farmer data).
☐ Very easy
☐ Easy
☐ Moderate
☐ Difficult
☐ Very difficult

Section 3: Data Quality and Validation

Q9
Does your company have a formal system or framework in place to manage how sustainability or traceability data is collected, stored, and used (for EUDR compliance)?
☐ Yes ☐ No ☐ Not sure

If yes:

Q10
Which type of data governance model best describes your company's current approach?
☐ Centralized (data collected and managed by your company or external providers)
☐ Farmer-led (data collected or managed by cooperatives, producer groups, or farmers directly)
☐ Hybrid (both centralized and farmer-led models coexist)
☐ Other: _____
☐ Not sure

Q11
Based on your experience, how do data error rates vary across different governance models?
☐ Higher in centralized models
☐ Higher in farmer-led models
☐ About the same
☐ Not applicable / no comparison
☐ Not sure

Q12
Since implementing a data governance framework, how has your company's transparency around deforestation-related data changed?
☐ No improvement
☐ Slight improvement
☐ Moderate improvement
☐ Significant improvement
☐ Drastically improved

Q13
How frequently does your company validate collected data before reporting for compliance?
☐ Never

☐ Rarely (less than 25% of data
☐ Occasionally (25-50% of data)
☐ Frequently (50-75%)
☐ Always (>75%)

Q14
Which of the following tools does your company use for data validation? Select all that apply.
☐ Automated validation systems (Meridia Verify…)
☐ Manual checks
☐ Third-party audits
☐ Cross-referencing with satellite imagery
☐ Other:

Section 4: Compliance and Reporting

Q15
How would you rate the performance of business intelligence tools (e.g. TraceX, SAP, Power BI…) in generating compliance reports?
☐ Very ineffective
☐ Ineffective
☐ Neutral
☐ Effective
☐ Very effective

IF Ineffective:

Q16
What are the main limitations or challenges you experience with your current business intelligence tools?
☐ Tools lack integration with key data sources
☐ Require high technical expertise or external support
☐ Reporting is rigid or difficult to customize
☐ Tools are expensive to implement or operate
☐ Internal data is not structured well for BI use
☐ Other (please specify): _____
IF Effective

Q17
What best explains why you consider these tools effective in generating EUDR compliance reports?
☐ Tools reduce manual reporting effort
☐ Integration with key data sources (e.g., satellite imagery, certification records, farmer data) works smoothly
☐ Reporting is fast and highly customizable
☐ Tools support automated validation or risk assessment
☐ Tools increase internal transparency and collaboration
☐ Other (please specify): _____

B. Interview Guide

Interview Guide: Expert Insights on EUDR Data Management

The following semi-structured interview guide was used to conduct expert interviews with professionals involved in EUDR compliance, sustainability reporting, and data governance. The guide focused on four thematic areas aligned with the research dimensions: data collection, data integration, data validation & governance, and compliance reporting.

Section 1: Data Collection

- Can you describe how deforestation-related data is collected in your organization?
- What role do smallholder farmers play in this process (if any)?
- Are farmer cooperatives or producer groups involved in facilitating data collection?
- Which sources (e.g., satellite, certification, field agents) are considered most reliable? Why?

Section 2: Data Integration

- How is deforestation-related data integrated into your internal systems (e.g., ERP, SAP, traceability platforms)?
- What are the main challenges you face in integrating diverse data sources (e.g., satellite, farmer, certification)?
- Do you use any external service providers or custom-built tools to support integration?

Section 3: Data Validation and Governance

- How does your organization validate the data before using it for EUDR reporting?
- What kind of governance model is in place (centralized, supplier-led, farmer-led, hybrid)?
- Have you encountered any issues related to data quality or trust in different parts of your supply chain?

Section 4: Compliance and Reporting

- Which business intelligence or reporting tools do you use for EUDR compliance?
- How effective are these tools in generating traceability and compliance reports?
- What are the main limitations or strengths of your current compliance workflow?

Closing

- Do you have any additional recommendations or lessons learned that could benefit others?

Statements

Official statement of original thesis

By signing this statement, I hereby acknowledge the submitted thesis (hereafter mentioned as "product"), titled:

A Deforestation-Free Future: Leveraging Farmer-Centric Data Governance for EUDR Compliance and Sustainabilityto be produced independently by me, without external help. Wherever I paraphrase or cite literally, a reference to the original source (journal, book, report, internet, etc.) is given.

By signing this statement, I explicitly declare that I am aware of the fraud sanctions as stated in the Education and Examination Regulations (EERs) of the SBE.

Place: Maastricht

Date: 20-06-2025

First and last name: Paul Geerlings

Study programme: International Business - Information Management & Business Intelligence

Course/skill: Thesis

ID number: i6406755

Signature: 

Statement on the use of Generative AI (GenAI) in the master thesis

I hereby certify that I adhered to the SBE guidelines on the use of GenAI tools such as ChatGPT in the master thesis. In the box below, I document how and for what purposes I used GenAI.

During the preparation of this work, I used GenAI for the following purposes:

- Search engine: ChatGPT/Perplexity; To help contextualize certain problems to help the search for academic sources.
- Explanation provider: ChatGPT; Explaining of concepts and give real world examples.
- Language assistant: ChatGPT; only to spot spelling errors.

After using any tool, I reviewed, quality-checked, and edited the content as needed and take full responsibility for the content of the thesis.

By signing this statement, I explicitly declare that I am aware of the fraud sanctions as stated in the Education and Examination Regulations (EERs) of the SBE.

Place: Maastricht

Date: 20-06-2025

First and last name: Paul Geerlings

Study programme: International Business - Information Management & Business Intelligence

Course/skill: Thesis

ID number: i6406755

Signature: 

Sustainable Development Goals (SDG) Statement

Name Paul Geerlings

ID I6406755

Supervisor Sidi Amar

Date 20-06-2025

Through the research conducted for this master's thesis, I seek to contribute to one or more of the 17 SDG(s) set forth by the United Nations (https://www.undp.org/sustainable-development-goals). Specifically:

SDG Code(s): 15

Explanation: The research for this thesis aligns with 'Life on Land', which focuses on sustainably managing forests, combating deforestation, and promoting biodiversity. By exploring how companies can optimize data management processes to comply with the European Union Deforestation Regulation (EUDR), the research directly supports efforts to reduce deforestation in global supply chains. The study emphasizes integrating farmer-centric governance, ensuring smallholder farmers play a key role in data collection and sustainable practices, which aligns with the goal of promoting inclusive and equal solutions