Community Data Governance Stack (CDGS): Engineering Community Co-Governance for Data-Intensive Interactive AI Systems

ANONYMOUS AUTHOR(S)

Artificial intelligence (AI) governance remains largely structured around risk mitigation and transparency rather than redistribution and repair. These approaches stabilize extractive data relations by preserving state and corporate control, leaving affected communities with limited recourse. We propose the Community Data Governance Stack (CDGS), a conceptual and infrastructural framework that treats governance as a first-class system design problem: specifying roles, rights, processes, and enforceable instruments through which communities can exercise meaningful control over AI data practices. Drawing on Indigenous Data Sovereignty, design justice, and labor scholarship, CDGS integrates sovereignty, refusal, fiduciary stewardship, benefit-sharing, and longitudinal oversight as coequal pillars of equitable governance. We outline implementable mechanisms—including community veto and withdrawal rights, fiduciary data trusts and cooperatives, benefit-sharing charters, and participatory review boards—that translate participation into co-governance rather than consultation. CDGS reframes evaluation beyond accuracy and compliance to assess whether AI systems deliver redistribution, redress, and repair through durable institutional and technical infrastructures.

CCS Concepts: • **Human-centered computing** → **Human computer interaction (HCI)**; **Collaborative and social computing**; • **Social and professional topics** → *Computing / technology policy*;

Additional Key Words and Phrases: CDGS; AI governance; data sovereignty; refusal rights; fiduciary stewardship; data trusts/co-ops; benefit-sharing; community review boards.

ACM Reference Format:
Anonymous Author(s). 2026. Community Data Governance Stack (CDGS): Engineering Community Co-Governance for Data-Intensive Interactive AI Systems. In *Proceedings of ACM Conference on Fairness, Accountability, and Transparency (ACM FAccT '26).* ACM, New York, NY, USA, 20 pages. https://doi.org/XXXXXXX.XXXXXXX

1 Introduction

Across the expanding terrain of artificial intelligence, governance has become the language of reassurance. Frameworks such as the EU AI Act and the U.S. Blueprint for an AI Bill of Rights aim to mitigate risk and protect individual rights, but they primarily operationalize governance through product-safety requirements and institutional compliance. In the EU context, questions of personal-data control are formally routed through data protection law (GDPR) and related regimes (e.g., IP and competition), yet these tools still largely encode individualized rights and uneven enforcement—leaving collective authority, refusal, and benefit-sharing under-specified as system-level obligations [46, 50, 25, 59]. By emphasizing risk documentation and compliance, these regimes often stabilize extractive infrastructures rather than transforming them [17, 16, 24]. The result is a governance paradigm that may recognize and even manage harm, but rarely redistributes power and equity. From an engineering interactive computing systems perspective, such gaps reflect a broader challenge: translating normative commitments into specifiable requirements, system roles, and validation practices across the interactive system life-cycle [44, 27].

This paper approaches governance from a different premise: AI governance must be reclaimed as a site of democratic and economic transformation. Building on traditions of Indigenous Data Sovereignty (IDS) [43], design justice [15], and labor ethics [30], we propose the **Community Data Governance Stack (CDGS)**—a portable framework for embedding collective community-based authority, fiduciary duty, and the right to refusal into the lifecycle of data and AI model

2026. Manuscript submitted to ACM

development. By *data-intensive interactive AI systems*, we refer to AI-enabled products and services whose interaction and decision-making depend on large-scale data pipelines—spanning data capture, labeling, aggregation, model training, and continuous feedback—such that upstream governance failures directly shape downstream user experience and harm [22, 3, 18]. We present CDGS as a governance requirements and specification model for these systems, defining roles, decision rights, workflows, and enforceable instruments for community co-governance.

CDGS treats data as a social, political, and economic artifact rather than a neutral resource, asserting that equitable governance begins upstream—where data are collected, labeled, and capitalized—rather than downstream in risk assessments or ethics reviews [5, 20].

This reframing also demands what Benjamin describes as a politics of imagination: the collective capacity to envision and prototype alternatives to extractive technological futures, not merely regulate their harms [4]. Rather than treating governance solely as a mechanism of constraint, imagination functions as a mode of institutional experimentation—opening space for communities to design new publics, new ownership arrangements, and new forms of collective stewardship. In this sense, governance becomes not only a site of regulation, but a site of world-building, where alternative data futures are actively constructed rather than passively awaited [6, 4].

Existing governance models often fall into what Selbst et al. (2019) [55] call the abstraction trap—isolating fairness and accountability from the structural conditions that produce inequity. CDGS resists that trap by aligning governance instruments with material redistribution. Through accountable infrastructures and mechanisms such as data trusts, community review/advisory boards, and enforceable refusal protocols, the framework repositions participation as co-governance rather than consultation. In this sense, CDGS is both conceptual and infrastructural: it specifies the roles, processes, and legal tools needed to institutionalize community control over AI systems.

Our contribution is threefold:

(1) **Governance Design Space.** We contribute the *Community Data Governance Stack (CDGS)* as a structured design space for engineering community-centered data governance into interactive AI systems. CDGS organizes governance options across key dimensions—sovereignty, refusal, stewardship, benefit-sharing, and longitudinal oversight—enabling systematic comparison and generative exploration of alternative governance configurations.
(2) **Governance Specification Model.** We provide a conceptual specification model that operationalizes collective governance as implementable system requirements. CDGS specifies the roles, decision rights, processes, and enforceable instruments (e.g., community veto and withdrawal rights, fiduciary data trusts/cooperatives, benefit-sharing charters, and participatory review boards) needed to institutionalize community co-governance beyond consultation.
(3) **Evaluation Criteria Beyond Accuracy.** We reorient evaluation from model-centric metrics toward governance-centric criteria that assess whether interactive AI systems materially redistribute power and support repair. CDGS proposes evaluative dimensions including the enforceability of refusal, accountability of stewardship, durability of oversight, and realized forms of redress, enabling assessment of governance performance rather than predictive accuracy alone.

This work is conceptual and grounded in interdisciplinary synthesis and illustrative precedents from Indigenous data governance, public data trusts, and community-based technology partnerships. CDGS is intended to support researchers and practitioners who engineer, deploy, and evaluate data-intensive interactive AI systems by providing an operational governance model for institutionalizing community participation as enforceable co-governance. CDGS contributes by sequencing these mechanisms into a stack of requirements—defining decision rights, enforceable processes, and evaluable obligations that can be carried through design, deployment, and maintenance.

By reframing governance as an act of reclamation, CDGS intervenes in a field dominated by compliance logic. The question animating this work is no longer how to make AI fair within extractive economies, but how to rebuild data economies that are themselves just in their underlying political–economic foundations.

2 Background

2.1 Data Power and the Limits of Consent

The contemporary data economy consolidates informational and economic power in state and corporate institutions that already possess structural advantage. Datafication is not a passive mirror of the world but an active infrastructure for shaping it—transforming everyday life into extractive inputs for predictive and generative systems [17, 16, 62, 5, 24]. Within this order, social harm is not an accidental by-product of bias but a feature of accumulation.

Prevailing governance regimes respond to these harms through risk-based frameworks—notably the EU AI Act and the U.S. Blueprint for an AI Bill of Rights—which emphasize documentation, transparency, and the classification of "high-risk" systems [25, 50]. Yet, as Veale and Zuiderveen Borgesius (2021) observe, such instruments often translate political conflict into procedural compliance [59]. They rarely disturb the upstream relations of ownership, control, and benefit that structure inequity. In this sense, governance stabilizes extraction rather than redistributes it.

In the EU, this division is partly intentional: the AI Act primarily imposes risk-based system obligations, while GDPR supplies a fundamental-rights framework for personal data (e.g., rights to object, delete, restrict processing), and competition law is increasingly used to contest platform-level data dominance [59, 25]. However, these regimes still tend to (1) treat harms through individualized rights and proceduralism, (2) struggle with meaningful consent at scale, and (3) provide limited pathways for collective decision rights, benefit-sharing, or community-defined refusal—especially for derived, aggregated, or 'non-personal' data flows that remain central to AI development.

The limits of individual consent further expose this imbalance. Consent assumes that data subjects act autonomously and transact on equal terms. But digital data are relational, co-created, and valuable only in aggregate—conditions that make truly informed consent impossible at scale [46, 61]. McNealy (2023) directly argues that consent-based models fail because data are networked representations embedded in social systems, not discrete property objects [46]. As a result, rights frameworks premised on individual notice and choice reproduce what Selbst et al. (2019) call abstraction traps: they locate accountability at the point of disclosure instead of at the point of power [55]. Empirical studies of NLP practitioners further demonstrate how these abstraction traps manifest in organizational practice: responsibility for data equity is frequently diffused across roles, governance mechanisms are perceived as external compliance burdens rather than internal design obligations, and practitioners report limited institutional authority to intervene in upstream data decisions [19]. Justice, therefore, demands collective and structural interventions that redefine who can set boundaries, refuse capture, and share in value creation.

2.2 Toward Collective and Relational Governance

Emerging traditions across disciplines already model what such interventions might entail. Indigenous Data Sovereignty (IDS) provides a living precedent: communities retain authority over how data about them are collected, used, and circulated—including the explicit right to refuse and withdraw [43, 60, 38]. IDS demonstrates that data rights can be collective, contextual, and enforceable through relational accountability. Its tools—Traditional Knowledge labels, CARE principles, and culturally embedded metadata—make provenance and authority visible [13].

Participatory and design-justice traditions likewise seek to reconfigure power in socio-technical design. Value Sensitive Design [26] and community-based HCI foreground the inclusion of stakeholder values, yet critiques within these traditions warn that "participation" often stops short of governance, mirroring longstanding EICS concerns about turning values and stakeholder needs into enforceable requirements and validated system properties [34, 12, 27]. Design Justice advances a more radical premise: justice arises when communities co-own design processes and outcomes, not merely contribute feedback [15]. Similarly, work on data labor exposes the hidden infrastructures of annotation, moderation, and platform micro-work that sustain AI systems while obscuring human contribution [30]. Proposals for worker data collectives and cooperatives extend this reasoning, linking equitable governance to economic redistribution [37].

Scholars in the AI Fairness communities have argued for re-embedding governance within these relational structures. Hoffmann critiques "terms of inclusion" that invite participation without shifting control [35]; Greene et al. (2019) warn that ethics initiatives risk moral laundering if they are not grounded in structural accountability [31]. Birhane and Guest call for decolonizing computational sciences by foregrounding situated epistemologies and community stewardship [8]. Collectively, these interventions signal a turn toward relational governance—a model that treats data as social obligations rather than commodities, underlining how data is embedded in ecological systems of actors, incentives, and power relations [46].

Parallel policy scholarship has similarly argued for reconstituting data governance around institutional controls and collective rights rather than market-driven ownership. Greenwood et al (2019) propose a *"New Deal on Data"* framework that emphasizes fiduciary stewardship, legally grounded data institutions, and public-interest governance as the basis for restoring democratic accountability in data systems. Rather than treating privacy as an individual transaction, this approach reframes data as a shared societal resource requiring durable institutional oversight and enforceable rights [52, 33]. While these proposals articulate the need for structural governance reform at the policy level, they stop short of specifying how such principles translate into operational governance architectures across socio-technical systems. CDGS builds on this institutional orientation by specifying the roles, processes, and instruments through which collective data governance can be enacted in practice [32, 44, 9].

From an engineering interactive computing systems perspective, these calls for collective and relational governance raise a concrete design challenge: how to translate justice-oriented commitments into requirements that can be specified, implemented, and validated within the life-cycle of data-intensive interactive AI systems. While many governance proposals articulate desirable principles (e.g., transparency, accountability, participation), they often remain under-specified as system-level obligations—leaving unclear who holds decision rights, which processes must be supported, and what artifacts or controls make governance enforceable in practice [44, 27, 9]. One partial precedent is documentation artifacts (e.g., dataset nutrition labels / data-sheets for datasets / model cards for ML-models) that render provenance and risks legible, but typically stop short of enforceable decision rights [**chmielinski2022datasetnutritionlabel2gen** , 28, 47, 36]. Addressing this gap requires governance to be treated as a first-class engineering concern, comparable to other quality attributes in interactive systems such as safety, reliability, and accountability [12].

The Community Data Governance Stack (CDGS) builds directly on this lineage by specifying a governance architecture that can be operationalized in system design. It integrates sovereignty, collective ownership, refusal rights, fiduciary stewardship, benefit-sharing, and longitudinal oversight into a coherent stack of roles, processes, and enforceable instruments. By naming and sequencing these commitments as a "stack," CDGS provides a practical scaffold for translating justice-oriented theory into implementable governance mechanisms—including data trusts, cooperative charters, review boards, and impact-and-repair plans—positioning community authority not as an ethical aspiration, but as an infrastructural requirement [27, 9].

2.3 Data Power, Refusal, and Sovereignty

Scholars of data power show how datafication stretches existing extractive relations into everyday informational life, baking social hierarchies into technical systems [17, 62]. Critical work on welfare, credit scoring, and policing makes clear that these systems fall hardest on poor and racialized communities [5, 24]. In that context, refusal is not simply opting out; it is an affirmative act of sovereignty and protection against extraction and surveillance [15, 5, 45, 29], and it reframes consent from a one-time click to an ongoing, collective right to set boundaries and withdraw when necessary.

Design and research traditions such as Value Sensitive Design and community-based participatory research offer practical ways to center impacted communities in agenda setting and evaluation [26, 53], while movements such as Data Justice Lab and Data for Black Lives put these commitments into practice through organizing and institutionally grounded trust-building. At the same time, the EU AI Act and the U.S. OSTP Blueprint focus on risk and transparency but say less about who owns data, who controls it, and who benefits [58, 10, 25, 50]. As a result, governance can end up stabilizing extractive arrangements instead of redistributing power.

Indigenous Data Sovereignty offers a different starting point: it asserts collective rights to govern data in line with cultural protocols, including the explicit right to say no [43], and shows that governance can be relational, reparative, and enforceable. Building on this, the Community Data Governance Stack (CDGS) brings refusal, sovereignty, and participatory design together, laying out concrete mechanisms—such as veto and withdrawal, fiduciary stewardship, benefit-sharing, and longitudinal oversight—for putting justice at the center of AI governance.

3 Conceptual Framework: The Community Data Governance Stack (CDGS)

3.1 Terminology and Scope

Throughout this paper, we differentiate *instruments* and *processes*. Instruments are concrete governance mechanisms such as data trusts or cooperatives, community review boards, governance charters, licensing terms, and impact-and-repair plans. Processes are the iterative activities that surround them, including co-design, ongoing consent and withdrawal oversight and audit, and repair planning. We use *community* to refer to collectives with shared stakes and governance capacity, such as tribal nations, neighborhood associations, labor unions, community-based organizations, and consortium, whose members can articulate boundaries, exercise refusal, and hold fiduciaries to account. We use "stack" deliberately, borrowing the engineering notion of interlocking layers that can be composed, audited, and adopted incrementally—so governance is not a single principle, but a set of interoperable requirements spanning roles, processes, instruments, and measurable outcomes.

3.2 Overview: Governance as Redistribution Infrastructure

The Community Data Governance Stack (CDGS) reframes data governance as infrastructure for redistribution, not a compliance layer for risk mitigation. It integrates insights from Indigenous Data Sovereignty, design justice, and participatory HCI into an enforceable architecture of power-sharing. The model is structured as a stack—a sequence of interlocking layers (principles → roles → processes → instruments → evaluation)—that communities, institutions, or networks can adopt incrementally or in full.

Where existing frameworks center documentation and transparency, CDGS foregrounds ownership, fiduciary duty, refusal, and benefit-sharing as preconditions for legitimacy. Governance becomes a performative act that redistributes authority across the AI lifecycle—from data collection to deployment—so that communities become governors rather than subjects of data systems.

3.3 Axioms (Why): From Neutrality to Relational Accountability

CDGS rests on three foundational commitments that invert conventional assumptions about governance:

Relationality. Data are not neutral facts but social artifacts embedded in histories, obligations, and asymmetries of power [46]. Governance must therefore begin with relational accountability—acknowledging who data represent and what responsibilities accompany their use [20].

Non-neutrality of design. Technical harm arises not from error but from power: sourcing, labeling, and modeling are political acts [5, 8].

Justice over optimization. Governance should pursue repair, redistribution, and refusal rather than post-hoc accuracy tuning [24].

These axioms define a paradigm shift: equitable governance starts upstream, where data are captured and valued, not downstream in audit or compliance reports.

3.4 Rights and Roles (Who): Reallocating Authority

CDGS assigns enforceable rights and duties across three categories of actors:

Community Holders. Collectives with shared stakes—tribal nations, neighborhood associations, unions, HBCU/HSI consortia—hold collective rights of refusal, withdrawal, veto, and benefit-sharing [43, 40].

Trustees / Stewards. Institutions entrusted with data (e.g., universities, CBOs, local governments) owe fiduciary duties to the communities they serve, not to funders or platforms. Fiduciary stewardship translates moral accountability into legal obligation.

Developers / Policymakers. Those building or regulating AI systems must operate under enforceable community terms rather than voluntary ethics [58].

This redistribution of roles transforms governance into a shared, contractual relationship where obligations flow downward and accountability upward.

Micro-vignette. A city contracts with a vendor to deploy an AI system using resident data. Community stakeholders define non-negotiable terms (permitted uses, veto/withdrawal conditions, benefit-sharing), trustees operationalize those terms through access controls and oversight procedures, and developers implement gated releases and auditable compliance with the charter. If the community exercises withdrawal, trustees trigger revocation and deletion workflows and developers must halt updates or retrain under the agreed terms.

3.5 Processes (How): Operationalizing Collective Power

CDGS defines an iterative governance process spanning the data and model lifecycle:

Collective Scoping and Prior Consent. Communities define boundaries—what data are collected, for what purposes, and under what values. Consent is ongoing, revisable, and collective, not one-time or individual [61, 46].

Micro-vignette. A city agency proposes using resident mobility and service-usage data to allocate public benefits. Under CDGS, community stewards scope permissible uses (e.g., service improvement only), prohibit secondary use for enforcement, and define non-negotiable exclusions (e.g., location traces at fine granularity). Consent becomes revisable: the community re-authorizes data use at defined intervals and can narrow or revoke permissions if conditions change.

Fiduciary Governance. Trustees operationalize community mandates through data trusts or cooperatives with auditable access terms, retention policies, and transparent revenue structures [30].

Micro-vignette. A community data trust receives a request from a university lab to access a locally curated dataset for model training. Trustees evaluate the request against chartered duties (purpose limitation, retention, benefit-sharing terms), issue time-bounded access with auditable conditions, and require downstream deletion and retraining triggers if the community later withdraws participation.

Benefit-Sharing. Value generated from data use is redistributed via royalties, dividends, or co-ownership models, recognizing data labor as labor [37].

Micro-vignette. A company licenses a community-annotated dataset to improve a speech system deployed in local services. CDGS routes a negotiated share of value (royalties or a dividend) to a community-controlled fund and requires public reporting on where value accrues. Benefit-sharing is treated as a standing obligation, not a one-time "gift" or acknowledgement.

Longitudinal Oversight and Redress. Community review boards conduct periodic audits, exercise kill-switch powers, and trigger repair plans for remediation and redistribution [24].

Micro-vignette. After deployment, a Community Review Board audits system outcomes and discovers harm concentrated in a specific neighborhood. The Board triggers a pause on updates, mandates remediation (e.g., retraining with revised inclusion rules), and requires a repair action (services, compensation, or decommissioning) tied to a documented timeline.

Across these stages, refusal remains first-class: communities can pause, revise, or terminate participation without penalty—a condition absent from most existing regimes.

3.6 Instruments (What): Making Governance Enforceable

CDGS culminates in a set of instruments—practical, traceable mechanisms that translate principles into enforceable obligations:

Refusal / Veto Protocols Enable collective opt-out, withdrawal, and data deletion within defined timelines. Example: binding clauses in research or licensing agreements.

In practice. Refusal is operationalized as a workflow: a community steward can initiate a veto or withdrawal, the system generates a revocation notice, and an audit log tracks propagation (access termination, deletion, and any required retraining). A concrete implementation can be as simple as a revocable license + time-to-deletion commitments + a public-facing status record.

Data Trust / Co-op Charters Establish fiduciary ownership, access terms, revenue shares, and stewardship duties. Example: legal trust deeds or cooperative bylaws.

In practice. The charter functions like a requirements document for governance: it defines who can approve access, what purposes are permitted, retention limits, and how revenues are distributed. Technically, these terms can be enforced through role-based permissions, time-bounded credentials, and mandatory logging of access and downstream sharing.

Community Review Boards Provide binding oversight on data use, model updates, and audits. Example: mandated review requirements in grant or procurement contracts.

In practice. A Review Board is paired with a recurring review cadence (e.g., quarterly), a standard agenda (new uses, model updates, incidents), and binding decision rights. Board actions (approve/deny/condition/pause) become system events that gate deployments and trigger required documentation or remediation steps.

Impact & Repair Plans Define ex-ante criteria for pausing deployment and mandating remediation. Example: kill-switch procedures or reparations schedules linked to project milestones.

In practice. Impact & Repair Plans define ex-ante thresholds (harm signals, governance breaches, missed deletion deadlines) that automatically trigger escalation: pause, rollback, retraining, compensation, or decommissioning. The plan is enforceable when linked to milestones, procurement terms, and auditable completion criteria.

These instruments make accountability traceable and allow communities to govern both the process and the profit of AI development [46, 56].

3.7 Portability and Adaptation

CDGS is intentionally portable. Its modular structure enables adaptation across institutional contexts:

In academic research, it can anchor data-sharing agreements, IRB protocols, and participatory ethics boards.

In industry, it can inform procurement, auditing, and benefit-sharing standards.

In policy and civic governance, it can support community data infrastructures such as municipal data trusts or cooperative datasets.

Adoption can be incremental—beginning with community review boards or refusal protocols—and expand toward full fiduciary data stewardship. This flexibility allows CDGS to serve as a living architecture that evolves with community capacity and legal reform.

3.8 Evaluation Beyond Accuracy

To determine whether CDGS meaningfully redistributes power, evaluation must shift from technical metrics to governance outcomes [11, 24]:

- **Distributional Impact:** Reduction in disparities or harms across groups; evidence of community influence on decisions.
- **Participation Quality:** Degree of agenda-setting power held by communities; responsiveness to recommendations.
- **Exercised Refusal:** Instances of veto, withdrawal, or data deletion executed without penalty.
- **Material Redress:** Tangible redistribution of value or reparations for harm.
- **Transparency of Process:** Completeness of public logs, audits, and repair reports.

These indicators make governance measurable as justice in action, aligning oversight with the framework's foundational goal: structural repair.

3.9 Summary

The Community Data Governance Stack transforms governance from symbolic ethics into enforceable infrastructure. By sequencing relational principles into actionable instruments, CDGS provides a portable model for building democracy into data systems themselves. It operationalizes what Benjamin calls abolitionist tools for the New Jim Code: institutions capable of refusal, repair, and redistribution [5]. Through CDGS, governance becomes not a means of managing risk, but an ongoing practice of reclaiming data power.

3.10 Validation and Fidelity

Since CDGS is proposed as a governance specification, we define *fidelity* as whether its layers are implemented as enforceable capabilities rather than symbolic participation. Fidelity can be evaluated by (1) confirming the presence of required artifacts (commitments, roles, instruments), (2) verifying traceability from governance decisions to system enforcement actions, and (3) reporting auditable governance outcomes beyond model metrics (e.g., withdrawal/refusal events, benefit-sharing distributions, redress actions, and process transparency). Fidelity may be reported by layer to support incremental adoption.

4 Applications and Illustrative Cases

Operationalizing the Community Data Governance Stack

The Community Data Governance Stack (CDGS) is designed as a portable architecture: its layers—principles, roles, processes, and instruments—can be adapted to various institutional, cultural, and economic contexts. While conceptual in origin, CDGS gains traction through application. This section illustrates how the stack's mechanisms can be operationalized across real and emergent governance arrangements. We expand the brief micro-vignettes introduced alongside the CDGS processes and instruments into fuller illustrative cases below. The examples that follow—drawn from Indigenous Data Sovereignty, community-based data cooperatives, and university–community research partnerships—demonstrate that the movement from compliance to co-governance is not hypothetical but already unfolding in fragments. CDGS serves to unify, scale, and render these practices enforceable. Relatedly, public-interest scholarship on participatory data stewardship offers implementation-oriented principles for involving publics in data governance decisions—useful baselines for translating CDGS's roles, processes, and instruments into operational practice [48, 51, 1, 2, 49].

4.1 Indigenous Data Sovereignty: The Right to Refuse as Governance

Indigenous Data Sovereignty (IDS) exemplifies what collective control over data can look like in practice. Across Aotearoa New Zealand, Canada, and Australia, Indigenous-led networks such as Te Mana Raraunga and the Maiam nayri Wingara collective assert the inherent rights of Indigenous peoples to determine how data about them are collected, used, and shared [43, 60]. These frameworks embed relational accountability, cultural protocols, and enforceable refusal into data governance. The CARE Principles for Indigenous Data Governance—Collective Benefit, Authority to Control, Responsibility, and Ethics—demonstrate the operationalization of what CDGS identifies as the "Rights & Roles" and "Instruments" layers [13].

Traditional Knowledge (TK) Labels signal culturally specific access and reuse restrictions, functioning as refusal protocols embedded directly into metadata. Such mechanisms move beyond the optics of inclusion toward sovereign governance. IDS initiatives illustrate that data can be stewarded through fiduciary duties and collective mandates rather than abstract openness. CDGS builds on this precedent by generalizing the logic of relational sovereignty across other forms of collective governance, emphasizing that refusal and redress are not special exceptions but governance defaults.

4.2 Community Data Cooperatives: Economic Redistribution through Fiduciary Governance

A second application emerges in civic and community data cooperatives, which aim to reclaim control over data labor and its economic value. Projects such as MIDATA in Switzerland and the DECODE Project in Europe establish participatory infrastructures through which individuals and communities share data under cooperative ownership, retaining rights to revoke, license, or monetize its use. These entities operationalize fiduciary stewardship and benefit-sharing—core features of CDGS's "Processes" and "Instruments" layers.

Within these cooperatives, members act as both data subjects and data trustees, deciding collectively how data may contribute to research or innovation [21]. Revenue from data use is redistributed through dividends or reinvested in public goods, aligning governance with material equity. Yet such projects often lack formal refusal mechanisms or enforceable oversight; withdrawal is permitted contractually but rarely supported through recurring review or kill-switch authority. CDGS addresses these limitations by codifying refusal, fiduciary duty, and longitudinal oversight as integrated obligations rather than optional ethics.

The stack thereby transforms data cooperatives from participatory experiments into full governance infrastructures. In doing so, it complements calls for equitable data economies that treat ethical stewardship as financially material to sustainability and trust [54].

4.3 A Creative-Work Data Trust: Singer Governance in Choral AI

A useful boundary case for CDGS is creative-work training data. In recent work on choral AI, a data trust structure was used to negotiate how singers' recordings could be collected, governed, and used, foregrounding consent, stewardship, and terms of reuse in a domain where identity, IP, and labor value are tightly coupled [39]. This maps cleanly onto CDGS: collective scoping (who counts as the contributing group), fiduciary stewardship (who holds the data and under what duties), enforceable refusal/withdrawal (revocation pathways), and benefit-sharing (terms for downstream reuse). The case also illustrates where GDPR/IP frameworks help, but do not fully specify collective governance workflows or community-facing interfaces—precisely the gap CDGS treats as an engineering concern.

4.4 University–Community Partnerships: From Consent to Co-Governance

Academic institutions have long been sites of extractive data practices, especially in research involving marginalized communities. Institutional Review Boards (IRBs) prioritize individual consent and risk minimization but rarely ensure community control over data use, reuse, or benefit distribution. Applying CDGS within university–community partnerships—such as collaborations with HBCU consortia [40], Tribal Colleges, or local community-based organizations—offers a pathway from consultation to co-governance. This translational move—building governance mechanisms with communities so they are usable, accountable, and durable in practice—aligns with translational ethics scholarship emphasizing community-centered infrastructures rather than one-off consent or consultation [23].

In this model, community representatives become trustees within governance charters that define collective rights to veto, withdraw, or share in benefits arising from data-driven research. Community Review Boards (CRBs)—analogous to IRBs but community-constituted—exercise oversight authority across the research lifecycle. Through fiduciary trust agreements, universities formalize duties to act in the community's best interest, including transparent auditing, shared data licensing, and reparative commitments when harms occur.

This approach extends participatory design into the realm of enforceable governance. Where IRBs ask, "Is this study ethical?", CRBs empowered under CDGS ask, "Is this governance just?"—a question that aligns research ethics with redistribution and accountability [34, 20].

4.5 Worker Data Collectives: Governance as Labor Power

CDGS also aligns with emerging efforts to organize the distributed labor that fuels AI. Platform workers who annotate, moderate, or label data—often under precarious conditions—have begun forming worker data collectives to assert collective rights over their contributions [30, 37]. These collectives, when structured as cooperatives or fiduciary trusts, embody CDGS's benefit-sharing mechanisms and refusal protocols.

Through fiduciary charters, workers can define terms of data use, negotiate royalties from model training, and demand transparency in value extraction. CDGS provides the architectural logic for such arrangements: data labor becomes governed through contracts of care rather than extraction. Governance merges with labor organizing—transforming data stewardship into a domain of economic and political empowerment.

4.6 Cross-Case Synthesis: From Fragmentation to Infrastructure

Taken together, these examples illustrate that the principles underlying CDGS already exist across multiple sectors, but in fragmented form. Indigenous networks exemplify sovereign refusal; cooperatives demonstrate fiduciary redistribution; universities model participatory oversight; and labor collectives signal economic redress. What is missing is a unifying governance architecture capable of translating these relational practices into enforceable, interoperable infrastructures.

CDGS provides that architecture. It connects the moral and political aspirations of these movements to specific mechanisms—charters, veto protocols, fiduciary duties—that render accountability traceable and repair measurable. In doing so, it moves governance from the rhetorical to the structural, showing that redistribution is not an aspiration external to AI but a design property of its institutions.

4.7 Reflection: Building a Reparative Data Future

These case illustrations point toward a reparative future in which communities define the boundaries, benefits, and obligations of data systems that shape their lives. Rather than proposing a new ethics checklist, CDGS offers a governance grammar: a way to sequence power, responsibility, and care. It recognizes that justice in AI requires not only fairer models but fairer infrastructures—those that embed refusal, fiduciary duty, and redistribution as first-class operations.

In this way, CDGS bridges theory and practice: it translates the relational ethics of data sovereignty and design justice into the enforceable mechanics of policy, contract, and institutional design. The challenge ahead is not whether such models can exist—they already do—but how to connect, resource, and scale them into the infrastructures emerging AI development life-cycle.

5 Discussion: From Risk to Redistribution — Reclaiming Data Power as Governance

Current AI governance paradigms continue to orbit risk and compliance—often adjacent to (but not resolved by) data protection and competition regimes—as if inequity in data systems could be managed through documentation and disclosure alone [58, 25, 50]. Yet the crises that animate the reproduction of harm, the appropriation of labor, and the erasure of refusal are not failures of implementation but of imagination. They arise from a governance logic that stabilizes existing power, converting extractive infrastructures into objects of technical optimization. In this landscape, risk becomes a management category rather than a symptom of dispossession.

The Community Data Governance Stack (CDGS) reframes governance not as a post hoc safeguard but as an act of reclamation. If most AI systems are built on unreciprocated data labor [30] and weakly consented corpora [10], then justice must begin where data power is produced—not where it is audited. The provocation here is simple but radical: governance should not merely manage risk; it should redistribute power.

5.1 Beyond Risk: Governance as Redistribution

Risk-based regimes such as the EU AI Act and the U.S. Blueprint for an AI Bill of Rights are structured around documenting and mitigating harm more than preventing its structural conditions. They center individual rights and transparency but rarely interrogate ownership, control, or the material conditions of data production [59]. By contrast, CDGS insists that equitable AI requires the upstream redistribution of data power through community-defined ownership, veto, and benefit-sharing. This orientation echoes policy proposals for a "New Deal on Data," which call for institutional controls, collective rights, and public-interest stewardship over data infrastructures rather than market-driven ownership alone [52, 32, 33].

This move from risk to redistribution aligns with what Selbst et al. (2019) call a shift away from abstraction traps—the tendency to separate technical problems from structural realities—toward socio-technical specificity [55]. CDGS operationalizes this shift by defining governance instruments that bind participation to material consequence [44]: review boards with veto authority, data trusts with fiduciary duties, and repair protocols that redistribute value. Governance becomes not a procedural layer, but a redistributive infrastructure.

5.2 Refusal as a Design Principle

Refusal, as articulated in Indigenous Data Sovereignty [43] and design justice scholarship [15], is not disengagement but governance. It is an affirmative act that asserts the right to opacity, to say no, to withdraw, and to be forgotten. In the current AI economy, refusal is treated as a defect—an interruption to the smooth flow of data accumulation. CDGS reverses that logic, framing refusal as a signal of collective agency and epistemic integrity.

This echoes Birhane's critique of "algorithmic colonization," where data extraction mirrors historical patterns of appropriation [7]. When refusal becomes a first-class mechanism—codified in licenses, withdrawal clauses, and charters—it transforms governance from permission-seeking into boundary-making. In doing so, it reclaims data relations as ethical and political negotiations rather than technical exchanges.

5.3 From Ethics to Economics: The Value of Stewardship

A persistent challenge in AI governance is the false dichotomy between ethics and innovation. Corporate frameworks often treat equity as external to performance, as if justice must be subsidized by profit rather than integrated with it [24, 62]. Yet the history of extractive industries shows that systems built on exploitation are also fragile: they produce backlash, mistrust, and long-term inefficiency.

CDGS foregrounds the economic materiality of ethical governance. When communities co-own data, participate in benefit-sharing, and retain rights of refusal, they not only mitigate harm but create new value systems grounded in legitimacy and sustainability. This is consistent with emerging arguments that responsible data practices can reduce downstream harms and strengthen institutional accountability [54]. By treating fiduciary stewardship and redistribution as infrastructure, CDGS reframes justice as a condition for innovation rather than its constraint.

5.4 Toward Infrastructures of Accountability

The question of "who governs AI" cannot be answered through ethics statements or compliance audits. Accountability requires infrastructure—legal, technical, and institutional—that communities can operate, not merely observe [46, 56, 42]. CDGS envisions this through a stack of enforceable instruments: veto protocols, cooperative charters, review boards, and impact-and-repair plans. These instruments translate normative commitments into traceable obligations, making governance performative rather than declarative.

Greene et al. (2019) argue that fairness without structural accountability risks reproducing harm under a new moral vocabulary [31]. Similarly, Hoffmann warns against “terms of inclusion” that invite marginalized participation without altering who sets the terms [35]. CDGS responds by embedding governance within relational infrastructures of community control—contracts, audits, and financial mechanisms that can be revoked or restructured. Accountability, then, is not a value but a practice of enforceable co-governance.

5.5 The Abolitionist Horizon: Governance as Structural Repair

Following Ruha Benjamin’s call for abolitionist tools for the New Jim Code, CDGS treats governance as a site of repair rather than regulation [5]. It rejects the notion that equitable AI can emerge from the same institutional logics that produced extractive data relations. To govern AI justly requires dismantling those logics—redistributing authority, reconstituting institutions, and building new publics around stewardship.

This abolitionist orientation reframes governance from a procedural safeguard to a structural reparation. It acknowledges that communities historically subjected to surveillance and dispossession—Indigenous nations, Black neighborhoods, labor collectives—must be resourced as governors, not stakeholders. In this sense, CDGS is not merely a technical model but a political claim: that data governance is the terrain on which democratic accountability in AI must be rebuilt.

5.6 Future Directions: Building a Reparative Data Economy

The implications of CDGS extend beyond governance to the broader political economy of AI. If data is the substrate of value creation, then redistributing control over data is an act of economic justice. The next frontier for research lies in developing metrics and mechanisms that make this redistribution legible—tracking not only accuracy or bias, but the flow of value, authority, and redress.

Building a reparative data economy means funding community-owned data infrastructures, formalizing fiduciary duties in AI procurement, and integrating repair costs into model lifecycles. This aligns with Keyes’ analysis of ‘misgendering machines’: systems trained on harm cannot produce justice without structural reinvention [41]. CDGS offers one blueprint for such reinvention—a stack that translates relational ethics into enforceable governance.

5.7 Closing Reflection: Reclaiming Data Power

To reclaim data power is to shift the locus of governance from compliance to co-creation, from harm management to structural repair. It is to recognize, as Taylor argues, that data justice requires connecting digital rights to global struggles for freedom and self-determination [57]. The provocation, then, is not merely to refine governance frameworks but to re-imagine them entirely: to build AI systems where communities hold the keys, where refusal is celebrated as integrity, and where accountability is measured in redistribution, not rhetoric.

In sum, the CDGS invites scholars and practitioners to treat AI governance as the site of democratic renewal. The question is no longer how to make AI fair, but how to make governance just.

6 Implications for Design, Policy, and Evaluation

6.1 Implications for Design and Policy

For designers / HCI practitioners. Treat participation as shared power, not consultation. Co-theorize the problem and define success criteria with community holders; encode these in system objectives and evaluation metrics. Make inclusion and exclusion decisions with communities, document rationales, and expose trade-offs [26, 20]. Build refusal into systems by design (opt-out, withdrawal, deletion), and keep these pathways available across the entire lifecycle.

For policymakers. Move beyond notice and consent toward collective rights. Recognize data trusts and cooperatives in statute; assign fiduciary duties to trustees [58]. Require community review boards for public and academic AI projects, and grant them binding authority over data access, model updates, and system decommissioning. Tie public funding and procurement to benefit-sharing (royalties, dividends, co-ownership) and to enforceable withdrawal terms [24].

For institutions. Invest in public-interest infrastructure. Resource local data repositories; train community stewards; publish standard charters and revocable licenses. Provide open audit tools and public logs; fund longitudinal oversight mechanisms that operate independently of principal investigators.

6.2 Risks and Limits (with Mitigations)

Institutional capture / participation-washing. Mitigations: independent funding for community trustees; conflict-of-interest rules; term limits; transparent audit trails.

Governance fatigue. Mitigations: stipends; predictable cadences (e.g., quarterly reviews); scoped decision rights; streamlined veto and withdrawal procedures.

Legal and contractual friction. Mitigations: template charters and licenses with revocation clauses; pre-approved procurement language; university counsel trained in fiduciary duties.

Over-generalization. Mitigations: mandate local scoping and co-definition of governance terms; treat CDGS as modular and configurable rather than prescriptive; encode cultural protocols in charters and operating procedures; re-authorize governance terms through periodic review.

6.3 Evaluation Beyond Accuracy

Assess whether governance redistributes power and delivers repair [11, 24, 42]:

- **Distributional impact:** changes in error or disparity rates across groups; decisions altered by community intervention.
- **Participation quality:** attendance and retention in review boards; proportion of agenda items set by communities; timeliness of responses to recommendations.
- **Exercise of refusal / withdrawal:** number and outcomes of vetoes; time-to-deletion; proportion of datasets with active withdrawal pathways.
- **Material redress:** revenue or value redistributed; remediation delivered after harm; decommissioning events triggered.
- **Process transparency:** completeness of public logs; frequency of audits; publication of impact-and-repair reports.

Implementing CDGS requires resourcing community stewardship, making refusal enforceable, and evaluating governance on repair and redistribution—not solely on risk or accuracy [26, 20, 58, 24, 43].

7 Conclusion: Governance as the Work of Reparation

Artificial intelligence now mediates access to credit, care, education, labor, and public representation. Yet its governance remains structured by logics of risk and compliance that treat harm as a technical failure rather than a social condition. This paper has argued that just governance cannot emerge from systems designed to extract; it must be built from architectures that redistribute authority and value. The Community Data Governance Stack (CDGS) advances such an architecture—one that redefines data governance as a site of relational accountability, fiduciary duty, and collective power.

CDGS reframes governance as a living infrastructure rather than a policy checklist. Its layered design—anchored in axioms of justice, collective rights, fiduciary processes, enforceable instruments, and redistributive evaluation—translates abstract commitments into actionable mechanisms and and governance requirements. Through case illustrations in Indigenous Data Sovereignty, civic cooperatives, academic partnerships, and worker collectives, we showed that building blocks of this paradigm already exist. What remains is to connect, scale, and institutionalize them under a coherent logic of repair.

To reclaim data power is to recognize that democracy in AI cannot be delegated to compliance regimes. It requires communities who can veto, withdraw, and share in the value their data produce. It requires trustees who are legally bound to act in the public's interest. It requires infrastructures that make accountability traceable, and refusal enforceable.

The provocation, then, is not to perfect existing governance models but to replace their foundations. Governance must evolve from a mechanism for managing harm to a practice of structural reparation. The future of equitable AI will not hinge on better documentation or more ethical models—it will depend on building institutions capable of justice.

In this vision, governance becomes design, design becomes policy, and policy becomes repair. CDGS offers one blueprint for that transformation: a stack through which communities reclaim authority, redistribute value, and remake the terms on which intelligence itself is governed.

7.1 Future Work

Future research should focus on translating the Community Data Governance Stack (CDGS) into operational prototypes that can be specified and evaluated in situ. For example, one prototype could implement a community veto-and-withdrawal workflow as an interactive system feature—pairing a revocable data license interface with an auditable log that records governance decisions, downstream data access, and time-to-deletion commitments. This includes developing governance metrics for redistribution, refusal, and fiduciary accountability; testing community-led review boards and data trust charters within academic and civic partnerships; and collaborating with policymakers to explore legal recognition of collective data rights.

The *CDGS Steward Console*—A minimal implementation could include:

(1) a role-based governance dashboard for designated community stewards to approve, deny, or revoke dataset access requests;
(2) a withdrawal interface that triggers automated propagation of deletion or suppression requests across downstream storage and training pipelines;

(3) an auditable decision log that records governance actions, justification fields, affected assets, and time-to-deletion service-level commitments;
(4) a public-facing transparency view that exposes aggregate governance outcomes without revealing sensitive deliberation details; and
(5) a compliance status monitor that verifies whether deletion, retraining, or decommissioning actions were completed within agreed timelines.

Beyond design and policy, this work aims to build a community of practice that treats governance itself as an act of repair—embedding justice not at the edges of AI systems but at their core.

References

[1] Ada Lovelace Institute. 2021. Exploring Legal Mechanisms for Data Stewardship. Report. Ada Lovelace Institute, (Mar. 2021). Retrieved Feb. 13, 2026 from https://www.adalovelaceinstitute.org/report/legal-mechanisms-data-stewardship/.

[2] Ada Lovelace Institute. [n. d.] Participatory and inclusive data stewardship (project). Project page. Accessed 13 Feb 2026. (). Retrieved Feb. 14, 2026 from https://www.adalovelaceinstitute.org/project/participatory-inclusive-data-stewardship/.

[3] Barbara Rita Barricelli, José Creissac Campos, Kris Luyten, Sven Mayer, Philippe Palanque, Emanuele Panizzi, Lucio Davide Spano, and Simone Stumpf. 2025. Engineering interactive systems embedding ai technologies (3rd workshop on). In *Companion Proceedings of the 17th ACM SIGCHI Symposium on Engineering Interactive Computing Systems* (EICS '25 Companion). Association for Computing Machinery, Trier, Germany, 71–75. isbn: 9798400718663. doi:10.1145/3731406.3735965.

[4] Ruha Benjamin. 2024. *Imagination: A manifesto*. WW Norton & Company.

[5] Ruha Benjamin. 2019. *Race after Technology: Abolitionist Tools for the New Jim Code*. Polity.

[6] Ruha Benjamin. 2022. *Viral justice: How we grow the world we want*. Princeton University Press.

[7] Abeba Birhane. 2020. Algorithmic colonization of africa. *SCRIPTed*, 17, 2, 389–409.

[8] Abeba Birhane and Olivia Guest. 2020. Towards decolonising computational sciences. *arXiv preprint arXiv:2009.14258*.

[9] Judy Bowen, Marco Winckler, and Jean Vanderdonckt. 2020. A glimpse into the past, present, and future of engineering interactive computing systems. *Proceedings of the ACM on Human-Computer Interaction*, 4, EICS, Article 71, 71:1–71:32. doi:10.1145/3394973.

[10] danah boyd danah and Kate Crawford. 2012. Critical questions for big data: provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society*, 15, 5, 662–679.

[11] Joy Buolamwini and Timnit Gebru. 2018. Gender shades: intersectional accuracy disparities in commercial gender classification. In *Proceedings of the Conference on Fairness, Accountability, and Transparency*. ACM, 77–91.

[12] Gaëlle Calvary, Jeffrey Nichols, José Creissac Campos, Nuno Jardim Nunes, and Pedro Campos. 2017. Editorial statement. *Proceedings of the ACM on Human-Computer Interaction*, 1, EICS, Article 1, 1:1–1:2. doi:10.1145/3106386.

[13] Stephanie Russo Carroll, Edit Herczog, Maui Hudson, Keith Russell, and Shelley Stall. 2021. Operationalizing the care and fair principles for indigenous data futures. *Scientific data*, 8, 1, 108.

[14] Kasia S. Chmielinski, Sarah Newman, Matt Taylor, Josh Joseph, Kemi Thomas, Jessica Yurkofsky, and Yue Chelsea Qiu. 2022. The dataset nutrition label (2nd gen): leveraging context to mitigate harms in artificial intelligence. *arXiv preprint arXiv:2201.03954*. https://arxiv.org/abs/2201.03954 arXiv: 2201.03954 [cs.CY].

[15] Sasha Costanza-Chock. 2020. *Design justice: Community-led practices to build the worlds we need*. The MIT Press.

[16] Nick Couldry and Ulises A Mejias. 2019. Data colonialism: rethinking big data's relation to the contemporary subject. *Television & new media*, 20, 4, 336–349.

[17] Nick Couldry and Ulises A. Mejias. 2019. *The Costs of Connection: How Data is Colonizing Human Life and Appropriating It for Capitalism*. Stanford University Press.

[18] Jay L Cunningham. 2024. *Methods of Designing Justice-oriented Interactive AI Systems*. Ph.D. Dissertation. University of Washington.

[19] Jay L Cunningham, Kevin Zhongyang Shao, Rock Yuren Pang, and Nathanael Elias Mengist. 2025. Advancing nlp data equity: practitioner responsibility and accountability in nlp data practices. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society* number 1. Vol. 8, 679–691.

[20] Catherine D'Ignazio and Lauren F. Klein. 2020. *Data Feminism*. MIT Press.

[21] Sylvie Delacroix and Neil D Lawrence. 2019. Bottom-up data trusts: disturbing the 'one size fits all'approach to data governance. *International data privacy law*, 9, 4, 236–252.

[22] Alan Dix, Sven Mayer, Philippe Palanque, Emanuele Panizzi, and Lucio Davide Spano. 2023. Engineering interactive systems embedding ai technologies. In *Companion Proceedings of the 2023 ACM SIGCHI Symposium on Engineering Interactive Computing Systems* (EICS '23 Companion). Association for Computing Machinery, Swansea, United Kingdom, 90–92. isbn: 9798400702068. doi:10.1145/3596454.3597195.

[23] Megan Doerr and Joon-Ho Yu. 2023. Translational research and communities. *Ethics & Human Research*, 45, 5, 34–38. doi:10.1002/eahr.500181.

[24] Virginia Eubanks. 2018. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor.* St. Martin's Press.

[25] European Commission. 2021. Proposal for a regulation laying down harmonised rules on artificial intelligence (artificial intelligence act). Brussels. (2021).

[26] Batya Friedman, Peter H. Kahn, and Alan Borning. 2002. Value Sensitive Design: Theory and Methods. Tech. rep. University of Washington.

[27] Mathias Funk, Rong-Hao Liang, Philippe Palanque, Jun Hu, and Panos Markopoulos. 2021. Designing and engineering interactive computing systems. *Proceedings of the ACM on Human-Computer Interaction*, 5, EICS, Article 193, 193:1–193:4. doi:10.1145/3457140.

[28] Timnit Gebru, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, Hal Daumé III, and Kate Crawford. 2021. Datasheets for datasets. *Communications of the ACM*, 64, 12, 86–92. doi:10.1145/3458723.

[29] Chris Gilliard and Hugh Culik. 2016. Digital redlining, access, and privacy. *Common Sense Education*. Essay.

[30] Mary L. Gray and Siddharth Suri. 2019. *Ghost Work: How to Stop Silicon Valley from Building a New Global Underclass.* Houghton Mifflin Harcourt.

[31] Daniel Greene, Anna Lauren Hoffmann, and Luke Stark. 2019. Better, nicer, clearer, fairer: a critical assessment of the movement for ethical artificial intelligence and machine learning. *Proceedings of the Conference on Fairness, Accountability, and Transparency*, 1–9.

[32] Daniel Greenwood, Arkadiusz Stopczynski, Brian Sweatt, Thomas Hardjono, and Alex Pentland. 2019. The new deal on data: a framework for institutional controls. In *Privacy, Big Data, and the Public Good: Frameworks for Engagement.* Julia Lane, Victoria Stodden, Stefan Bender, and Helen Nissenbaum, (Eds.) Cambridge University Press, 192–210. doi:10.1017/CBO9781107590205.012.

[33] Thomas Hardjono, David L. Shrier, and Alex Pentland, (Eds.) 2019. *Trusted Data: A New Framework for Identity and Data Sharing.* MIT Press. ISBN: 9780262356053.

[34] Christina N. Harrington, Sheena Erete, and Anne Marie Piper. 2019. Deconstructing community-based collaborative design: toward more equitable participatory design engagements. *Proceedings of the ACM on Human–Computer Interaction*, 3, CSCW, 1–25.

[35] Anna Lauren Hoffmann. 2021. Terms of inclusion: data, discourse, violence. *New Media & Society*, 23, 12, 3539–3556.

[36] Sarah Holland, Ahmed Hosny, Sarah Newman, Joshua Joseph, and Kasia Chmielinski. 2018. The dataset nutrition label: a framework to drive higher data quality standards. *arXiv preprint arXiv:1805.03677.* https://arxiv.org/abs/1805.03677 arXiv: 1805.03677 [cs.CY].

[37] Jyun-Yu Hsieh et al. 2024. Worker data collectives as a means to improve accountability, combat surveillance and reduce inequalities. In *Companion Publication of CSCW 2024.* ACM, 1–6. doi:10.1145/3678884.3681829.

[38] Maui Hudson, Stephanie Russo Carroll, Jane Anderson, D. Blackwater, F. Cordova-Marks, J. Cummins, R. K. Rowe, et al. 2023. Indigenous peoples' rights in data: a contribution toward indigenous research sovereignty. *Frontiers in Research Metrics and Analytics*, 8, 1173805. doi:10.3389/frma.2023.1173805.

[39] Victoria Ivanova and Jennifer Ding. 2025. Choral data 'trust' experiment white paper. (Feb. 2025). doi:10.5281/zenodo.14859320.

[40] Jazette Johnson, Lucretia Williams, Jaye Nias, Saurav K Aryal, and Gloria Washington. 2025. Centering black voices: lessons learned and reflections from a large-scale aave data collection at a historically black university. In *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, 1–7.

[41] Os Keyes. 2018. The misgendering machines: trans/hci implications of automatic gender recognition. *Proceedings of the ACM on human-computer interaction*, 2, CSCW, 1–22.

[42] Joshua A. Kroll. 2021. Outlining traceability: a principle for operationalizing accountability in computing systems. *arXiv preprint arXiv:2101.09385.* https://arxiv.org/abs/2101.09385.

[43] Tahu Kukutai and John Taylor, (Eds.) 2016. *Indigenous Data Sovereignty: Toward an Agenda.* ANU Press.

[44] Víctor Manuel López Jaquero, Radu-Daniel Vatavu, Jose Ignacio Panach, Oscar Pastor, and Jean Vanderdonckt. 2019. A newcomer's guide to eics, the engineering interactive computing systems community. *Proceedings of the ACM on Human-Computer Interaction*, 3, EICS, Article 1, 1:1–1:9. doi:10.1145/3300960.

[45] Tressie McMillan Cottom. 2015. Who do you think you are? when marginality meets academic microcelebrity. *Ada: Journal of Gender, New Media, and Technology*, 7.

[46] Jasmine McNealy. 2023. An ecological approach to data governance. *Notre Dame JL Ethics & Pub. Pol'y*, 37, 523.

[47] Margaret Mitchell, Simone Wu, Andrew Zaldivar, Parker Barnes, Lucy Vasserman, Ben Hutchinson, Elena Spitzer, Inioluwa Deborah Raji, and Timnit Gebru. 2019. Model cards for model reporting. In *Proceedings of the Conference on Fairness, Accountability, and Transparency.* Association for Computing Machinery, 220–229. doi:10.1145/3287560.3287596.

[48] Roshni Modhvadia and Octavia Field Reid. 2024. Participatory and Inclusive Data Stewardship: A Landscape Review. Report. Contributing authors: Valentina Pavel and Luke Patterson. Ada Lovelace Institute, (Dec. 2024). Retrieved Feb. 14, 2026 from https://www.adalovelaceinstitute.org/report/participatory-inclusive-data-stewardship/.

[49] Mozilla Foundation. 2024. A practical framework for applying ostrom's principles to data commons governance. Blog post. Accessed 13 Feb 2026. (Dec. 2024). Retrieved Feb. 13, 2026 from https://foundation.mozilla.org/en/blog/a-practical-framework-for-applying-ostromsprinciples-to-data-commons-governance/.

[50] Office of Science and Technology Policy. 2022. Blueprint for an ai bill of rights: making automated systems work for the american people. Executive Office of the President. (2022).

[51] Reema Patel, Aidan Peppin, Valentina Pavel, Jenny Brennan, Imogen Parker, and Cansu Safak. 2021. Participatory Data Stewardship: Data for the Public Good. Report. Ada Lovelace Institute, (Sept. 2021). Retrieved Feb. 14, 2026 from https://www.adalovelaceinstitute.org/wp-content/uploads/2021/09/Participatory-data-stewardship_Final-report.pdf.

[52] Alex Pentland. 2009. Reality mining of mobile communications: toward a new deal on data. In *The Global Information Technology Report 2008-2009 – Mobility in a Networked World*. World Economic Forum.

[53] Vinodkumar Prabhakaran and Daniel Martin. 2020. Participatory machine learning using community-based system dynamics. *Health and Human Rights Journal*, 22, 2, 71–74.

[54] Inioluwa Deborah Raji, Andrew Smart, Rebecca N. White, Margaret Mitchell, Timnit Gebru, Ben Hutchinson, Jamila Smith-Loud, Daniel Theron, and Parker Barnes. 2020. Closing the ai accountability gap: defining an end-to-end framework for internal algorithmic auditing. In *Proceedings of the ACM Conference on Fairness, Accountability, and Transparency (FAccT)*. ACM.

[55] Andrew D Selbst, Danah Boyd, Sorelle A Friedler, Suresh Venkatasubramanian, and Janet Vertesi. 2019. Fairness and abstraction in sociotechnical systems. In *Proceedings of the conference on fairness, accountability, and transparency*, 59–68.

[56] Jatinder Singh, Jennifer Cobbe, and Christopher Norval. 2018. Decision provenance: harnessing data flow for accountable systems. *IEEE Access*, 7, 6562–6574. doi:10.1109/ACCESS.2018.2887201.

[57] Linnet Taylor. 2017. What is data justice? the case for connecting digital rights and freedoms globally. *Big Data & Society*, 4, 2, 2053951717736335.

[58] Michael Veale and Reuben Binns. 2017. Fairer machine learning in the real world: mitigating discrimination without collecting sensitive data. *Big Data & Society*, 4, 2.

[59] Michael Veale and Frederik Zuiderveen Borgesius. 2021. Demystifying the draft eu artificial intelligence act. *Computer Law Review International*, 22, 4, 97–112.

[60] Maggie Walter, Robert Lovett, B. Maher, B. Williamson, Jane Prehn, G. Bodkin-Andrews, and V. Lee. 2020. Indigenous data sovereignty in the era of big data and open data. *Australian Journal of Social Issues*, 55, 3, 1–14. doi:10.1002/ajs4.141.

[61] William H. Wong, John Duncan, and David A. Lake. 2024. Why data about people are so hard to govern. *Regulation & Governance*, 19, 1, 236–252. doi:10.1111/rego.12591.

[62] Shoshana Zuboff. 2019. *The Age of Surveillance Capitalism*. PublicAffairs.

Received 20 February 2007; revised 12 March 2009; accepted 5 June 2009

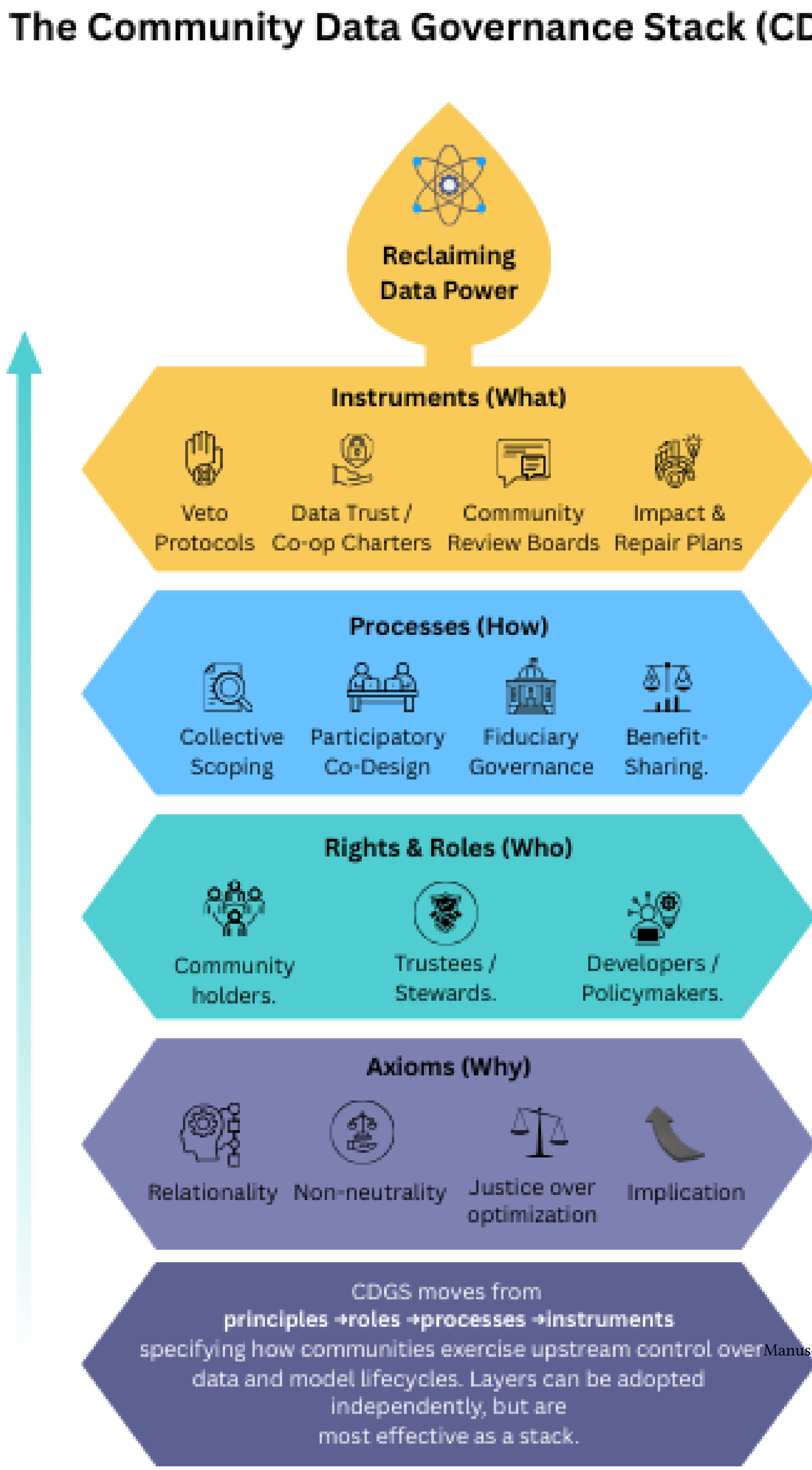

Fig. 1. The Community Data Governance Stack (CDGS). From principles to power—how communities reclaim data governance. The stack visualizes layered transitions from axioms (why) through rights and roles (who), processes (how), and instruments (what), culminating in evaluation (outcomes) that centers redistribution and repair. A vertical axis moves from compliance toward co-governance and redistribution, while a parallel axis indicates a shift from technical oversight to democratic stewardship.

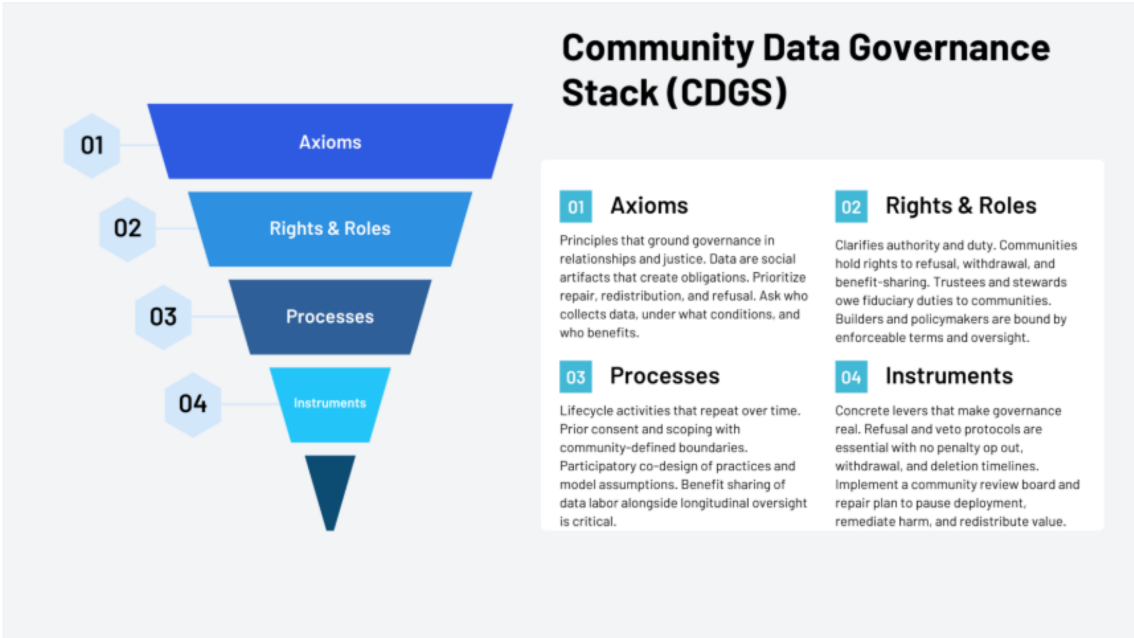

Fig. 2. The Community Data Governance Stack (CDGS) represented as a layered funnel. The figure illustrates how principles flow upward into roles, processes, and instruments, enabling communities to exert upstream control over the data and model lifecycle. While layers can be adopted individually, the stack is most effective when implemented together as a redistributive governance architecture rather than risk management.