Landscape Analysis

EXISTING MECHANISMS, GAPS, AND OPPORTUNITIES FOR AGRICULTURAL DATA SHARING

STRUCTURE OF THE LANDSCAPE ANALYSIS

Assumptions
A set of foundational assumptions that framed - and were affirmed by - the research for this landscape analysis. These were developed through practitioner interviews, first-phase analysis of the GLCDI, and analogical research drawn from adjacent sectors.

Leading Organizations
A non-exhaustive list of organizations in the beef and soil carbon sector that help to collect, aggregate, analyze, store, host, increase access to, and define standards for soils and management data.

Interview and Research Findings
Interviews with leading practitioners across regenerative agriculture, soil health research, and carbon markets, which surfaced consistent tensions and emergent opportunities within agricultural data governance.

Gaps
An outline of key observed gaps, including the structural absence of coherent data governance, interoperability standards, and value-aligned incentive mechanisms.

Opportunities
Descriptions of potential opportunities integrating findings from case studies, the assumptions section, interviews, and additional research.

Next Steps
Reflection questions assigned to the Steering Committee to advance thinking beyond what is contained in this landscape review.

Appendix: Example Case Studies
An overview of insights garnered from relevant case studies, including meteorological data governance, open-source software architectures, and Creative Commons digital image rights.

Addendum: Steering Committee Commentary
Additional comments or conversation prompts generated by the steering committee during review of this document in draft form.

Framing and Context of Emergent Tensions in Data Governance

The first phase of the Grazing Lands Carbon Data Initiative (GLCDI) revealed not only the magnitude of the data gap in soil carbon monitoring in the US, but also exposed deeply entrenched frictions in how data is collected, owned, and utilized across the agricultural sector. This initial discovery underscored that the challenge is not solely technical, nor merely a function of data sufficiency; it is fundamentally shaped by the architectures of governance, the distribution of power, lack of trust, and the underlying (and lack of) incentives that drive data collection and data sharing.

The data gap can be broadly understood as a misalignment between the amount, quality, and accessibility of desired data versus existing data - among data collectors and data users -that results in poor data value among stakeholders, inhibits effective knowledge-sharing, and results in negative perceptions of the data economy like owner-take-all mindsets. In the case of the livestock and grazing industries the data gap is the result of all three factors, as well as ranchers' lack of true incentives for data collection and mistrust of industry actors–especially technology providers.

Understanding the contours of the data gap, as well as existing infrastructure and governance that enables the sharing of said data, is critical for meeting the needs of various stakeholders in the beef industry, including but not limited to: using data to monitor soil health and soil carbon, optimize ranch profitability in line with sustainability, enable insights into conservation efforts, and support claims about such insights.

These early findings also raised critical unresolved questions: Does the private sector, which often holds key data assets, possess sufficient motivation to participate in open or semi-open data systems? Or is the construction of such systems destined to fall predominantly on public institutions, research bodies, and non-profit actors who are structurally better aligned with data commons ideologies? And, likely most crucially, what is necessary to increase rancher participation in data collection and sharing?

IN THIS SECOND PHASE OF THE GLCDI, WE POSED THE QUESTION:

What are the gaps in data sharing infrastructure and governance regarding soil carbon and grazing management data?

This Landscape Analysis phase of work seeks to interrogate these questions and assumptions directly. It tests the hypothesis that private actors (such as corporations and ag-tech service providers) lack compelling incentives to contribute to data sharing infrastructure unless direct commercial value or competitive advantage is realized. However, we further investigate whether this assumption is universally valid across all segments of the agri-food system. Are there emergent models or edge cases—perhaps visible in other industries—where hybrid governance arrangements have unlocked new forms of value, participation, or cooperation? The exploration of analogous sectors including meteorological data systems and the evolution of open-source software development platforms serve not as direct blueprints, but as provocations: Where have private actors engaged in public-good data ecosystems, in effect to close data sharing gaps, and under what governance terms are those gaps closed and maintained? Moreover, where have efforts at open data sharing collapsed under the weight of rent-seeking behavior or inadequate incentive structures?

EFFICIENCIES IN CONNECTIVITY INFRASTRUCTURE

To describe data, data sharing, and data sharing infrastructure, we can use maps as a visual analogy. Locations where data is held can be thought of as cities. Within a storage location, stakeholders are able to use the data that exists there for various different needs. But how does data travel from one location to another? What vectors are available to the users, or recipients, of that data between locations? (Fig. 1.)

Roads might represent this flow of data from one location to another. Roads may be supplanted by alternative ground transportation methods, or by no methods at all, and therefore reliant on communications and knowledge infrastructure to enable inferring of key aspects of the data. Further, roads are not the only form of transportation infrastructure. Railroad tracks serve as another example; cars cannot travel on train tracks and trains cannot travel on roads, so interoperability for transporting the units (humans, in the roads/trains example) that each method supports needs to be considered in the effective development of pathways between locations.

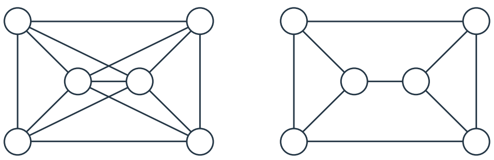


FIG.1. In the real world, efficient roadways are designed to maximize connectivity without developing duplicative pathways. This demonstrates the importance of considering the efficiency of connectivity for data sharing.
TERRA GENESIS 2025

Using this analogy, this Landscape Analysis of the GLCDI aims to understand the efficiency and interoperability of the infrastructure (roads) and governance (transportation policies and social norms) for sharing (moving) soil carbon data, as well as other environmental data.

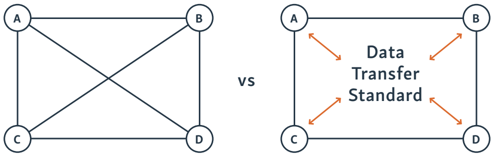


FIG.2. Effective data transfer emphasizes data flows between actors, and in the method best-suited to the nature of data sharing between those actors. A Data Transfer Standard may enable different forms of data to more easily interoperate.
TERRA GENESIS 2025

Further, a data transfer standard may facilitate the sharing of data and increase interoperability. Using the roads analogy further, one can acknowledge that while trains serve as an important form of transportation, the products they carry can most often only reach their end destination (and therefore generate value) if they can be transferred to a truck for road transport. A Data Transfer Standard may define how different data types may need to be formatted or packaged in order to be usable by other stakeholders or forms of sharing infrastructure. (Fig. 2.)

PRESENT CHALLENGES IN DATA GOVERNANCE

This phase of the GLCDI challenges the framing of data governance as purely a matter of technical standards or platform interoperability. It probes the cultural and epistemic dimensions of data sharing—particularly the language, metaphors, and power structures embedded in how agricultural data is framed, owned, and circulated. Does the persistence of data as private property, often guarded by technological intermediaries, reflect only market logics? Or does it also reveal deeper legacies of extractivism and control that resist simplistic technical fixes? The landscape analysis that follows uses these provocations as both lens and scalpel—testing assumptions from the GLCDI Discovery Phase against the lived realities of practitioners, the structural constraints of the sector, and the speculative possibilities of alternative governance models.

It is through this lens that the landscape analysis unfolds, not as a definitive mapping of the current state, but as an inquiry into the structures, incentives, and imaginaries that shape agricultural data governance. The analysis deliberately foregrounds case studies of adjacent industries, practitioner insights, and experimental models to stress-test assumptions surfaced in the Discovery Phase.

Rather than offering prescriptive solutions, the sections that follow expose the underlying architectures of data flow, the frictions of interoperability and sovereignty, and the governance tensions that must be navigated if more participatory, equitable, and outcome-focused data systems are to emerge.

Where possible, the analysis returns to a central, unresolved question: What conditions—not only technical, but political, economic, and cultural—must be created for agricultural data systems to evolve beyond the logics of privatization, toward models that distribute both the risks and benefits of data production and use more equitably?

Assumptions for Effective Data Sharing

A set of foundational assumptions developed through practitioner interviews, first-phase analysis of the GLCDI, and analogical research drawn from adjacent sectors form our core Assumptions. These assumptions are not intended to prescribe a fixed model for agricultural data governance. Rather, they surface working hypotheses about incentive structures, stakeholder roles, regulatory limitations, and the maturity of technology infrastructure in order to help orient future inquiry and program design. The purpose of this section is to provide the steering committee and research collaborators with a framework from which to test, refute, or refine governance options as the initiative evolves.

STAKEHOLDERS | **INCENTIVES** | **CHALLENGES** | **OPPOURTUNITIES**

ASSUMPTION 0

Without emphasis on value proposition for increasing producer data collection, data-sharing efforts will be futile
Data-sharing is downstream of the fundamental issue facing the industry; very few ranchers believe there is a good business case for collecting management and/or soils data. All realistic challenges and opportunities in the data sharing effort are informed by this barrier. Addressing the necessary incentives amongst stakeholders to increase producer data is key.

ASSUMPTION 1
Private sector motivation is currently insufficient to aid producer management and land stewardship.

ASSUMPTION 3
Disproportionate burden is placed on public sector and civil society.

ASSUMPTION 2
Public and non-profit actors dominate infrastructure development.

ASSUMPTION 4
Incentives for private sector engagement exist, but are limited.

ASSUMPTION 5
Rancher sharing data requires additional incentives beyond improving management decisions.

ASSUMPTION 6
Practice incentives are necessary, but are insufficient for systems-level change.

ASSUMPTION 9
Policy and regulation will likely not evolve fast enough to meet sector needs.

ASSUMPTION 10
Governance models must align with FAIR Principles and protect data sovereignty.

ASSUMPTION 7
A successful system requires multiple, interrelated benefits for involved actors.

ASSUMPTION 8
Governance models must align with FAIR Principles and protect data sovereignty.

ASSUMPTION 11
This sector is shaped by its language which must be critically examined.

Preliminary research for this Landscape Analysis provided the following assumptions. Unless otherwise noted, subsequent interviews and research reinforced these premises.
TERRA GENESIS 2025

These assumptions were shaped by input from technical stakeholders in data collection, management, and policy; insights from open-source and commons-based data initiatives; and the early discovery efforts conducted under the Discovery Phase.

They reflect both the current barriers observed in the field and potential leverage points. While some assumptions are already being tested through coalition engagement and comparative research, others remain open questions requiring validation. They are included here as a transparent basis for iteration, critique, and alignment.

Assumption 0

Without Emphasis On Value Proposition ? For Increasing Producer Data Collection, Data-Sharing Efforts Will Be Futile

While this workstream focused on barriers and opportunities for increasing data sharing, it is crucial to acknowledge that there are not sufficient incentives for producers to collect said data in the first place. Data-sharing is downstream of the fundamental issue facing the industry; very few ranchers believe there is a good business case for collecting management and/or soil data. An overemphasis on increasing data sharing pathways without addressing this pervasive issue will result in fragmented (and likely failed) efforts.

Assumption 1

Private Sector Motivation Is Currently Insufficient To Aid Producer Management And Effective Land Stewardship

The private sector lacks sufficient intrinsic incentives to invest heavily in the development of shared data infrastructure and governance. Many existing models rely predominantly on proprietary data systems, with limited evidence that private actors will voluntarily contribute to open data ecosystems without direct, short-term commercial benefits.

Assumption 2

Public and Non-Profit Actors Dominate Infrastructure Development

The majority of data standards development, infrastructure maintenance, enforcement, and data sharing are currently driven by open-source initiatives, government agencies, non-profits, and academic institutions. It is likely that meaningful data sharing infrastructure and protections would require policy intervention or legislation to enforce equitable participation. It should be noted that while public and non-profit actors lead the charge, additional funding is needed to support their work.

Assumption 3

Disproportionate Burden on Public and Civil Society Sectors

Given the inertia within the private sector, the burden of developing interoperable, trusted data-sharing infrastructure is falling disproportionately on public institutions, civil society, and non-profit organizations. While interviews highlighted the importance of engaging private actors, no clear governance mechanism is yet in place to operationalize this.

Assumption 4

Incentives for Private Sector Engagement Exist, But Are Limited

Currently, corporations are facing pressure from certain high-regulation markets (eg, the EU) to collect and provide agricultural data. But these regulations don't require this data be shared with industry-wide consortiums, scientific efforts, conservation orgs, etc. Potential additional incentives for this kind of private sector engagement include access to larger, aggregated datasets and potential reductions in data collection and acquisition costs through shared systems. Regardless, these incentives predominantly appeal to corporate entities and do not translate into mechanisms for systemic rancher participation.

Assumption 5

Rancher Data Sharing Requires Additional Incentives Beyond Improving Management Decisions

For producers, the incentive landscape is fundamentally different. Despite increasing focus on marketplace mechanisms (e.g., selling to sustainable beef markets or accessing carbon credits), there remains virtually no inherent and far-reaching incentive for farmers to share data. And as stated in Assumption 0, very few ranchers collect data in the first place#. This signals the need for distinct incentive structures—such as benchmarking tools, operational improvements, or community-led governance models—that are designed around producer priorities and trust.

Assumption 6

Practice Incentives Are Necessary, But Not Sufficient for Systems Change

Practice-level incentives (e.g., payments to ranchers for data collection and sharing) may be effective as entry points for rancher participation but are unlikely to create enduring behavior change. Unless the broader system becomes inherently advantageous for data sharing—through convenience, security, and realizing of value —participation will diminish if/when up-front cash incentives for collecting and sharing data end. Enduring change requires systemic shifts, not just transactional incentives. For example, services on data (as opposed to simple payments for data) may be a necessary longer-term play.

Assumption 7

A Successful System Requires Multiple Interrelated Benefits For Involved Actors

It is insufficient for the benefit of an effort simply to be "increased ease of data sharing" or "increased security of data sharing". This is because the easiest and most secure form of data sharing is likely no sharing at all. Thus, additional benefits must be layered in order to incentivize said sharing in the first place. Furthermore, the focus on multiple use and benefits ensures no single mechanism needs to succeed in order for progress to be achieved. For example, if sharing data helps ranchers to access markets and agronomic insights and ecosystem credits, or if sharing data helps researchers develop models and access high quality data reserves and increase representation and insights.

Assumption 8

Governance (specifically a lack thereof) and Incentives are the Primary Bottlenecks

While investments in technical infrastructure are important for generating value to stakeholders (including producers), the primary constraint to advancing agricultural data sharing lies in governance and generation of clear incentives. Without clear, inclusive, and enforceable governance models, even the best technical infrastructure will fail to create equitable, trusted, adoptable, compelling, or useful systems.

Assumption 9

Policy and Regulation Will Likely Not Evolve Fast Enough to Meet Sector Needs

Given the pace of technological and market developments, policy and regulatory evolution will likely lag behind present needs. Thus, interim governance models must be driven by multi-stakeholder coalitions, voluntary frameworks, and innovative governance experimentation that can operate ahead of formal regulatory shifts.

Assumption 10

Governance Models Should Align with FAIR Principles and Protect Data Sovereignty

Any emerging data sharing and governance systems must embed FAIR principles (Findable, Accessible, Interoperable, Reusable) while simultaneously ensuring data security, producer control, and sovereignty. This balance is essential to building trust and fostering participation across diverse actor groups.

Assumption 11

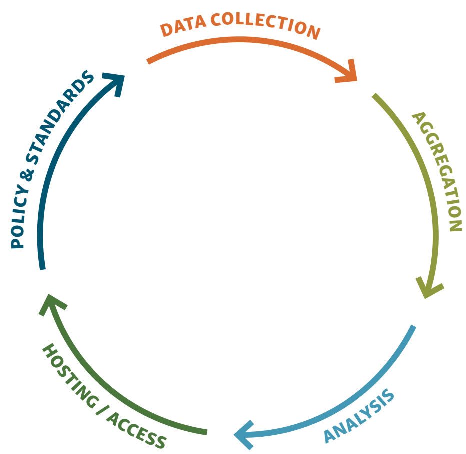
The Sector is Shaped by its Language, Which Must Be Critically Examined

The discourse around data governance is heavily influenced by terms like “data ownership,” which carry colonial and extractive connotations. Exploring alternative framings—such as “data sovereignty” or “data stewardship”—is critical, as the language chosen will shape governance paradigms, stakeholder relations, and perceptions of agency.

Leading Organizations and Mechanisms

A resilient and equitable agricultural data governance ecosystem requires the coordination of five interdependent stakeholder functions: Policy and Standards, Data Collection, Aggregation, Analysis, and Hosting and Access. These categories serve to clarify the operational roles within data ecosystems, and also reveal the governance gaps and structural imbalances that currently limit scalability, interoperability, and trust across sectors.

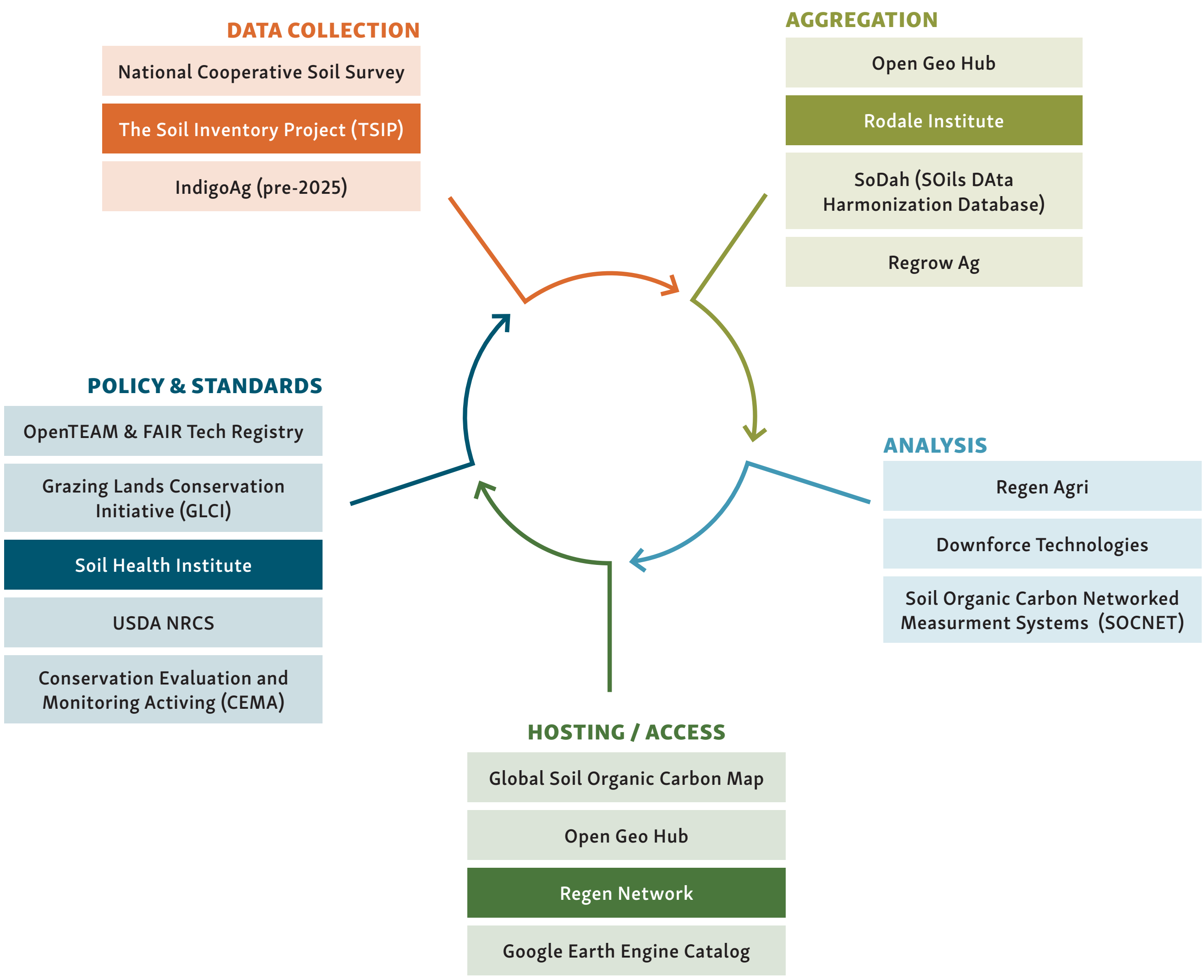
Policy and Standards actors—such as government agencies and ISO-aligned bodies—define the regulatory scaffolding and technical schemas that underpin data legitimacy and interoperability. **Data Collection** stakeholders, including research institutes and producer networks, generate field-level data that form the empirical foundation for downstream claims. **Aggregation** entities work to harmonize fragmented data streams, either via federated models or centralized repositories, providing metadata structure and consent traceability. **Analysis** actors, often academic or technocratic in orientation, apply modeling, benchmarking, and verification logic to these datasets, while **Hosting and Access** actors ensure long-term stewardship, security, and discoverability of data assets. Understanding which organizations currently hold, bridge, or underperform in these roles is critical for designing scalable and equitable governance mechanisms.



In an optimal state. Stakeholders across this cycle of actors are able to find, utilize, and share soil carbon and management data effectively.

Alongside these discrete roles lies a sixth domain: **Interpretation**. Interpretation is conceived of as a key action of each role, and enables them to function effectively in an optimal state: shaping how data outputs are contextualized, communicated, and applied. It encompasses translating technical data into actionable insights, the mediation of meaning between stakeholders with different epistemologies, and the construction of narratives that determine what data counts, for whom, and to what end.

CONTEMPORARY ACTORS WORK IN THE US AGRICULTURAL DATA SPACE



Leading actors across roles in Data Sharing within the context of grazing management and soils data. Organizations highlighted in blue were interviewed to vet research findings and provide additional insights.
TERRA GENESIS 2025

POLICY AND STANDARDS

Soil Health Institute (SHI) - Interviewed

2015 | soilhealthinstitute.org

Role, Value and Potential for Expansion

- U.S.-based non-profit research and policy organization focused on advancing soil health measurement frameworks and decision tools.
- SHI developed the Soil Health Measurement Framework and plays a key role in aligning U.S. institutions (e.g., NRCS, Land Grant Universities, extension services) on standardized soil health indicators.
- SHI is well-positioned to bridge scientific research and implementation pathways for federal soil health programs. It could facilitate consensus-building on minimum viable datasets and data integration protocols into USDA systems.

Cultural, Academic, and Economic Priorities

The Soil Health Institute is anchored in a technocratic and agronomic tradition with strong alignment with a traditional (industrial) agricultural paradigm. While it provides credibility through peer-reviewed research and cross-institutional validation of metrics its primary work is in influencing resource allocation and program design within USDA, making it a high-leverage actor for shaping public investments in soil health and for mediating between policy and practice.

USDA NRCS (Natural Resources Conservation Service)

1935 | nrcs.usda.gov

- Collects and shares U.S. soil carbon data from conservation programs (Soil Carbon Monitoring Network, CEMA, Dynamic Soil Survey, and other shorter-term efforts, such as Partnerships for Climate Smart Commodities Projects.
- Potential expansion of farmer participation and data sharing through incentives.

NRCS's efforts are deeply embedded in the U.S. agricultural cultural landscape, where farmer trust, data sovereignty, and local knowledge remain paramount. NRCS are recognized as datasets that serve as gold standards for soil health research and policy modeling, offering often unparalleled temporal depth and spatial resolution. NRCS's potential to accelerate farmer data contributions—if paired with equitable incentives and privacy guarantees—positions it as a cornerstone actor in unlocking value across soil health markets and regenerative finance initiatives.

FAIR Data Principles (Findable, Accessible, Interoperable, Reusable)

2016 | go-fair.org

- Promotes data-sharing best practices for scientific data.
- Encourages more widespread corporate adoption in soil carbon reporting.

The FAIR principles represent a cultural "north star" for open data stewardship, embedding values of transparency, equity, and interoperability into scientific practice. FAIR principles underpin the reproducibility and scalability of research outputs, especially in cross-disciplinary and multi-sector initiatives. The widespread adoption of FAIR principles in corporate reporting could standardize ESG data disclosures, de-risk data-driven investments, and reduce compliance costs, while also fostering more equitable participation in emerging data economies.

DATA COLLECTION

The Soil Inventory Project (TSIP) - Interviewed

2010 | thesoilinventoryproject.org

- Multi-stakeholder initiative consolidating distributed soil datasets across academic, public, and private sectors into a unified inventory.
- Operates as a metadata aggregation and alignment effort, curating dataset descriptors, data provenance information, and governance characteristics. TSIP supports visibility into fragmented datasets and enables alignment across emerging coalitions, with established tech infrastructure.
- TSIP is poised to serve as a backbone structure for federated soil data networks. It can anchor discussions around data trust architecture, credentialing of datasets, and identification of structural gaps in U.S. and international inventories.

TSIP reflects a collaborative ethos built on transparency and mutual accountability among data custodians by providing a meta-framework for understanding dataset quality, scope, and stewardship, essential for research reproducibility and decision-maker trust. TSIP further enhances the discoverability of data assets that might otherwise remain siloed, reducing redundancy in data collection and enabling strategic investment in high-need regions or themes.

AGGREGATION

Rodale Institute - Interviewed

1947 | rodaleinstitute.org

- Leading research, education, and advocacy institution advancing regenerative organic agriculture, with more than 70 years of field trials and producer engagement.
- Maintains some of the longest-standing comparative trials in organic vs. conventional farming. Disseminates practical guidance, policy recommendations, and standards development for regenerative systems.
- While not a core data infrastructure actor, Rodale serves as a trusted intermediary capable of legitimizing new frameworks for data stewardship, especially among regenerative and organic farmers skeptical of conventional institutional channels.

Rodale embodies a values-driven, farmer-aligned perspective that often resists technocratic or extractive data models. Further, it grounds regenerative agriculture in empirical evidence and supports translational science. Rodale's influence on policy and certification pathways shapes market access and trust for a growing segment of producers who prioritize soil health beyond yield-centric metrics.

SoDaH (SOils DAta Harmonization) Database

First published 2020 | LTER

- A global initiative aimed at harmonizing soil datasets from different research sources.
- Allows researchers and policymakers to integrate and compare standardized soil carbon data across studies.
- Could expand partnerships with private sector data holders.

SoDaH plays a pivotal role in bridging fragmented academic research traditions that often prioritize local or national datasets over globally harmonized standards. By fostering a culture of shared scientific norms, SoDaH enhances cross-regional comparability, supporting not only research integrity but also policy harmonization. The platform's target model is to lower the cost of data discovery and analysis for both public and private sectors, enabling more efficient R&D investments into climate-smart agriculture, soil health valuation, and ecosystem services monetization.

OPENGEOHUB

2019 | opengeohub.org

- Develops open-source geospatial tools for soil data analysis.
- Potential to expand integrated soil carbon models with existing open climate datasets.

OPENGEOHUB fosters a culture of open-source innovation, bridging gaps between academic, civic, and private sector users of geospatial data. Its tools democratize access to high-end analytics, enabling under-resourced actors—including smallholder networks and NGOs—to engage in soil health assessments. Academically, it accelerates methodological experimentation and teaching. From an economic perspective OPENGEOHUB interoperability focus supports efficiencies in data workflows for consulting, certification, and carbon market services.

ANALYSIS

SOCNET (Soil Organic Carbon Networked Measurement System)

2022 | wicst.wisc.edu/socnet

- Developing standardized measurement protocols for soil carbon.
- Aims to unify soil carbon data reporting, making it easier for different stakeholders to share data, in particular between farmers and university researchers.
- Potential for expanding data-sharing beyond research institutions to include corporations.

SOCNET confronts the entrenched cultural challenge of fragmented data collection methodologies by positioning standardization as both a scientific necessity and an economic enabler. It enhances methodological rigor and reproducibility, creating pathways for meta-analyses and cross-site comparisons critical to climate modeling. From an economic perspective, SOCNET's protocols could serve as the scaffolding for harmonized soil carbon credits, providing credible pathways for both voluntary and compliance market mechanisms.

HOSTING AND ACCESS

Regen Network (RND) - Interviewed

2016 | regen.network

- A blockchain-enabled ecological data platform for verifying ecosystem services, particularly soil carbon and biodiversity, with global reach and producer-centric ethos.
- Offers smart-contract-based data verification, ecological claim issuance, and incentive mechanisms through tokenization. Distinctive for integrating cryptographic trust with ecological science.
- Regen Network may serve as a testbed for next-gen data sovereignty models, federated ledgers, and verification logic. It represents one of the most complete examples of interoperable architecture between data collection, governance, and transaction.

Regen Network has strived to exemplify a decentralized, post-institutional governance model that challenges conventional power asymmetries in data ownership by building connections between digital governance research, ecological modeling, and Web3 infrastructure. Regen Network opens pathways for market design experiments—especially around voluntary ecosystem markets—enabling farmers to derive revenue from verified data while retaining agency.

Global Soil Organic Carbon Map (GSOCmap) (FAO)

2017 | gsocmap

- Provides publicly accessible soil organic carbon maps, updated periodically.
- Supports global soil carbon tracking with open data contributions from countries.
- Encouraging private entities to contribute anonymized datasets would improve accuracy.

GSOCmap reflects the FAO's mission to promote data democratization as a public good, embedding cultural values of transparency and equity into global soil health governance. It serves as a foundational reference for modeling, reporting, and verification (MRV) in carbon accounting, supporting interdisciplinary collaborations across agronomy, climate science, and environmental economics. GSOCmap underpins emerging carbon markets by providing essential baseline data, although its expansion into private-sector partnerships would elevate both its precision and its commercial utility.

European Soil Data Centre (ESDAC)

2012 | esdac.jrc.ec.europa.eu

- European Commission's soil data repository, providing public soil carbon datasets.
- Push for more integration with corporate soil carbon data.

ESDAC operates within a European governance context that values data sovereignty and harmonized environmental reporting, aligning with the EU's digital and green policy frameworks. It enables cross-border research and supports directives on soil health, biodiversity, and climate adaptation. Further, ESDAC has the potential to reduce transaction costs for multinational agrifood corporations and financial institutions needing high-confidence datasets for compliance, reporting, and impact assessments.

Australia's National Soil Strategy & CSIRO

2021 | csiro.au and agriculture.gov.au

- Funds research on soil carbon measurement and supports national databases.
- Increase public-private collaborations for data-sharing initiatives.

Australia's National Soil Strategy reflects both a cultural and political recognition of soil as a national asset intertwined with the country's agricultural identity and climate resilience goals. CSIRO's leadership drives innovation in soil carbon methodologies, particularly in arid and variable climates, influencing global best practices. Its primary economic role is expanding public-private collaborations and would position Australia as a data-rich ecosystem for regenerative agriculture investments and nature-based solutions finance.

Interview and Research Findings

SYNTHESIS OF KEY INTERVIEWS

Generally speaking, interviews and subsequent research affirmed influences documented in Assumptions. Further findings from these discussions are documented below.

Interviews with leading practitioners across regenerative agriculture, soil health research, and carbon markets surfaced consistent tensions and emergent opportunities within agricultural data governance. At a high level, interview findings can be framed as two primary tensions:

High degree of rigor an standardization necessary for increasing utility, confidence, and credibility of data ⟷ Overly-strict standards make data sharing and exchange systems prohibively difficult to engage with

Since technology is not the bottleneck efforts should focus on the development of good governance ⟷ Technology serves a critical role by lowering the barrier to data sharing and exchange

Tensions for data governance, as highlighted by key interview findings
TERRA GENESIS 2025

A dominant theme in interviews was the persistent mistrust or fear among farmers and ranchers regarding data sharing, rooted in historical patterns of data misuse, lack of transparency, and concerns about punitive actions from regulatory bodies and/or markets. While data sharing is broadly recognized as essential for advancing climate-smart agriculture and accessing carbon or biodiversity markets, practitioners emphasized that farmer participation remains contingent on clear, credible, and immediate value propositions—including secondary revenue streams, benchmarking tools, and market access. Additional surveys and interviews are required to discern specific motivations driving mistrust among farmers and ranchers, and whether those motivations are proxies.

The interviews revealed a strong consensus around the need for flexible, tiered data standards that enable participation from diverse producers without imposing prohibitive technical or financial burdens. There is widespread concern that overly rigid standards will entrench exclusion, limit scalability, and exacerbate existing inequities within the sector. Conversely, innovative governance structures—such as group-level claims, insurance mechanisms, and standards development boards—are seen as vehicles to balance producer protection with downstream data utility.

Data quality variability emerged as both a technical and governance challenge. Practitioners acknowledged that without contextual metadata and transparency on collection conditions, soil carbon data risks becoming scientifically unreliable and undermining the credibility of carbon markets and certification systems. The complexity of soil science further compounds this risk, necessitating that data systems evolve with both scientific rigor and practical accessibility in mind.

Among all interviewees shared the critique of continuing dominance of privatized data ecosystems, with calls to reframe data governance around open-source models and collaborative coalitions. Interviewees saw coalition building, participatory design processes, and clear articulation of mutual benefits as essential to breaking historical paradigms of extractive economics.

SOIL HEALTH INSTITUTE
Christine Morgan

SOIL HEALTH INVENTORY
Kris Covey

REGEN NETWORK
Gisel Booman

RODALE INSTITUTE
Andrew Smith

FIG.1. Primary interviewees and associated organizations. Refer to “Leading Organizations” for breakdown of role in data sharing and data governance networks.
TERRA GENESIS 2025

KEY INTERVIEW AND RESEARCH FINDINGS

We conducted several interviews with leading practitioners, researchers, and data managers operating in the agricultural sector. In addition, we combined insights from previous aligned projects focused on data governance and the application of governance for managing data infrastructure, data analysis, and data monetization. (Fig. 1.)

These primary viewpoints reflect both alignments and frictions with the working assumption that data governance models in agriculture remain undefined and fragmented, despite clarity on data availability and actor roles:

Finding 1

No Shared Governance Model Exists Across Stakeholders

Multiple stakeholders (interviewees) explicitly noted that while data sharing is acknowledged as necessary, there is no clear, accepted governance structure defining how data should be managed, shared, or protected across actors.

Finding 2

Producers Prioritize Data Sovereignty in Absence of Trusted Frameworks

Each interviewee conveyed the shared sentiments we have heard from producers in the first phase of this scope of work; specifically their desire for control over data use and sharing, but this emphasis stems largely from a lack of trust and the absence of clear governance protocols to guarantee agency or benefit sharing.

Finding 3

Corporate Focus on Interoperability Lacks Counterbalancing Governance Safeguards

Corporate actors expressed the need for scalable and interoperable systems; they did not propose governance models that resolve producer concerns (e.g. data ownership, replication or corruption of data, and data privacy), reinforcing a gap between infrastructure desires and governance commitments.

Finding 4

Researchers Highlight Proprietary Data Blockages Without Clear Governance Resolutions

Researchers voiced frustration over the inaccessibility of private data, but the interviews show no clear pathway being proposed to resolve these tensions through governance innovation, only critique of current barriers.

Finding 5

Data Quality and Consent are Recurring Governance Gaps

Across groups, there is concern that even when data exists, its quality, consent pathways, and provenance are unclear—reflecting governance voids in addition to data scarcity.

Finding 6

Privacy and Trust are Systemic Barriers with No Governance Consensus

Persistent farmer skepticism about data misuse continues to be named as a core challenge, yet no stakeholder group articulated a universally accepted governance mechanism to address this beyond generic calls for transparency and consent.

Finding 7

Practitioners Agree on the Need for Incentives but Not on How to Govern Their Allocation

While there is alignment that data sharing requires farmer incentives (e.g. farmers benefitting from sharing their data specifically), no governance model is yet in place to ensure these are equitable, sustained, or tied to clear participation rules.

Finding 8

Group-Level Claims and Buffer Pools for Risk Mitigation are Mentioned, but Not Institutionalized

Group-level claims and risk pools are possible governance tools to protect producers, yet there are few domestic (U.S.) models that are readily-scalable, and which have data sharing models (turnkey or bespoke) that conform with existing limited data sharing among industry stakeholders.

Finding 9

Ranchers with Documented Management Plans are a Biased Cohort

As documented in a research collaboration between USRSB and Future in Food, only 10% of ranchers have a documented (sharable) grazing management plan. These ranchers skew above average with regard to their willingness to try new technologies. Less than half of these producers have documented their management plans digitally.

OPPORTUNITIES

Opportunity 1

Balance Standardization with Flexibility

There's an opportunity to develop flexible data standards that propose different tiers of rigor/ methodology depending on the data's intended use. These shall be designed in a way that still enables interoperability and improves data confidence. Innovate with Governance: The convening of a Standards Development Board could provide value.

Opportunity 2

Incentivize Farmer Participation

Ultimately, attaching crucial metadata to soil data (such as sample coordinates, management decisions, etc) requires farmer willingness to share data. Secondary revenue streams from data and access to carbon/biodiversity markets are opportunities to increase farmers' willingness to share data. Benchmarking tools could also incentivize farmers by helping them improve their practices and positioning them competitively.

Opportunity 3

Supplemental Opportunity

Creating reciprocal value for farmers through models (Multi-Model Ensembles) that enable farmers to synthesize foundational data on landscape parameters, ecological health, soil composition, management histories, and biophysical patterns at the field-level. MMEs, when available through integrated platforms attuned to farmer and rancher needs, demonstratively exhibit how rancher data contribute to broad ecosystem and industry insights.

CHALLENGES AND RISKS

Risk 1

Privacy and trust continue to be critical barriers.

Farmers fear misuse of their data, and other stakeholders - such as researchers - don't want to share their data without confidence it will contribute credible value to the industry.

Risk 2

Holding the industry to a high scientific rigor can make scaling impractical.

Overly-rigid standards risk excluding farmers at different readiness levels. For example, some stakeholders - such as researchers - may need a level of data quality that is highly rigorous and therefore difficult to scale. Other stakeholders may require less rigor to accomplish their aims, but this approach may also render data quality insufficient for the research community.

Risk 3

Data quality varies generally.

Without standards, data quality varies significantly, undermining downstream credibility for carbon markets, certifications, or scientific analysis.

Risk 4

Detail matters.

The science of soil carbon measurement is quite complicated. Without high detail regarding the conditions that soil data were collected under, the takeaways generated from that data could be inaccurate.

Risk 5

The data sharing ecosystem continues to fall into a historic paradigm of privatization.

Finding ways to encourage collaboration over proprietary approaches would be valuable. Consider, what can data-sharing entities stand to gain by taking an open-sourced approach?

Key Observed Gaps

The persistent data gap in soil carbon monitoring and modeling across U.S. grazing lands represents not only a deficit in information but reflects the structural absence of coherent data governance, interoperability standards, and value-aligned incentive mechanisms. This gap has been consistently reinforced through engagement with producers, researchers, ag-tech firms, and sustainability professionals throughout the GLCDI process (although the participants during the discovery phases emphasized their enthusiasm for data collection, as well as expanding the means to collect data).

Consent and Data Use Rights

While the spatial and quantitative dimensions of the gap—for example, the 612 million acres lacking digital farm management data—present a significant limitation, the qualitative dimensions may be even more critical. Data integrity, provenance, contextual metadata (e.g., soil characteristics, management practices), and clarity on consent and data usage rights remain underdeveloped or absent from existing datasets, particularly those managed by private platforms or decentralized farm record systems. This trend has expanded in the past decade with the ascension of "Ag Tech" proprietary functions intended to meet market needs versus data integrity standards.

Furthermore, the lack of transparency regarding the volume, quality, and consent of privately held data by ag-tech companies exacerbates this data gap (again, in potential conflict with those same actors supporting the expansion of data collection and distribution). This opacity limits researchers' ability to calibrate or validate models such as Century, RothC, or emerging machine learning approaches[1], while also undermining the potential of these models to be applied confidently across diverse grazing landscapes.

The fragmentation of data collection methodologies, ranging from producer-led pen-and-paper records to proprietary software platforms, creates a patchwork of non-interoperable datasets. Compounding this is the absence of an industry-wide consensus on which datasets are foundational for which models or claims, whether those pertain to carbon market participation, biodiversity co-benefits, or ecosystem services reporting. The interviewees confirmed that this ambiguity impedes the sector's ability to move toward scalable and trusted data systems, as stakeholders continue to operate in parallel, disconnected streams of data generation and analysis. This is both in terms of data quality, and the effectiveness with which data is used to solve real-world challenges.

[1] A multi model ensemble reveals net climate benefits from regenerative practices in US Midwest croplands

Data Standards

Efforts to close the data gap cannot proceed without confronting the systemic challenge of establishing widely accepted, flexible, and producer-centered data standards. Interviews and working groups in both the Discovery and Landscape Analysis phases, as well as in other data governance projects, reiterated the need for open-source data schemas and interoperable APIs that align with existing standards such as ISO 28258 (SoilML), OpenAPI protocols for soil health data, and NRCS data systems, while still enabling the nuanced, localized data collection practices necessary for land stewards and smallholder producers. (Fig. 1.)

The development of such standards cannot be isolated from broader governance considerations, as stakeholders emphasized that standards function not only as technical tools but as mechanisms of inclusion, power, and access. This is especially important outside a domestic (U.S.) context where smallholder farmers and producer communities are at greater odds with how their data is collected and used, and in many cases without their input or consent. Without participatory, transparent standard-setting processes, the risk persists that standards will favor the data requirements and technical capacities of downstream actors (e.g., corporations, certifiers) while marginalizing producers—especially those operating without advanced data infrastructure, data literacy, or in marginalized regions more generally.

Interoperability must be understood as both a technical and political challenge, requiring that emerging standards are designed to bridge existing data ecosystems (e.g., PODS, Yardstick, OpenTEAM platforms) while also integrating legacy systems used by USDA, NRCS, and other public agencies. This will require not only technical

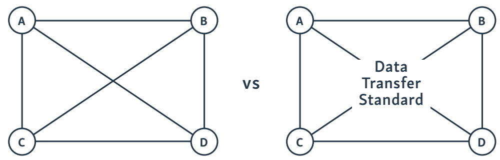


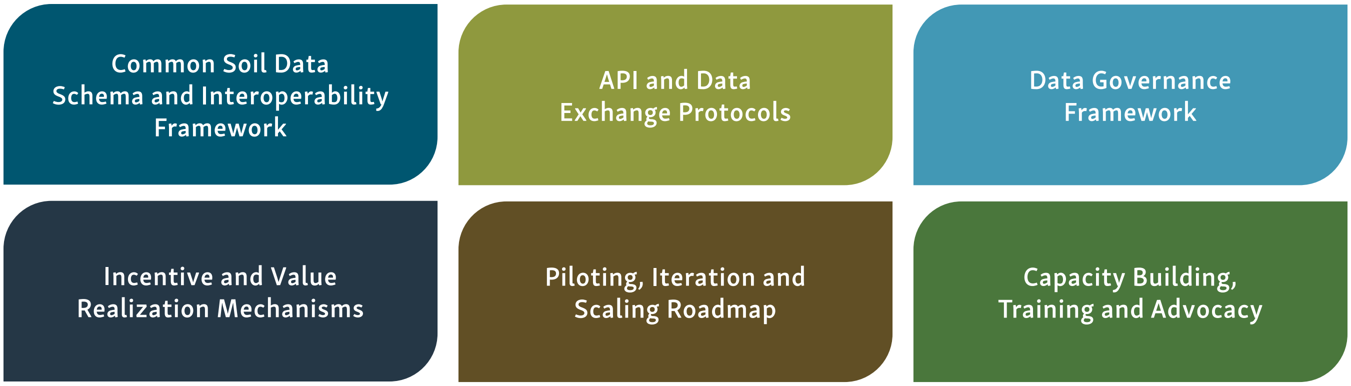
FIG.1. Concept: A data transfer standard would streamline sharing
TERRA GENESIS 2025

harmonization of data formats (e.g., JSON, XML) but also alignment of governance principles with global norms such as the FAIR data principles (Findable, Accessible, Interoperable, Reusable), ensuring that data remains open, traceable, and usable across sectors and geographies.

Standards Development

Standards development must anticipate future scalability and the integration of emergent data types and verification technologies, such as remote sensing, blockchain-based verification, or producer-managed data trusts. Modular, flexible standards architectures—allowing for tiered participation levels and adaptive data schemas—will be essential to ensuring that data systems can evolve alongside scientific, technological, and policy developments, rather than becoming static or exclusionary.

It is imperative that any standard effort be embedded within an enforceable data governance framework, ensuring clear protocols for data consent, usage rights, attribution, and benefit-sharing mechanisms, particularly for producers and Indigenous communities. Without such governance alignment, the technical advancement of data standards will fail to resolve the underlying tensions of data sovereignty, equity, and trust that continue to define the agricultural data landscape.



Core Components of Ag Data Governance and Sharing Infrastructure
TERRA GENESIS 2025

CORE COMPONENTS OF AG DATA GOVERNANCE AND SHARING INFRASTRUCTURE

Common Soil Data Schema and Interoperability Framework

Objective

Establish a universally accepted, open-source soil data schema, aligned with existing standards (e.g., ISO 28258 SoilML, OpenAPI protocols), to enable interoperability across platforms (USDA, NRCS, FAO, PODS, Yardstick, Pasture Map, OpenTEAM).

Supporting Context

The current landscape is fractured by proprietary data models, incompatible systems, and inconsistent formats. A harmonized schema is foundational to bridge these divides, ensuring datasets from diverse sources—including producers, corporations, and research bodies—can be aggregated, validated, and utilized without costly conversion or quality degradation.

Critical Features

- Modular design to support evolving data types (soil health, biodiversity, management practices).
- Tiered schema allowing both minimal data contributors and high-resolution data systems to coexist.
- Metadata standards to ensure traceability, consent documentation, and provenance tracking.
- Outcomes visualized for producers and based on model inputs (implications of management).

API and Data Exchange Protocols

Objective

Define secure, standardized API protocols for seamless data exchange between platforms, stakeholders, and systems.

Supporting Context

While some systems offer APIs, these are not harmonized, limiting cross-system functionality and compounding governance and data integration challenges. OpenAPI for soil health data, aligned with broader OGC standards, will enable machine-readable, automated data exchanges while embedding necessary data consent and privacy flags.

Critical Features

- Open-source, documented API specifications.
- Built-in governance hooks for consent management, user attribution, and data access restrictions.
- Compliance compatibility with FAIR principles and institutional mandates (USDA, NRCS).

Data Governance Framework

Objective

Co-develop a participatory, enforceable data governance framework that defines data ownership, consent, access, benefit-sharing, and dispute resolution mechanisms.

Supporting Context

The absence of governance alignment remains the primary bottleneck to scaling data sharing. This framework must not only codify rights and responsibilities but also protect data sovereignty for producers, Indigenous communities, and marginalized groups. The framework should align with:

- Global norms (e.g., FAIR, CARE principles).
- Producer-centered data stewardship models.
- Emerging practices from data trusts and cooperative governance experiments.

Critical Features

- Transparent protocols for data contribution, usage rights, and attribution.
- Governance participation pathways for producers and community-led entities.
- Conflict resolution and grievance mechanisms.

Incentive and Value Realization Mechanisms

Objective

Design incentives that drive both private and public sector engagement while centering producer value.

Supporting Context

As the synthesis has shown, practice-level incentives alone are insufficient for long-term behavior change. Systems-level value realization must be built into the architecture of data sharing agreements and governance frameworks.

Critical Features

- Tiered participation benefits (e.g., benchmarking tools, preferential access to carbon markets).
- Data-as-a-service models with clear revenue pathways or market access improvements for data contributors.
- Mechanisms ensuring that contributors—especially farmers—share in the downstream value created by their data.

Piloting, Iteration, and Scaling Roadmap

Objective

Test governance, incentive mechanisms, technical components, and value realization through pilots, iteratively refine, and scale via global alignment and certification processes.

Supporting Context

Following the proposed roadmap, pilots with OpenTEAM, USDA, and FAO would stress-test both technical interoperability and governance frameworks under real-world conditions. These pilots would also serve to explore critical gaps in consent enforcement, cross-sector data flows, and stakeholder trust-building.

Critical Features

- Pilots scoped for diverse geographies and stakeholder constellations.
- Evaluation of technical, social, and governance KPIs.
- Post-pilot alignment with ISO, OGC standards, and advocacy for institutional adoption.

Capacity Building, Training, and Advocacy

Objective

Equip stakeholders—particularly producers and data stewards—with the skills, tools, and knowledge to engage effectively in emerging data ecosystems.

Supporting Context

Education and trust-building were consistently cited as critical prerequisites for farmer participation. Without targeted capacity-building, new systems risk replicating existing inequities and failing to engage the very actors whose data is most essential.

Critical Features

- Practitioner-focused toolkits and training on data standards, rights, and benefits.
- Producer and community-led data literacy programs.
- Global advocacy for institutional adoption of the governance and technical frameworks.

Opportunities

INTEGRATING CASE STUDIES, ASSUMPTIONS, AND PRACTITIONER INSIGHTS FOR AGRICULTURAL DATA GOVERNANCE MODELS

An examination of meteorological data governance, open-source software ecosystems, and creative commons frameworks—augmented by insights from practitioners in agricultural data collection systems and research—illustrates the complexity of designing equitable, resilient, and adaptive governance architectures in the agricultural sector. Across all examples there exists a constant: the interplay between infrastructure sufficiency, governance modalities, and stakeholder incentives is central to the evolution of sustainable data ecosystems.

Practitioner insights further reinforce that infrastructure sufficiency alone is insufficient to address these systemic tensions. That is to say, while we know the capacity to collect, exchange, and examine data within the sector can be accomplished, the underlying intentions of individual stakeholders must be reconciled (i.e. incentives alignment). In particular, challenges around data quality, variable standards, and the complexity of soil carbon science demand that any governance model balances scientific rigor with accessibility and farmer inclusion. The historical paradigm of data privatization persists, and unless actively countered through innovative governance models—such as standards development boards, federated commons models, or modular governance architectures—agricultural data ecosystems risk replicating or trending toward extractive structures familiar from other sectors.

The work on GLCDI to-date converges around the need for tiered, flexible governance structures (including novel structures) that allow differentiated participation, contribution, and benefit capture across diverse stakeholder groups and readiness levels. This includes incentive alignment mechanisms such as secondary revenue streams (e.g., credits, benchmarking, and MRV monetization) that position data contributors as co-owners of value chains, rather than passive data providers. These models must be explicitly designed to navigate the frictions between standardization and contextual specificity, enabling interoperability without excluding stakeholders unable to meet rigid data collection or formatting standards.

The emerging synthesis points toward the necessity of, at minimum, a hybrid, federated governance models for agricultural data systems, blending public custodianship of core infrastructures, peer-driven commons stewardship for adaptive governance, and dynamic consent and incentive architectures that secure producer participation, trust, and benefit. This approach must be accompanied by deliberate, early-stage design decisions around interoperability, licensing, contributor recognition, and governance authority that reflect both the risks and opportunities of the sector's data ecosystem evolution.

OPPORTUNITIES FOR EXPANDING DATA SHARING IN AGRICULTURE

Opportunity 1

Establishing a Public-Private Data Sharing Framework

Drawing lessons from open weather data models, a shared soil carbon data repository could be established through public-private cooperation. Government agencies, private sector actors, and non-profits would contribute toward policies supporting the anonymization of personally-identifying data, as well as creating a common infrastructure for scientific, regulatory, and commercial applications. This could address current fragmentation by building a neutral, trust-based platform modeled after precedents like NOAA data partnerships.

Opportunity 2

Incentivizing Private Sector Contributions

Public and corporate incentives remain essential to stimulate participation. Governments could subsidize data submission by farmers and ranchers, while food and agriculture companies might receive sustainability certifications or preferential procurement treatment for transparently disclosing supply chain soil carbon data. However, incentives must be designed to overcome persistent private-sector hesitation rooted in data control, competitive advantage, and liability concerns. Examples like zero-knowledge (ZK) technologies and blind-trust models offer technical capabilities suited for addressing privacy concerns specifically–an issue cited frequently from farmers and ranchers.

Opportunity 3

Creation of a Global Open Soil Carbon Database

Inspired by global climate modeling and initiatives such as the FAO's GSOCmap, a non-profit-led global soil carbon database could provide universal access to harmonized data. Unlike current efforts, this database would explicitly integrate private-sector data through governance models that protect proprietary interests while ensuring data accessibility for public benefit. This model would demand carefully crafted data-sharing agreements, anonymization protocols, and tiered access frameworks.

Opportunity 4

Open Data Alliance Model for Participatory Governance

A federated alliance model could operationalize the above efforts, requiring participants to contribute data in order to gain access to the collective pool—echoing successful open-source governance models such as Linux. While access to data remains non-monetary, major beneficiaries (e.g., large agrifood or tech firms) would be expected to underwrite operational costs through voluntary contributions, ensuring long-term sustainability of the commons without commoditizing data access itself.

PROPOSED ROADMAP TOWARD DATA GOVERNANCE AND SHARING ALIGNMENT

As is outlined in "Next Steps" this landscape analysis does not aim to define a prescriptive approach to advancing the industry. Thus, the roadmap below describes a potential order of operations inclusive of recommendations from this landscape assessment.

PHASE 1	PHASE 2
Landscape Assessment and Stakeholder Engagement (present scope)	Standard Development and Prototyping
PHASE 3	**PHASE 4**
Piloting Governance Framework and Technical Interoperability	Scaling, Certification and Advocacy

Phase 1 - Landscape Assessment and Stakeholder Engagement (Present Scope)

- Conduct comparative gap analysis of existing data models and corresponding business models (e.g., NRCS, WoSIS, PODS) to determine efficacy of data management and financial viability of the supporting enterprise
- Identify and convene key stakeholders (forming a 'Trust Alliance')—including producers, scientists, agronomists, policymakers, private sector, and ag-tech developers.
- Establish multi-stakeholder working groups to surface barriers, opportunities, and alignment needs.

Phase 2 - Standard Development and Prototyping

- Co-develop industry-wide standards for soils data, including collection methodologies and storage formats
- Such a standard may take a "tiered" approach, where lower tiers of the standard emphasize accessibility to enable scale and the higher tiers of the standard emphasize rigor for precise, scientific insights
- A standard may be designed in open-source formats and aligned with existing standards like ISO 28258 (SoilML) and FAIR principles.
- Define API protocols for soil health data exchange using OpenAPI specifications.
- Pilot integration of these standards within selected data systems (e.g. NRCS, TSIP, Yardstick, OpenTEAM).

Phase 3 - Piloting Governance Frameworks and Technical Interoperability

- Implement pilot projects with key partners to test interoperability, governance protocols,incentive alignment, value creation, and stakeholder compliance.
- Formalize a data governance framework ensuring transparency, producer agency, ethical data sharing, and equitable cost-sharing mechanisms.
- Address complex issues of consent, privacy, and data quality through iterative feedback and adaptive governance models.

Phase 4 - Scaling, Certification, and Advocacy

- Pursue ISO (International Organization for Standardization) and OGC (Open Geospatial Consortium) certification of the developed standards and governance models.
- Develop practitioner-facing resources, training, and documentation to enable broad adoption.
- Launch coordinated advocacy campaigns targeting USDA, FAO, NRCS, and global scientific bodies to embed the new governance frameworks into institutional and policy contexts.

Next Steps

Rather than prescribing pathways forward based on this research alone, it is encouraged that the Steering Committee of the GLCDI use this document and their collective opinions and experiences to discern and co-create next steps. This section, therefore, identifies useful activities to reflect on content and describe aspects of the roadmap ahead.

CONSIDER THE FOLLOWING PROMPTS

1. After reviewing the "Assumptions" section:

a. Which statements feel the most pressing to you?

b. Are there any you would like to challenge?

2. When reviewing the "Opportunities"" section:

a. What can you identify as challenges and opportunities within each proposed pathway?

b. What might help to reconcile these tensions so that progress is possible?

3. Much of the recommendations within this report point towards the development of a flexible yet credible data standard.

a. Does this recommendation concern you in any way?

b. What would the presence of a *flexible yet credible* standard make possible?

   i. In addition to the terms "flexible yet credible" above, what other adjectives would you use to describe an ideal data governance standard?

c. What stakeholders are needed to advance the development of this standard?

4. Regarding other gaps and opportunities:

a. What key value is currently not being realized in the industry? What would unlock that value, and who must be engaged or introduced to this industry to realize that value?

5. After reviewing the "Example Models" sections in the Appendix:

a. What principles and practices would you highlight as useful for advancing this industry?

b. In which direction does the agricultural data industry need to head in order to reach a state where public and private development allows for innovation and market opportunities like those explored in the case studies? What does this "state of the art" look like when it is fully realized?

6. Regarding the composition of the steering committee (and its shared aims):

a. Can the steering committee define a specific goal, or series of goals, that must be reached in order to realize the preferred state of affairs with respect to agricultural data sharing and monetization?

7. In order to move forward:

a. Who is most capable of setting the conditions that are necessary for this industry to build off of? What purpose are they serving? *(Is it private entities (e.g. corporations) with their broad access to capital, public entities (e.g. government) with strong ethical and science-backed foundations, or public stakeholders and networks (e.g. participatory communities) who push innovation in unforeseen ways?)*

b. What instruments, specifically, need to be employed to make the proposed models work, and what instruments, specifically, are not sufficient? Where are their barriers to access and what would it take to remove those barriers.

Appendix A: Case Study Synthesis

IMPLICATIONS FOR AGRICULTURAL DATA SHARING GOVERNANCE MODELS

Each case study—meteorological data governance, open-source software architectures, and Creative Commons digital image rights—offers distinct yet interlinked perspectives on the governance, valuation, and stewardship of data within hybrid public-private ecosystems. For each case study the stakeholders involved, or the market, obviates gaps between providers and users. These nominally represent possible outcomes for agricultural data governance within the GLCDI, although the ubiquity and maturity of the agricultural sector functionally exceeds that of any one case study. Each illustrates critical lessons for the agricultural data domain, particularly where foundational infrastructures are fragmented, governance mechanisms are emergent, and value flows remain asymmetrically distributed–all concerns raised by participating members of the GLCDI, as well as those interviewed.

METEOROLOGICAL DATA: PUBLIC-PRIVATE STEWARDSHIP

Meteorological data governance underscores the centrality of foundational public infrastructure as a precondition for equitable data access and utility. The public sector's enduring custodianship of collection integrity, quality assurance, and long-term stewardship via Federal entities (acting fundamentally in the interest of the public) provides the backbone upon which private sector actors build specialized, often high-value, services.

However, the shift toward commercial-first access models and third-party hosting platforms has diluted the public sector's influence over data dissemination, raising equity and sovereignty concerns. Who has first access to the data? What data may be restricted or paywalled? What ethical concerns do these changes raise? And who, ultimately, governs these changes?

This model emphasizes the necessity of clearly articulated governance boundaries, ensuring that while private actors may optimize and add value to data, the integrity, accessibility, and accountability of primary data systems should remain a public mandate. Moreover, the evolution and purpose of the evolving relationship between the public and private sectors should be defined around whether commercial, academic, or quasi-commercial endeavors underpin the cooperative relationship.

Key Takeaways

1. Public (federal) collection, maintenance, and verification of a high-value data resource
2. Systems and function improvement via efforts to expand/simplify data acquisition and access
3. Subsequent proliferation of commercial viability through reduced friction for data acquisition

Addressing Gaps Relevant to Agricultural Data

1. Explore formalization of public mandates for baseline agricultural data collection and verification, ensuring primary datasets remain accessible and uncompromised by commercial gating.
2. Consider implementation of tiered data access protocols that preserve public-good use while enabling value-added services by private actors under transparent licensing.
3. Clearly define interoperable governance boundaries between federal custodians (e.g., USDA, NRCS) and private service providers to avoid erosion of public control over data infrastructure.

OPEN SOURCE SOFTWARE: RESEARCH FOUNDATION TO COMMERCIAL REVOLUTION

The OSS ecosystem demonstrates a decentralized, peer-driven governance model, where stewardship is embedded within community norms and contributor dynamics. While this has fostered extraordinary innovation and platform ubiquity, it also exposes structural vulnerabilities, including underinvestment in maintenance and the risk of de-commoning–the process by which mutual agreements to uphold the original premise and purpose of the resource erode as competitive framing of its value emerger–as private actors commercialize derivatives without proportional reinvestment.

For agricultural data systems, this case illustrates the importance of proactive governance scaffolds that balance contributor freedoms with safeguards against exploitative appropriation, ensuring that the commons is both resilient and adaptive to emergent user needs. This is especially true as data originators, users, providers, stewards, and fiduciaries encompass a broad set of stakeholders, including farmers and producer communities who historically are among the most exploited and resistant to data sharing.

Key Takeaways

1. Academic sourcing of practical framework to underpin research and economic applications
2. Emergent peer-to-peer governance models with origins in "benevolence" albeit murky intentions
3. Extremely high proportional investment output from foundational resource base

Addressing Gaps Relevant to Agricultural Data

1. Consider the introduction of contributor governance charters or data stewardship agreements to preempt extractive appropriation of community-generated datasets.
2. Establish minimum reinvestment thresholds or benefit-sharing models for commercial users of open agricultural data platforms.
3. Anchor early-stage agricultural data commons in purpose-built institutional scaffolding (e.g., open data cooperatives or federated trusts) to sustain operational viability and preserve alignment with the interests of producers and the public.

CREATIVE COMMONS: AT RISK OF ERODING CONFIDENCE AND EXTRACTION

The Creative Commons case highlights the limits of consent-based, minimally governed models, where the permissiveness of licensing structures has enabled widespread redistribution and derivative use but has done little to ensure creator agency or equitable value capture. In agricultural data contexts, this warns against over-reliance on voluntary participation models without enforceable governance mechanisms, particularly where asymmetries in technical capacity, bargaining power, and data literacy exist (crucially: between farmers and producer communities and agri-businesses). This disparity signals the need for dynamic, enforceable consent architectures that evolve with technological and market developments, inhibiting extractive dynamics wherein producers' data is commodified without reciprocal benefits.

Key Takeaways

1. Minimal governance subject to intensive interference from commercial model evolution
2. Governance is strongly, if not entirely, subject to technological interference
3. Capacity of the creative class (those whose value increase) as centered value producers

Addressing Gaps Relevant to Agricultural Data

1. Move beyond static licensing toward dynamic, machine-readable consent frameworks that allow data contributors—especially farmers—to retain meaningful control over data reuse.
2. Embed community-defined governance principles into platform architecture to prevent consent erosion and strengthen equitable participation.
3. Create traceability protocols that link downstream use of shared data to upstream contributors, enabling transparency, attribution, and potential revenue-sharing pathways.

Synthesis

Synthesizing these lessons, the implications for agricultural data governance are clear: robust, transparent, and enforceable governance mechanisms must be embedded from the outset, whether through centralized public stewardship, federated commons models, or hybrid frameworks that balance openness with protection. Furthermore, governance must be adaptive, recognizing that data infrastructures are not static but evolve alongside technological, regulatory, and market shifts. This requires deliberate early-stage architectural decisions—around interoperability, licensing, data sovereignty, and contributor governance—that will shape long-term equity, access, and value realization pathways.

Crucially, all three cases highlight the interplay between infrastructure sufficiency and governance efficacy. Where infrastructure is robust and well-maintained (as in meteorology), governance can focus on ensuring equitable access and preventing monopolistic capture. Where infrastructure is decentralized and emergent (as in OSS and Creative Commons), governance must work harder to prevent degradation, exploitation, and inequity. For the agricultural sector, which currently lacks cohesive infrastructure, this demands a dual investment in both infrastructural capacity and governance frameworks, ensuring that as data systems scale, they do so in ways that center producer agency, data sovereignty, and shared value creation.

ADDITIONAL MODELS FOR INVESTIGATION

In addition to meteorological data, open-source software, and Creative Commons licensing, the following industry analogs support additional investigation for determining applicable data governance and sharing models.

1. **Health data Systems** offer a critical analog—particularly in the use of anonymized data sharing models (ZK and blind-trust) designed to protect individual privacy specifically, while enabling broad-scale research and policy insights. Applied industry privacy frameworks–HIPAA in the U.S. and GDPR in the EU–demonstrate how robust governance, consent architectures, and de-identification techniques can support ethical data ecosystems across diverse stakeholders.

2. **Financial credit reporting systems** provide a precedent for structured, regulated data aggregation with tiered access and standardized reporting, albeit with cautionary lessons around surveillance and inequity. However, financial credit reporting systems also bear the brunt of negative public perception, driven in part by failing security infrastructure which has resulted in major privacy breaches in recent years.

3. **Satellite imagery platforms**, where hybrid governance models (e.g., NASA-private partnerships) enable free, high-resolution data access globally, catalyzing innovation while safeguarding core infrastructure via public mandate. Each of these sectors illustrates distinct mechanisms for aligning data stewardship, access, and public value—offering applicable lessons for agricultural data governance architectures. The extent to which privacy is guaranteed in satellite imagery, however, is unevenly applied.

Appendix B: Case Studies Detailed

INDUSTRY ANALOGS FOR AGRICULTURAL DATA SHARING GOVERNANCE MODELS

METEOROLOGICAL DATA SHARING, GOVERNANCE AND DEVELOPMENT

The governance of meteorological data represents a unique and instructive model in understanding the complexities of data generation, management, and monetization within a dual public-private framework. Rooted in the premise that meteorological data serves a critical public safety function, the United States federal system—through agencies such as NOAA and sub-agencies like NCEI—has historically maintained the position that all data generated with public funds should be accessible without restriction.[2] This orientation has established an infrastructure that privileges data as a public good, inherently open, optimized, and stewarded for the collective benefit. However, the overlay of private sector interests has introduced nuanced frictions, particularly in how this data is accessed, processed, and monetized by commercial actors who do not bear the cost of collection but stand to generate significant economic value from its use.

A defining characteristic of the meteorological data ecosystem is the inverse governance model that has emerged over the past two decades. Public agencies retain responsibility for data collection integrity, quality assurance, and long-term stewardship, while private entities increasingly assume roles in data refinement, sector-specific application development, and the creation of proprietary models and services layered atop publicly sourced data. This bifurcation of roles is both practical and fraught. It has permitted accelerated innovation in forecasting technologies and risk analytics but has also precipitated a de facto privatization of access pathways, as evidenced by the prioritization of commercial cloud platforms like AWS over public portals such as NCEI for data dissemination.[3]

This shift toward commercially-mediated access illustrates the tensions between public mission accountability and the market-driven imperatives of private actors. While the federal infrastructure remains accountable to legislative oversight and largely insulated from direct economic returns, private sector actors, including forecasting services and insurance industries, capitalize on data streams initially generated for public safety purposes. The inherent risk-attending nature of federal agencies allows them to subsume major initial investment and research in areas otherwise out-of-scope of private capital; moreover this first-mover position enables those agencies and organizations to establish premises of good governance and management.[4,5] The

[2] Private companies couldn't match NOAA's vast public weather data powers
[3] Unlocking the Potential of NEXRAD Data through NOAA's Big Data Partnership
[4] Private companies couldn't match NOAA's vast public weather data powers
[5] Public-Private Partnership: NOAA Taps Commercial Sources for Improved Weather Predictions

establishment of initiatives like the Big Data Partnership (BDP) reflects this dynamic, where cost underwritings for commercial cloud hosting were justified by the anticipated public and private utility of more widely available data, despite minimal direct fiscal benefit to federal agencies.

At a structural level, meteorological data governance is underpinned by the application of FAIR principles, ensuring that data is findable, accessible, interoperable, and reusable across sectors and use cases. However, the embedding of these principles within a system that permits commercial-first access points introduces questions of equity, particularly for smaller-scale users or entities lacking the capacity to engage with privatized data infrastructures. Moreover, the expansion of third-party data modeling and forecasting services further detaches accountability for data accuracy and usage from its primary sources, creating layers of abstraction wherein the risks associated with misinterpretation or misuse of data are increasingly borne by end-users or intermediaries, rather than the original data stewards.

The operational dynamics of meteorological data governance demonstrate a critical differentiator from other sectors, notably agriculture, wherein private entities are both data producers and beneficiaries, yet lack comparable public infrastructure underwriting the costs of collection and stewardship. This absence of foundational public infrastructure in agriculture exacerbates concerns around data accessibility, equity, and control, particularly where smallholders and marginalized communities are concerned. More importantly, the threat of data censorship and erasure by ideological or political means undermines general confidence in the traditionally-stewarded body of data and subsequent data value generation.[6] By contrast, the meteorological sector benefits from a century-long investment in open, government-managed and standardized data systems, enabling the layering of private sector innovation atop a resilient public commons.

This case underscores the potential and limitations of quasi-open governance models, wherein private actors contribute to, but do not fundamentally govern, core data infrastructures. Despite the democratization of access achieved through platforms like AWS, the ownership of processing architectures and commercial derivatives remains firmly within private domains, challenging notions of data as a commons. This dualism raises critical inquiries into whether similar structures could be feasibly replicated in agricultural data ecosystems or whether the inherent differences in data origination, sectoral incentives, and infrastructure investments preclude such analogies from holding.

OPEN SOURCE SOFTWARE, UNIX, AND THE EVOLUTION OF DATA GOVERNANCE MODELS

The development of UNIX and its derivative open-source software (OSS) platforms, most notably Linux, offers a complex analog for examining the interplay of public and private interests in data governance and infrastructure stewardship.[7,8] Originating within academic and

[6] Why the Environmental Data & Governance Initiative is archiving public environmental data
[7] The Evolution and Impact of Open Source Software (OSS) on Modern Technology
[8] The history behind Christine Peterson's term 'open source software'

research institutions, UNIX was designed not as a commercial product, but as a functional response to emergent computing needs, with its evolution underpinned by the four freedoms of software use and distribution. This foundation set the stage for what can be regarded as one of the most successful commons-based technological ecosystems of the modern era, with estimates suggesting that OSS underpins as much as 95-99% of all contemporary codebases.[9,10]

Unlike meteorological data, where federal entities retain custodianship, the governance of OSS platforms is decentralized, emergent, and ideologically motivated. The architecture and base code of UNIX—and later Linux—enabled adaptive use cases to proliferate, creating pathways for both public utility and private monetization without centralized oversight. Governance is embedded not in regulatory frameworks, but in peer-produced, community-enforced norms, with contributors often wielding disproportionate influence relative to their technical contributions or institutional affiliations. This has produced a governance model characterized by benevolent maintenance, where stewardship is voluntary, and where the boundaries between governance and participation are deliberately porous.[9]

Despite its framing as a commons, OSS development is not without its tensions. The inherent openness of the code creates vulnerabilities to what could be termed de-commoning effects, wherein private entities leverage open architectures to create proprietary, closed applications or services, contributing minimally, if at all, back to the core codebase. This dynamic reflects a tragedy of the commons risk within digital infrastructures, where the depletion is not of the resource itself, but of the motivations, incentives, and capacities required to sustain collective maintenance and innovation.[11,12] Private sector actors, while benefitting enormously from the OSS ecosystem, often exacerbate this depletion by prioritizing proprietary extensions and vertical integrations, sidestepping the communal responsibilities embedded in the original governance ethos.[13]

In the context of data governance, the OSS ecosystem provides a critical perspective on user-contributor dynamics, where the intent of contributors directly shapes governance structures, often in the absence of formal mechanisms. This is particularly relevant when considering how agricultural data systems might evolve—where producers, if given agency and infrastructural support, could emerge as informal governance leaders by virtue of their contributions. However, OSS also illustrates the limits of such emergent models, particularly in sectors where contributors are asymmetrically positioned in terms of technical capacity, economic leverage, or institutional support.

Importantly, OSS governance remains contested terrain, with philosophical divisions between those advocating for rigid licensing protections (ensuring all derivatives remain open) and those favoring permissive licensing models that allow private appropriation without mandatory reciprocity.[14] This division underscores a critical

[9] Measuring the Economic Value of Open Source
[10] Measuring the Economic Value of Open Source
[11] Revealing Value: The Economic Power of Open Source Software
[12] The Value of Open Source Software
[13] A Guide to the 6 Open Source Governance Models
[14] Creative Commons and Culture

governance fault line relevant to agricultural data governance debates, particularly in framing whether future systems should privilege openness and reciprocity by design or accept hybrid models that permit selective proprietary use cases.

The OSS case also raises pivotal questions regarding the implicit governance embedded within foundational architectures themselves. By establishing the baseline affordances and constraints of a platform, base code exerts an enduring influence over how future systems, applications, and monetization pathways evolve. In agriculture, where data infrastructures are nascent, this suggests that early-stage design choices—around interoperability, licensing, and contributor governance—may have outsized influence on long-term data sovereignty and economic equity outcomes, reinforcing the need for deliberate and inclusive architectural choices from the outset.

CREATIVE COMMONS & DIGITAL IMAGE RIGHTS IN ERA OF SYNTHETIC REPRODUCTION

The evolution of digital image rights under the framework of Creative Commons licensing provides a provocative case study in the governance of creative data assets and the shifting dynamics of value retention and loss as works transition from proprietary to public spheres. At its core, the Creative Commons (CC) model seeks to democratize access to creative works through consent-based licensing structures, wherein authors explicitly define the terms under which others may use, modify, or distribute their works. This model reflects an incentive-driven commons, wherein creators voluntarily cede certain exclusive rights in exchange for broader dissemination, visibility, and, theoretically, secondary value capture through derivative use cases. [15]

However, the mechanisms of governance within the CC framework expose significant tensions that challenge its efficacy as a sustainable model for creator equity. The open, permissive nature of CC licenses has led to an erosion of the primary value of original works, particularly as they are reposted, remixed, and recontextualized across digital platforms with little attribution or compensation flowing back to original creators. The proliferation of synthetic reproduction technologies, specifically generative AI (GAI), exacerbates this dynamic, enabling the rapid generation of derivative works that may displace or devalue the original asset in both creative and economic terms.[16] The focus on GAI in this case study is to root the evolution of creative commons licenses in a tangible, present example; however, the issue of GAI as diluting or negatively impacting a creative field is far more pervasive than can be covered in this report.

This dynamic reflects a critical differentiator from sectors like meteorology or OSS, where data integrity, provenance, and utility are structurally anchored to original sources. In the creative sector, the detachment of the work from its creator is both technically easy and culturally normalized, particularly in digital ecosystems optimized for frictionless content sharing. Governance structures, while nominally embedded in licensing protocols, lack robust enforcement mechanisms, leaving creators reliant on platform goodwill or litigation to

[15] Preventing Others from Commercializing Your Innovation: Evidence from Creative Commons Licenses

[16] Governance of open source software: state of the art

protect their rights—an often inaccessible avenue for smaller or independent creators.[17]

Furthermore, the governance of creative commons is inherently non-centralized and non-adaptive, lacking institutional stewardship bodies capable of responding to emergent technological disruptions such as GAI. This has precipitated debates around the adequacy of existing licensing models to account for synthetic derivative works, algorithmic reproduction, and the commodification of creative aesthetics divorced from original creators' intent or participation. Such gaps in governance not only diminish creator agency but also challenge the sustainability of the commons model itself, as creators perceive the risk-reward calculus to be increasingly unfavorable.

Of particular relevance to agricultural data governance is the parallel in consent-based participation models, where producers may be invited to contribute data under frameworks that promise democratization but deliver limited direct benefits or protections. The creative commons case warns of the dangers of over-relying on voluntaristic data contributions without corresponding governance safeguards, enforcement mechanisms, and equitable value return structures. It highlights the imperative for consent to be informed, dynamic, and enforceable, rather than static and permissive by default.

The creative sector exemplifies the risks of market-led devaluation of communal resources, wherein the ubiquity of freely available works paradoxically erodes the economic viability of the very creators that fuel the commons. In agricultural data contexts, this underscores the need to preemptively design governance structures that balance openness with mechanisms to prevent data extractivism, ensuring that producers are not simply providers of raw data for others to commercialize, but beneficiaries of the ecosystems they help create.

[17] Time to change the copyright creative commons licensing terms?