



TECHNIUM
SOCIAL SCIENCES JOURNAL

Vol. 79/2026
A New Decade for Social Changes



PLUS
COMMUNICATION P



International
Communication & PR

Artificial Intelligence and Predictive Food Insecurity: A Conceptual Framework for Ethical Transparency and Trust (BTT)

Saada Reuveni

Researcher in economics and business in developing countries

saadare@gmail.com

Abstract. This theoretical study introduces the Bias–Transparency–Trust (BTT) framework, a conceptual model for examining the ethical dimensions of artificial intelligence (AI) in predictive food-security systems across developing countries. As organizations such as FAO and WFP increasingly employ machine learning to anticipate hunger crises, ethical challenges surrounding data bias, opacity, and public trust remain insufficiently explored. The BTT framework proposes that ethical legitimacy in AI-based foresight arises only when bias mitigation, transparent communication, and stakeholder trust function in equilibrium. Drawing from theories of data justice, anticipatory governance, and communicative ethics, the paper conceptualizes AI as an ethical infrastructure, a moral architecture that shapes how societies envision and act upon the future. Through normative reasoning and conceptual synthesis, the study introduces the concept of predictive legitimacy, emphasizing the moral credibility of anticipatory decision-making. It argues that democratizing predictive governance requires participatory transparency, co-created data practices, and iterative trust-building mechanisms. The framework thus bridges technical and social ethics, offering policymakers and humanitarian agencies a relational model for responsible innovation. Ultimately, this study positions AI not merely as a forecasting tool, but as a vehicle for collective resilience and justice in global hunger governance.

Keywords. Predictive Legitimacy; Bias-Transparency-Trust (BTT) Framework; Artificial Intelligence Ethics; Food Security Governance; Data Justice and Anticipatory Governance

Chapter 1: Introduction

Artificial intelligence (AI) has become a cornerstone of twenty-first-century humanitarian practice. Agencies such as the **Food and Agriculture Organization (FAO)** and the **World Food Programme (WFP)** now deploy machine-learning models to forecast hunger crises, optimize logistics, and guide anticipatory funding mechanisms. Platforms like **HungerMap LIVE**, **FEWS NET**, and the **Integrated Food Security Phase Classification (IPC)** transform satellite data, mobile-survey indicators, and market prices into real-time food-insecurity maps. These systems promise unprecedented speed and foresight in managing hunger emergencies.

Yet they also raise profound **ethical, epistemic, and political questions** about whose knowledge is represented, who controls algorithmic decisions, and how predictive authority interacts with local sovereignty.

AI-driven humanitarian forecasting thus embodies a paradox. On one hand, it reflects scientific progress and innovation; on the other, it risks deepening structural inequities in global knowledge production. Machine-learning models depend on data quality, coverage, and labeling. Because most developing countries experience chronic data scarcity-especially in informal economies and conflict zones-predictive systems often extrapolate from incomplete or biased datasets. The very populations most vulnerable to hunger may therefore be the least visible to the algorithms designed to protect them. This produces what scholars term **algorithmic bias**-systematic distortions arising from unbalanced or unrepresentative data inputs (Mehrabi et al., 2021).

Beyond technical accuracy, predictive humanitarian systems influence power structures. Early-warning alerts can trigger donor mobilization, shape national policy, and direct humanitarian flows. If the models generating those alerts are opaque or exclusive, they risk replacing one form of inequality (material scarcity) with another (informational dependency). As **Guston (2014)** argues, anticipatory systems enable institutions to “govern through the future,” concentrating decision-making power in expert and donor networks. Similarly, Eubanks (2018) shows how algorithmic prediction can displace local knowledge, producing new forms of dependency and exclusion.

• 1.1 Research Gap, Purpose, and Contribution

While the technical dimensions of AI in food-security forecasting are increasingly studied, the **ethical and governance dimensions remain under-theorized**. FAO and WFP reports frequently call for “responsible AI,” yet there is limited conceptual guidance on how to operationalize fairness, transparency, and trust in contexts marked by extreme inequality. Existing frameworks-such as **Responsible AI** (Dignum, 2020) and **Data Justice** (Heeks & Renken, 2022)-offer valuable principles but lack an integrated structure for analyzing predictive humanitarian systems.

Most existing research implicitly assumes that predictive accuracy translates directly into ethical legitimacy. This conflation overlooks the fact that **trust in AI forecasts is socially constructed**, depending not only on technical performance but also on transparency, participation, and perceived fairness. Without meaningful community involvement, even a statistically robust model may fail to generate legitimacy or effective action.

Thus, a critical gap emerges between **technical optimization** and **ethical optimization**. Bridging this divide requires a theoretical framework that unites three dimensions often treated separately:

- **Bias** - data and algorithmic distortions that exclude or misrepresent populations;
- **Transparency** - the clarity, interpretability, and communicability of predictive systems;
- **Trust** - the relational confidence linking institutions, decision-makers, and affected communities.

This study introduces the **Bias–Transparency–Trust (BTT)** framework as a **normative model** for understanding and improving AI-based hunger prediction systems. The BTT framework reframes the ethics of predictive governance as a **dynamic equilibrium** among technical integrity, epistemic inclusion, and social legitimacy.

The central proposition advanced here is that AI in humanitarian forecasting **cannot be ethically neutral**. Its legitimacy depends on whether it reduces or amplifies structural

asymmetries in knowledge and power. The BTT model provides an analytical lens through which researchers and policymakers can evaluate how algorithmic systems may evolve from **technocratic prediction** toward **participatory anticipation**-where communities help define and interpret the future, rather than being passively forecasted by it.

The study addresses the overarching question:

Can algorithmic early-warning systems for food insecurity empower local resilience, or do they reinforce donor-centric biases through opaque predictive authority?

To explore this question, the analysis employs **conceptual synthesis**, drawing on theories from AI ethics, data justice, and development governance. Rather than relying on empirical testing, the study develops **theoretical propositions** explaining how transparency mediates the relationship between bias and trust, shaping the ethical legitimacy of predictive systems.

Developing countries stand at the frontline of climate shocks, market volatility, and political instability-factors that intensify hunger vulnerability. As humanitarian agencies increasingly depend on AI to anticipate crises, the governance of predictive systems becomes both a technical and moral imperative. The **BTT framework** contributes by:

- Clarifying ethical priorities through a distinction between technical bias and structural injustice;
- Providing analytical criteria for evaluating transparency mechanisms in complex AI ecosystems; and
- Offering a conceptual bridge between global predictive governance and local knowledge systems.

In doing so, the study extends the literature on **anticipatory governance** (Boyd & Holmes, 2015), proposing that anticipation must be **co-produced and ethically grounded**, not merely technologically enabled. Collectively, the chapters that follow articulate a coherent framework uniting **theoretical insight, ethical reasoning, and practical relevance** for AI-based food-security governance.

Chapter 2: Literature Review: AI, Ethics, and Food Security

The intersection between **artificial intelligence (AI)** and **food security** represents one of the most consequential transformations in contemporary development governance. In the past decade, global organizations have embraced data-driven prediction systems to anticipate food crises, allocate aid, and inform policy interventions. However, this technological optimism has been accompanied by a persistent set of **ethical and institutional dilemmas**. As predictive algorithms increasingly influence humanitarian decision-making, scholars and practitioners alike have called for frameworks that integrate *ethical reasoning, transparency, and social legitimacy* into technical models (Floridi, 2019; Dignum, 2020).

This literature review surveys the principal intellectual streams that inform the theoretical foundations of the current study. It highlights three overlapping domains: **AI ethics and responsible innovation**, focusing on bias, explainability, and moral accountability; **Data justice and epistemic inequality**, addressing the structural asymmetries in digital knowledge production; and **Humanitarian governance and anticipatory systems**, exploring how predictive models reshape decision-making and sovereignty in developing contexts. By synthesizing these perspectives, this chapter identifies the conceptual void the **BTT framework** seeks to fill: the absence of an integrated ethical model for predictive food-security governance.

- **2.2 Artificial Intelligence and Ethical Governance**

The field of **AI ethics** has evolved rapidly since the mid-2010s, propelled by growing public concern over algorithmic decision-making and its societal impacts. Foundational contributions from scholars such as **Luciano Floridi** and **Virginia Dignum** emphasize that AI systems should be developed under the principles of *beneficence, non-maleficence, autonomy, justice, and explicability*-collectively known as the “**AI4People**” framework (Floridi et al., 2018). Within humanitarian settings, these principles translate into obligations to ensure that predictive models neither harm marginalized populations nor reinforce systemic exclusion.

A critical subfield within AI ethics examines **algorithmic bias** - systematic distortions that produce unfair or inaccurate outcomes (Mehrabi et al., 2021). In machine-learning contexts, bias may emerge from unrepresentative training data, flawed feature selection, or unequal weighting of variables. In food-security forecasting, such biases often stem from **geographical and socio-economic gaps** in data collection. Studies show that regions with weak statistical infrastructure or informal agricultural markets are routinely underrepresented in predictive models, leading to distorted assessments of food stress (Jerven, 2013; Watmough et al., 2019). This exclusion reflects a deeper pattern in remote sensing and global forecasting systems, which often fail to capture community-level coping mechanisms or informal trade networks. Consequently, AI systems may overestimate stability in data-poor areas while amplifying crisis alerts in data-rich regions - a form of epistemic distortion with direct humanitarian consequences.

Recent literature extends the ethical discussion to **explainability and transparency**. As **Jobin, Ienca, and Vayena (2019)** show in their review of over 80 global AI ethics guidelines, transparency consistently ranks among the top three principles across institutions. Yet, operationalizing transparency remains challenging. In the context of complex neural networks, full technical transparency may be infeasible or counterproductive, requiring instead *interpretability* and *communicability*-ensuring that stakeholders understand how and why predictions are generated.

Humanitarian agencies have begun to acknowledge these concerns. The **World Food Programme (WFP)**, for example, has launched initiatives under its *Innovation Accelerator* to enhance the explainability of its **HungerMap LIVE** platform (WFP, 2023). However, as scholars argue, transparency alone does not guarantee **trust**. Without participatory validation and contextual understanding, transparency risks becoming performative - an exercise in disclosure rather than genuine accountability (Aizenberg & van den Hoven, 2020).

- **2.3 Data Justice and Epistemic Inequality**

Beyond technical ethics, the literature on **data justice** expands the discussion toward the structural dimensions of inequality embedded in digital systems. **Heeks and Renken (2022)** define data justice as the pursuit of fairness in how data is collected, processed, and used, emphasizing that informational inequities often mirror - and reinforce - existing social hierarchies.

In the context of developing countries, **epistemic inequality** - the unequal ability to produce and validate knowledge-has profound implications for AI governance. Predictive systems often depend on data sourced from global organizations, remote sensing, or commercial providers, rather than from local communities or informal economies. This pattern generates what **Taylor (2017)** calls “data colonialism,” in which peripheral actors remain subjects rather than co-producers of data-driven decisions.

Empirical research supports these critiques. Jones et al. (2013) argue that global food-security frameworks such as **Famine Early Warning Systems Network (FEWS NET)** often

marginalize local indicators and coping mechanisms, leading to incomplete or distorted assessments. Similarly, HungerMap LIVE integrates machine-learning outputs with household-survey and satellite data but rarely incorporates qualitative insights from local actors, limiting its contextual sensitivity (WFP, 2023). From a data-justice perspective, “more data” does not automatically translate into “better governance” (Taylor, 2017); rather, ethical foresight demands epistemic pluralism and participatory infrastructures. Within this stream, the BTT model’s emphasis on transparency and trust serves as a corrective: transparency enables epistemic inclusion, while trust provides the social validation required for equitable data ecosystems.

- **2.4 Humanitarian Governance and Predictive Systems**

Parallel to AI ethics and data justice, a growing literature on humanitarian governance examines how predictive technologies reshape institutional authority. Scholars such as **Chaves-González et al. (2022)** describe the rise of *anticipatory action*—a form of governance that seeks to manage crises **before** they fully materialize. In this paradigm, humanitarian agencies rely on predictive analytics and early triggers to allocate resources in anticipation of disasters, effectively governing “through uncertainty.” While such systems promise efficiency and responsiveness, they also risk reinforcing **donor-centric hierarchies**. As **Martin-Shields and Kudejira (2021)** argue, global institutions often retain control over model design, data governance, and validation processes, leaving local governments and communities as passive recipients of forecasts. This dynamic produces a new form of **algorithmic dependency**, in which local actors defer to AI-generated outputs without the capacity to interpret, question, or adapt them to local realities (Duffield, 2016).

When AI predictions drive aid allocation, errors raise profound questions of moral accountability. If a model anticipates famine that never occurs, who is responsible for misallocated resources? Conversely, if a crisis unfolds without warning, does responsibility lie with the model, the operator, or the donor institution? As O’Neil (2016) warns, algorithmic decision-making often obscures responsibility, creating powerful systems that govern without accountability. The literature reveals two competing tendencies. On one side lies the **technocratic narrative** of optimization and speed; on the other, the **participatory narrative** of inclusion and legitimacy. The challenge lies in reconciling these paradigms through an integrative theoretical model that recognizes both the efficiency imperatives of humanitarian response and the moral imperatives of democratic governance.

- **2.5 Integrative Insights and Theoretical Gap**

The convergence of these three literatures - AI ethics, data justice, and humanitarian governance - reveals a shared yet unfulfilled aspiration: the alignment of **technical precision** with **ethical legitimacy**. Each field provides partial solutions: AI ethics defines moral principles, data justice critiques power asymmetries, and governance studies examine institutional behavior. Yet, none offers a **composite framework** that simultaneously explains how bias, transparency, and trust interact to shape the ethical outcomes of predictive humanitarian systems. This theoretical vacuum motivates the **BTT** framework developed in this study. BTT does not introduce new ethical values but reorganizes existing ones into a **relational model**. It posits that:

Bias undermines legitimacy when unchecked.

Transparency mediates between algorithmic opacity and social comprehension.

Trust emerges not automatically but through sustained institutional communication and participatory validation.

The BTT framework thus moves beyond normative checklists toward **systemic ethics**-understanding AI not as a discrete tool but as a socio-technical assemblage embedded in structures of power, culture, and governance.

Chapter 3: The Bias-Transparency-Trust (BTT) Conceptual Framework

This chapter introduces the **BTT** framework - the central conceptual contribution of this study. The framework provides a normative model for understanding how **ethical legitimacy** can be achieved in artificial intelligence (AI)-based hunger prediction systems.

While AI offers unprecedented precision for anticipating food crises, it also raises critical ethical questions: Who defines risk? Whose data count as knowledge? Who trusts whom in algorithmic decision-making?

The BTT model proposes that the ethical credibility of predictive food-security systems emerges only when **bias, transparency, and trust** coexist in dynamic balance. Rather than treating these as isolated ethical principles, the framework conceptualizes them as *interdependent variables* that shape the legitimacy of AI-driven foresight.

- **3.1 Conceptual Foundations, Core Constructs and definitions**

The BTT framework builds upon and integrates insights from three major intellectual traditions:

AI Ethics and Responsible Innovation - Scholars such as **Floridi (2019)** and **Dignum (2020)** argue that AI systems must be “responsible by design.” This means embedding moral reasoning in technical architecture. However, their models often remain principle-based - listing fairness, transparency, and accountability as separate values. BTT moves beyond this checklist approach to explore *how* these principles interact dynamically.

Data Justice and Epistemic Inclusion - Theories of **data justice** (Heeks & Renken, 2022) emphasize that algorithmic systems can reproduce structural inequalities if they rely on biased datasets. BTT extends this notion by linking *data fairness* to *social trust*: justice in representation enhances legitimacy only when the process is transparent and participatory.

Communicative and Relational Ethics - Drawing from **Habermas (1984)** and **O’Neill (2002)**, BTT treats transparency and trust as communicative acts. Ethical governance requires deliberation - an open dialogue where affected actors can question and interpret algorithmic outputs.

By synthesizing these traditions, the BTT framework establishes a **relational ethics of anticipation**: legitimacy is not granted by compliance, but *earned through dialogue and co-production*.

The framework defines three interdependent constructs:

Bias - Bias refers to systematic distortions in how data are collected, processed, and interpreted. In predictive food-security systems, bias arises from: Geographic unevenness in data coverage (urban vs. rural); Overrepresentation of formal economic sectors; Cultural and linguistic exclusion in data labeling. Ethically, bias is not simply an error; it reflects *epistemic inequality* - whose knowledge is considered valid. The normative goal is **epistemic justice**: ensuring that marginalized data sources (informal economies, indigenous knowledge, local markets) inform predictive modeling.

Transparency - Transparency signifies more than technical explainability; it embodies *communicative openness*. It is achieved when algorithms and data processes are intelligible, contestable, and inclusive. BTT distinguishes two levels:

- **Procedural Transparency** – public access to model documentation and assumptions.
- **Dialogic Transparency** – active interpretation through participatory review and feedback.

Transparency thus functions as the *mediating mechanism* between bias reduction and trust formation.

- **Trust** - Trust is the relational confidence that connects data producers, decision-makers, and affected communities. It is neither automatic nor purely cognitive - it is socially constructed through repeated experiences of transparency and responsiveness. In the humanitarian context, trust arises when AI tools respect local expertise and avoid paternalism. Conversely, opacity and bias foster skepticism, dependency, or outright rejection.

• 3.2 Conceptual Relationships: The Equilibrium Model

The BTT relationship can be visualized as an **equilibrium triangle** (Figure 1). Each side of the triangle represents a feedback mechanism that sustains or destabilizes ethical legitimacy.

Figure 1 The Bias–Transparency–Trust (BTT) Model

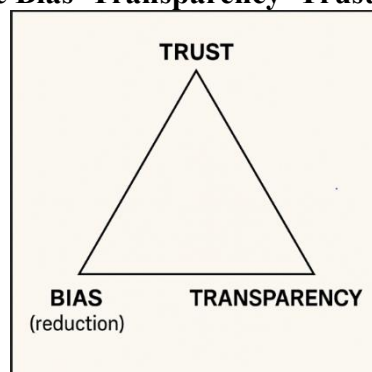


Table 1 summarizes how each dimension of the Bias–Transparency–Trust (BTT) model reinforces the others. Ethical legitimacy emerges when these feedback loops operate in balance, sustaining participatory and transparent governance in predictive food-security systems.

Table 1 Interdependent Relationships among Bias, Transparency,

Relationship	Mechanism	Ethical Function
Bias → Transparency	Reducing bias clarifies data provenance and model logic.	Enhances openness and accountability.
Transparency → Trust	Communicative clarity allows stakeholders to understand and engage.	Builds confidence and participatory legitimacy.

Trust → Bias	Inclusive trust encourages broader participation in data collection.	Expands diversity, reducing systemic distortion.
---------------------	--	--

and Trust in the BTT Framework

When these feedback loops are balanced, **ethical legitimacy** is achieved - predictive systems are both technically sound and socially credible. If one component weakens, the equilibrium collapses:

- Low transparency → technocratic opacity.
- High transparency but low trust → epistemic resistance.
- Persistent bias → donor dominance or algorithmic dependency.

Thus, legitimacy is not a static state but a **dynamic process** of continuous recalibration.

3.4 Predictive Legitimacy and the BTT Framework

From the relational equilibrium of bias, transparency, and trust arises the concept of **predictive legitimacy**-the moral credibility of anticipatory systems. Whereas traditional theories of legitimacy, whether **Weberian** or **institutional**, emphasize retrospective justification and compliance with established authority, predictive legitimacy concerns **future-oriented governance**: how systems acting on uncertain futures sustain ethical credibility and public trust.

A predictive system achieves legitimacy when its **data foundations are inclusive and fair**, its **processes are transparent and communicatively open**, and its **outputs are collaboratively interpreted and trusted** by the communities they affect. Predictive legitimacy therefore reframes AI ethics from a retrospective to a prospective paradigm-shifting attention from *how decisions were made* to *how futures are anticipated together*.

To demonstrate the practical significance of this concept, consider the **World Food Programme's HungerMap LIVE** and the **FAO's Agricultural Stress Index System (ASIS)**. Both rely heavily on satellite imagery, climatic data, and market indices, yet they frequently **under-represent informal and non-market food systems**, introducing structural bias. Their outputs are typically presented through technical dashboards that limit accessibility for local stakeholders, weakening transparency. Consequently, local actors may distrust forecasts that contradict lived experience, eroding trust and responsiveness.

Applying the **BTT framework**, such predictive systems could evolve by:

- conducting **bias audits** to reveal under-represented regions and populations;
- developing **explainability layers** that narrate model logic in local languages; and
- organizing **participatory validation workshops** where communities and national agencies jointly interpret forecasts.

Through these mechanisms, AI becomes a facilitator of **co-created foresight** rather than an instrument of **technocratic prediction**. While numerous ethical frameworks address artificial intelligence, few examine the **interactive dynamics** among principles or their implications for **predictive legitimacy** in humanitarian governance. The BTT model distinguishes itself by emphasizing **relational equilibrium** rather than static compliance, integrating **AI ethics, data justice, and anticipatory governance** into a unified conceptual architecture that is analytically rigorous and contextually adaptable.

Table 2 Comparative Overview of Ethical Frameworks Relevant to AI-Based Food-Security Systems

Framework	Focus	Limitation	BTT Contribution
AI4People (Floridi et al., 2018)	Universal ethical principles	Static and principle-based	Adds relational dynamics and contextual ethics
Responsible AI (Dignum, 2020)	Governance of AI processes	Emphasis on organizational accountability	Integrates communicative ethics and social trust
Data Justice (Heeks & Renken, 2022)	Fair data representation	Lacks anticipatory dimension	Adds foresight and participatory transparency
Anticipatory Governance (Boyd & Holmes, 2015)	Managing uncertain futures	Weak on ethical theory	Embeds explicit moral reasoning through BTT

Table 2 contrasts leading ethical paradigms with the BTT framework, highlighting its added value in embedding transparency and trust as active, relational processes within bias reduction. The **BTT framework** thus functions as a **bridge model** connecting technical ethics, social justice, and anticipatory governance. It situates AI ethics within a **relational and developmental paradigm** responsive to inequality and context. Ethical governance, in this view, cannot be imposed externally through compliance checklists but must be **cultivated internally** through iterative communication and participatory interpretation.

For theory, BTT offers a **testable conceptual nucleus** that future empirical studies can operationalize through indicators of bias, transparency, and trust. For practice, it provides humanitarian organizations with a **diagnostic tool** to assess whether:

- their data are inclusive (bias),
- their processes are communicatively open (transparency), and
- their forecasts are trusted by the communities they serve (trust).

Ethical success is measured not solely by algorithmic accuracy but by the **resilience of relationships** that sustain predictive systems. Accordingly, BTT reconceptualizes the ethics of AI in food-security governance as a **relational equilibrium**, defining bias as epistemic inequality, transparency as communicative justice, and trust as relational legitimacy. Together, these elements constitute the **moral architecture of predictive governance**, determining whether humanitarian foresight achieves genuine ethical credibility-or predictive legitimacy—in its quest to anticipate and alleviate hunger.

Chapter 4: Research Questions and Theoretical Propositions

The purpose of this chapter is to articulate the **conceptual logic** that drives the inquiry into AI-based hunger prediction systems. Since this study is theoretical, it does not advance testable hypotheses in the empirical sense; instead, it formulates **theoretical propositions** derived from the **BTT** framework. These propositions serve as normative signposts that structure analysis, offering explanations of how ethical legitimacy can emerge - or fail - in predictive food-security governance.

The following sections transform the overarching research question: *Can algorithmic early-warning systems for food insecurity empower local resilience, or do they reinforce donor-centric bias through opaque predictive authority?* into a series of interrelated sub-questions and propositions, together, these define the **ethical architecture** of predictive governance.

- **4.1 Theoretical Question and Supporting Sub-Questions**

At the center of this study lies one primary theoretical question: *How does the interaction between bias, transparency, and trust determine the ethical legitimacy of AI-based hunger prediction systems in developing countries?* This question presupposes that predictive systems are not neutral tools but *governance mechanisms* whose ethical quality depends on relational dynamics. It implies that the legitimacy of such systems cannot be measured solely by their accuracy or timeliness, but must also consider **who participates, who benefits, and who bears the risks** when predictions shape humanitarian decisions. From the main inquiry arise three subsidiary questions, each reflecting one dimension of the BTT triad:

Bias Dimension: *In what ways do data and algorithmic biases reproduce or challenge historical inequities in the governance of hunger? What ethical responsibilities accompany the recognition of bias as both a technical and moral construct?*

Transparency Dimension: *How can transparency move beyond disclosure toward communicative understanding among technical experts, policymakers, and communities? What institutional mechanisms enable meaningful transparency without compromising privacy or operational efficiency?*

Trust Dimension: *How is trust cultivated, maintained, or eroded in contexts where predictive authority is centralized in international organizations? What are the conditions under which trust translates into local ownership rather than passive compliance?*

Each sub-question targets a conceptual domain rather than an empirical variable. Together they form the foundation for the **normative reasoning** advanced through the BTT framework.

- **4.2 Theoretical Propositions**

The following propositions emerge logically from the relationships defined in Chapter 3. They serve as theoretical statements that describe how ethical outcomes are produced within predictive food-security systems.

- **Proposition 1: Algorithmic bias is both a technical flaw and an ethical signal.** Bias indicates not only inaccurate predictions but also the exclusion of marginalized perspectives. Recognizing bias as a moral indicator reframes mitigation from a purely statistical exercise to an act of epistemic inclusion. The ethical imperative, therefore, is to design data ecosystems that reflect social heterogeneity and plural knowledge forms (Taylor, 2017).

Theoretical implication: Ethical legitimacy begins with acknowledging whose experiences are omitted from data narratives. Reducing bias requires structural reforms in data governance—open data sharing, participatory validation, and local capacity building.

- **Proposition 2: Transparency mediates the relationship between bias and trust.** Transparency functions as the **ethical bridge** that connects algorithmic integrity with social legitimacy. When transparency is procedural (open algorithms, accessible datasets) and communicative (clear interpretation and explanation), it allows stakeholders to identify bias and judge institutional reliability. Conversely, when transparency is absent or performative, bias remains concealed and trust deteriorates.

Theoretical implication: Effective transparency is dialogic rather than unilateral; it transforms AI systems from black boxes into shared epistemic spaces (Habermas, 1984).

- **Proposition 3: Trust is an emergent and relational outcome.** Trust cannot be engineered; it must be **earned** through iterative interaction. It accumulates when predictive systems demonstrate fairness, responsiveness, and openness to correction. In contexts of humanitarian urgency, trust becomes the social currency that determines whether forecasts are acted upon or dismissed (O’Neill, 2002).

Theoretical implication: Sustained trust depends on continuous engagement - feedback loops, participatory scenario planning, and shared monitoring mechanisms.

- **Proposition 4: Ethical legitimacy depends on equilibrium among bias, transparency, and trust.**

A predictive system achieves legitimacy when transparency effectively mediates between bias and trust. Excessive bias erodes legitimacy even if transparency is high; transparency without trust yields compliance but not commitment; trust without transparency risks paternalism. The moral equilibrium occurs when bias is acknowledged and mitigated through transparent communication, generating informed trust (Dignum, 2020).

Theoretical implication: Governance should focus on balancing-not maximizing—each dimension to maintain ethical stability.

- **Proposition 5: Participatory anticipation enhances both transparency and trust.** When local communities participate in the interpretation and validation of predictive outputs, they co-produce knowledge rather than passively receive it. This participatory anticipation aligns with **anticipatory governance** (Boyd & Holmes, 2015), in which foresight becomes a collective rather than technocratic function.

Theoretical implication: Inclusion transforms transparency from disclosure to dialogue and converts trust from expectation into co-ownership.

- **4.3 Conceptual Model of Relationships and Analytical Reasoning**

Building on the theoretical foundations of the Bias-Transparency-Trust (BTT) framework, this section summarizes the conceptual relationships that link the core variables and their expected ethical outcomes. The purpose is to illustrate how bias reduction, transparency enhancement, and trust formation interact as a dynamic system that determines predictive legitimacy in AI-based food-security governance. Table 3 presents a synthesized model of these relationships, highlighting the directionality, interaction strength, and anticipated ethical implications derived from the theoretical propositions.

Table 3 Conceptual Relationships among Bias, Transparency, and Trust in the BTT Model

Relationship Direction	Type of Interaction	Expected Effect	Ethical	Illustrative Implication
Bias → Transparency	Causal and corrective	Reduction in informational distortion; openness	enhanced	Bias audits improve clarity and accountability in data use.
Transparency → Trust	Mediating and reinforcing	Increased confidence and participatory	stakeholder and legitimacy	Clear model communication

			fosters institutional credibility.
Trust → Bias	Feedback and generative	Broader inclusion of local data sources; mitigated epistemic inequality	Community trust leads to richer, more diverse datasets.
Bias ↔ Transparency ↔ Trust	Recursive and dynamic	Stable ethical equilibrium sustaining predictive legitimacy	Continuous recalibration among all three variables ensures long-term system credibility.

Table 3 conceptualizes the relational logic of the BTT framework. Each pathway demonstrates how ethical legitimacy in predictive AI systems arises from iterative feedback among bias reduction, transparency enhancement, and trust maintenance.

The transformation of research questions into theoretical propositions allows this study to proceed as **conceptual analysis** rather than empirical testing. Each proposition functions as a lens for evaluating real or hypothetical scenarios of AI-based food-security systems. The analysis that follows (Chapter 6) will apply these propositions to explore how ethical legitimacy can be constructed, maintained, or lost in predictive governance. The methodological chapter (Chapter 5) next explains how such conceptual reasoning is structured - detailing how analytical coherence and philosophical rigor substitute for data-driven inference in a theoretical inquiry.

Chapter 5: Analytical Methodology: Conceptual and Normative Reasoning

In empirical research, methodology often refers to the techniques used for collecting and analyzing data. In theoretical inquiry, however, methodology signifies something broader: it concerns **the logic of reasoning**, the **architecture of conceptual construction**, and the **criteria of validity** by which ideas are evaluated. This study is intentionally non-empirical. It aims not to test hypotheses through data, but to **construct and justify a conceptual framework** - the Bias-Transparency-Trust (BTT) model- that explains ethical dynamics in AI-based predictive food-security systems. The methodology adopted here draws from three intellectual traditions: **conceptual synthesis**, **comparative reasoning**, and **normative analysis**. Together, these methods ensure that the theoretical arguments are both logically coherent and practically relevant to humanitarian governance.

- **5.1 Conceptual Research Design**

Conceptual research seeks to **clarify, integrate, and extend theoretical constructs**. Rather than observing empirical phenomena, it operates through *abstraction* and *theoretical modeling*. As **Jaccard and Jacoby (2020)** note, conceptual models serve as “cognitive maps” that guide understanding by identifying relationships among ideas. The design of this study therefore follows four sequential steps:

- a. **Concept Identification** – Defining the core constructs (bias, transparency, trust) from interdisciplinary sources (AI ethics, data justice, and governance theory).
- b. **Conceptual Synthesis** – Integrating these constructs into a coherent triadic model that captures their interactions.

- c. **Normative Reasoning** – Assessing the moral implications of these relationships through ethical and philosophical argumentation.
- d. **Theoretical Application** – Illustrating the model’s utility through thought experiments and comparative reasoning across humanitarian contexts.

These steps constitute a systematic approach to theoretical construction that balances analytical rigor with normative depth.

- **5.2 Methodological Foundations**
- **Conceptual Synthesis**-The first methodological foundation is **conceptual synthesis**, a process of weaving together diverse theoretical strands into a unified framework. Here, synthesis occurs at the intersection of three disciplines:
 - **Artificial Intelligence Ethics**, which provides language for understanding fairness, accountability, and explainability (Floridi et al., 2018; Dignum, 2020).
 - **Data Justice Theory**, which situates technological systems within structures of social power and inequality (Heeks & Renken, 2022).
- **Humanitarian Governance Studies**, which examine how digital and predictive technologies reconfigure authority, accountability, and decision-making power (Madianou, 2019; Duffield, 2016). Through conceptual synthesis, the BTT model transforms dispersed ethical principles into a relational framework that links bias, transparency, and trust as interdependent dimensions of legitimacy.
 - **Comparative Reasoning**-The second foundation is **comparative reasoning** - the systematic juxtaposition of ideas, frameworks, or cases to identify patterns and contrasts. While this study does not employ empirical comparison, it uses *conceptual comparison* to test the internal coherence of the BTT model against existing frameworks.

Three comparative axes guide the reasoning:

Between frameworks: Comparing BTT with established models such as AI4People, Responsible AI, and Data Justice, highlighting where BTT extends or refines them.

Between governance paradigms: Contrasting technocratic anticipation (centralized, donor-driven forecasting) with participatory anticipation (locally co-produced foresight).

Between ethical logics: Examining how consequentialist (outcome-focused) and deontological (duty-focused) ethics intersect in predictive governance.

This comparative method ensures that BTT is not developed in isolation but as part of a critical dialogue with existing traditions.

- **Normative Reasoning**-The third foundation, **normative reasoning**, concerns moral justification - explaining why certain configurations of bias, transparency, and trust are ethically desirable. Normative reasoning follows a three-tiered logic (Rawls, 1971; O’Neill, 2002):

Descriptive Layer: Identify the ethical issue (e.g., opaque AI forecasts).

Evaluative Layer: Interpret the issue through moral principles (e.g., fairness, participation, autonomy).

Prescriptive Layer: Derive normative guidance (e.g., design participatory transparency mechanisms).

This structure enables the translation of abstract ethical principles into **actionable theoretical prescriptions**, bridging philosophy and policy. For example, when an algorithm trained primarily on high-quality urban data underrepresents rural food insecurity, normative reasoning leads to the conclusion that justice requires compensatory inclusion measures—such

as locally validated indicators or participatory data audits. Such reasoning demonstrates how ethical theory can guide policy even in the absence of quantitative metrics.

- **5.3 Analytical Logic of the Study**

The **analytical logic** of the study unfolds in three interconnected phases: conceptualization, relation mapping, and theoretical inference. Table 4 summarizes the sequential analytical logic employed in developing the BTT framework. Each phase contributes to constructing a coherent theoretical model linking bias reduction, transparency enhancement, and trust formation to predictive legitimacy.

Table 4 Analytical Phases and Outputs of the Theoretical Reasoning Process

Analytical Phase	Objective	Output
Conceptualization	Define the core constructs of Bias, Transparency, and Trust based on interdisciplinary theory.	Achieves theoretical clarity and conceptual precision.
Relation Mapping	Identify interconnections and causal directions among the three variables.	Develops a structural model of ethical equilibrium (BTT).
Theoretical Inference	Derive normative propositions and anticipated ethical implications.	Generates testable conceptual insights and theoretical propositions.

This structure allows the paper to function as both a theoretical and normative inquiry, providing **logical coherence** comparable to empirical robustness.

- **5.4 Validity, Ethical Reflexivity and Positionality**

Unlike empirical studies, where validity is established through measurement reliability, theoretical research relies on **conceptual validity** - the degree to which constructs and relationships are logically consistent, comprehensive, and plausible. Following **Dubin (1978)** and **Whetten (1989)**, this study ensures validity through:

- **Internal Consistency:** Each construct is defined precisely, and their interrelations are theoretically justified.
- **Comprehensiveness:** The model integrates technical, ethical, and social dimensions.
- **Parsimony:** Despite its breadth, the model remains simple enough to be generalizable.
- **Theoretical Utility:** The framework generates actionable insights for real-world governance.

Furthermore, the reasoning process is transparent: every conceptual link is accompanied by normative justification, enabling replication by other theorists.

An often-overlooked aspect of conceptual research is **ethical reflexivity** - awareness of the researcher's positionality in shaping theoretical constructs. This study recognizes that theories of AI and governance are not value-neutral; they emerge from particular intellectual and geopolitical standpoints. As **Duffield (2016)** argues, digital humanitarian infrastructures often enable forms of **remote governance**, where global institutions use predictive technologies to manage crises from a distance while local populations become passive recipients of data-driven decisions. Rather than fostering participation, these systems can entrench **technocratic**

dependency, reducing communities to sources of data rather than co-creators of knowledge or strategy.

To counter this structural imbalance, the methodology adopts a **reflexive posture**, positioning the researcher as an *ethical participant* rather than an external observer. This entails:

- Acknowledging the limitations of applying **Eurocentric theories** to Global South contexts;
- Drawing upon **African, Asian, and Latin American scholarship** on food systems, data sovereignty, and technology governance;
 - Emphasizing **dialogic ethics**, wherein legitimacy arises from inclusive discourse rather than prescriptive authority (Habermas, 1984).

Such reflexivity strengthens both the **credibility** and **universality** of the BTT model by rooting it in **pluralistic epistemic traditions**, rather than reinforcing top-down technocratic assumptions.

• 5.5 Methodological Advantages and Limitations

The conceptual methodology adopted in this study offers both notable strengths and inherent constraints. Its primary advantage lies in the ability to integrate diverse theoretical traditions—AI ethics, data justice, and governance theory—into a coherent analytical model without the need for empirical data.

However, as a theory-driven approach, it also faces limitations in external validation and contextual generalization, which future empirical research must address.

- **Advantages:**
 - Enables theoretical innovation without dependence on incomplete datasets.
 - Facilitates interdisciplinary synthesis across ethics, AI, and governance.
 - Provides conceptual precision that can guide empirical research in the future.
- **Limitations:**
 - Absence of quantitative validation may limit immediate policy adoption.
 - Abstract reasoning requires careful interpretation when applied to real-world systems.
 - Risk of normative idealization—concepts may appear achievable in theory but remain difficult in practice.

However, these limitations are intrinsic to conceptual work and do not diminish its value. As **Hirschman (1970)** argued, the purpose of theory is to “open new spaces for thought,” offering intellectual scaffolding upon which empirical research can later build.

Table 5 summarizes the main methodological strengths and weaknesses of this conceptual study. While the approach advances theoretical innovation, future empirical research will be essential to validate and contextualize its propositions.

Table 5 Methodological Advantages and Limitations of the Study

Dimension	Advantages	Limitations
Conceptual Scope	Integrates multiple theoretical traditions (AI ethics, data justice, governance).	Lacks empirical testing or statistical verification.
Analytical Flexibility	Enables normative reasoning and interdisciplinary synthesis.	Generalizations may overlook contextual specificities.

Ethical Insight	Illuminates moral structures underpinning predictive governance.	Difficult to quantify ethical dynamics or outcomes.
Policy Relevance	Offers transferable conceptual tools for humanitarian organizations.	Implementation requires adaptation to institutional realities.

In summary, the conceptual approach adopted here prioritizes analytical coherence over empirical generalization, aiming to build a structured ethical lens rather than test hypotheses.

By clarifying the logical relationships among bias, transparency, and trust, the methodology establishes a solid foundation for the theoretical exploration that follows. The next chapter applies this framework analytically to derive expected ethical configurations and institutional insights within the **BTT model**.

Chapter 6: Theoretical Analysis and Expected Insights

Building on the conceptual foundations and methodological reasoning presented earlier, this chapter applies the **BTT** framework to analyze the ethical dynamics of predictive food-security systems. It interprets how variations in bias reduction, transparency enhancement, and trust formation generate distinct ethical configurations and institutional outcomes. Through theoretical reasoning, the chapter identifies the **expected insights** of the model-how ethical equilibrium can be sustained, disrupted, or restored-and develops a typology of governance scenarios that illustrate the moral architecture of predictive legitimacy.

- **6.1 Theoretical Scenario Logic**

The **Theoretical Scenario Logic** outlines how interactions among bias, transparency, and trust produce varying ethical configurations within the **BTT framework**. Rather than treating these variables as static principles, this logic conceptualizes them as dynamic forces that combine to shape the moral legitimacy of predictive systems. By analyzing different permutations of equilibrium and distortion, this section develops a typology of scenarios-ranging from *inclusive foresight* to *technocratic opacity*-that illustrates how ethical balance determines the credibility and justice of AI-driven food-security governance.

To analyze how BTT operates, we can imagine four theoretical scenarios-ideal types that illustrate possible interactions among the three dimensions. Each scenario represents an ethical configuration within predictive governance. The interaction among bias, transparency, and trust produces four ethical configurations of predictive governance. These theoretical scenarios are summarized in Table 6.

The interaction among the three ethical dimensions-bias, transparency, and trust-creates distinct moral configurations that define how predictive systems function within institutional contexts. Each configuration represents a particular ethical equilibrium or imbalance, revealing how shifts in one dimension can either reinforce or undermine overall legitimacy. The following table summarizes these **theoretical scenarios**, outlining the characteristic patterns and their corresponding ethical implications within the BTT framework.

Table 6 Ethical Configurations and Expected Outcomes in the BTT Framework

Scenario	Bias Level	Transparency Level	Trust Level	Expected Outcome	Ethical
----------	------------	--------------------	-------------	------------------	---------

1. Inclusive Foresight	Low	High	High	Ethical equilibrium: participatory legitimacy
2. Technocratic Opacity	Low–Moderate	Low	Low	Efficiency without legitimacy
3. Algorithmic Dependency	High	Moderate	Moderate	Fragile trust, donor dominance
4. Epistemic Resistance	Moderate	High	Low	Transparency mismatch; local skepticism

Table 6 highlights how each configuration represents a distinct ethical equilibrium within the BTT model, illustrating the continuum between participatory legitimacy and technocratic control. This typology provides a conceptual map of how predictive systems can succeed or fail ethically. Each type reflects a distinct theoretical pattern derived from the interaction of BTT dimensions.

- **Scenario 1: Inclusive Foresight**

Configuration: *Low Bias- High Transparency - High Trust*

Ethical Outcome: Participatory legitimacy and epistemic justice.

This ideal configuration represents **ethical equilibrium**, where predictive governance functions as a partnership between global expertise and local knowledge. Bias is minimized through diverse datasets and participatory data validation; transparency is achieved through communicative clarity; and trust is maintained through shared accountability. In this scenario, AI becomes a **facilitator of co-production**, not a replacement for human judgment. Local actors—farmers, municipal authorities, and NGOs—are involved in interpreting model outputs, grounding predictions in contextual understanding. Ethical legitimacy arises because transparency is not mere disclosure but an ongoing dialogue (Habermas, 1984).

Expected insight: Predictive legitimacy requires distributed intelligence—systems that respect both algorithmic precision and human wisdom.

- **Scenario 2: Technocratic Opacity**

Configuration: *Low–Moderate Bias - Low Transparency - Low Trust*

Ethical Outcome: Efficiency without legitimacy.

In this case, the algorithm is technically accurate but institutionally opaque. Forecasts may correctly anticipate food shortages, yet stakeholders remain unaware of how predictions are generated or validated. Decision-making authority remains concentrated in central agencies or donors, while local actors lack epistemic access to model logic. Such opacity produces what Miller (2019) calls “**governing by numbers**”—a form of abstraction in which predictive systems replace relationships with algorithmic metrics. As a result, trust remains weak: communities may comply with alerts, but they do not internalize them as shared or legitimate knowledge (Eubanks, 2018).

Expected insight: Technocratic opacity produces operational efficiency but undermines ethical legitimacy, creating fragile governance that depends on compliance rather than confidence.

- **Scenario 3: Algorithmic Dependency**

Configuration: *High Bias - Moderate Transparency - Moderate Trust*

Ethical Outcome: Fragile trust and donor dominance.

Here, transparency mechanisms exist (e.g., dashboards, open datasets), but underlying biases persist due to structural inequality in data sources. AI predictions appear credible to donors yet remain disconnected from local realities. Over time, institutional actors develop **algorithmic dependency** - relying on machine forecasts even when contradictory local evidence arises (O’Neil, 2016)). Trust, though present, becomes **misplaced trust** - a form of epistemic deference that replaces accountability with automation. Donor agencies may cite model results to justify decisions, effectively displacing human responsibility.

Expected insight: When transparency exposes process but not power, it legitimizes bias instead of correcting it. True transparency must reveal not only *how* AI works but also *who* defines its parameters

- **Scenario 4: Epistemic Resistance**

Configuration: *Moderate Bias - High Transparency - Low Trust*

Ethical Outcome: Transparency mismatch and local skepticism.

In some cases, international agencies may implement strong transparency mechanisms—publishing algorithms, sharing data, and inviting local review—yet still face **trust deficits**. This occurs when transparency reveals discrepancies between external models and local epistemologies. For example, predictive alerts may contradict farmers’ lived experiences, leading to resistance or indifference (Taylor, 2017). This configuration demonstrates that transparency, though necessary, is insufficient without cultural translation and participatory interpretation. Ethical legitimacy depends not only on what information is shared but also on how it is *understood* and *trusted*.

Expected insight: Transparency without contextual resonance leads to epistemic alienation—stakeholders see the system but do not see themselves within it.

- **6.2 The Ethical Dynamics of BTT and Theoretical Implications**

The BTT model thus describes a **dynamic moral economy** in predictive governance. Each dimension influences and is influenced by the others: **Bias Reduction** improves factual accuracy, which enhances transparency by clarifying causal logic; **Transparency** enables stakeholders to identify and question bias, promoting iterative correction; **Trust** motivates continued participation, sustaining feedback loops that further reduce bias.

However, these positive feedback loops depend on **institutional will** and **social capital**. If any component fails—if bias is concealed, transparency becomes performative, or trust erodes—the equilibrium collapses. Predictive governance then reverts to technocratic or paternalistic modes, where decisions are made *for* rather than *with* communities.

Expected insight: Ethical AI governance is a recursive process, not a fixed state. Its legitimacy must be continuously reproduced through openness, dialogue, and responsiveness.

The theoretical implications of the BTT model extend beyond food-security systems to the broader field of AI governance in development. Three comparative insights stand out:

From Compliance to Co-Production: Traditional governance frameworks emphasize compliance with ethical standards. BTT reframes ethics as co-production—where legitimacy arises from collective meaning-making rather than top-down regulation (Dignum, 2020).

From Metrics to Meaning: Conventional AI ethics relies on measurable indicators such as accuracy or fairness scores. BTT introduces interpretive dimensions—dialogue, understanding, and trust - that transcend quantitative evaluation (Floridi, 2019).

From Technocracy to Relational Governance: While most predictive systems are designed to optimize performance, BTT centers *relational ethics* - how algorithms mediate relationships between donors, states, and citizens.

In each case, the BTT model transforms ethical reasoning from a static code into a dynamic practice grounded in human relationships and institutional transparency. The implementation of the **BTT** framework within humanitarian agencies is expected to generate several institutional transformations. These changes reflect a shift from compliance-oriented structures toward adaptive, dialogue-driven, and ethically reflexive governance. The anticipated ethical impacts.

Table 7 summarizes the primary transformations, their operational descriptions, and the anticipated ethical impacts.

Table 7 Expected Institutional Transformations through the Implementation of the BTT Framework

Transformation	Description	Ethical Impact
From secrecy to structured openness	Replace ad-hoc disclosure with participatory transparency mechanisms.	Builds epistemic legitimacy and fosters public accountability.
From algorithmic dependence to interpretive partnership	Create local advisory panels to interpret predictions collaboratively.	Enhances contextual accuracy and inclusiveness.
From donor accountability to shared accountability	Shift reporting metrics from donor satisfaction to community validation.	Reinforces relational trust and participatory legitimacy.
From static ethics to iterative learning	Institutionalize ethics reviews at every update cycle.	Ensures continuous moral calibration and adaptive governance.

The anticipated ethical impacts.

Table 7 outlines the expected institutional shifts required to operationalize the BTT framework in humanitarian settings. Each transformation embodies a move toward participatory transparency, interpretive collaboration, and sustained ethical learning.

Expected Insight: Institutional ethics must evolve from **compliance-oriented** to **dialogue-oriented** cultures, embedding transparency and trust as enduring organizational practices rather than symbolic public-relations gestures. The BTT framework thus repositions ethics as an *infrastructure of governance* - a continuous process of reflection, participation, and shared responsibility.

The analysis of scenarios and mechanisms yields several key theoretical insights:

Equilibrium Principle: Ethical legitimacy emerges only when bias, transparency, and trust exist in balance.

Dialogic Principle: Transparency achieves ethical power only through communicative interaction, not mere disclosure.

Reflexive Principle: Trust is iterative and contingent; it must be nurtured through institutional humility and responsiveness.

Justice Principle: Reducing bias is not only a technical task but a form of epistemic justice—recognizing marginalized knowledges as legitimate data sources.

Transformation Principle: AI governance must evolve from technocratic control toward participatory co-production, embedding ethics in daily institutional routines.

These insights articulate a **theoretical pathway** toward ethically resilient AI systems for predictive food security. They also set the foundation for the next chapter, which explores how the BTT model contributes to broader policy and theoretical debates.

Chapter 7: Integrating Ethical Theory and Policy: The Contributions of the BTT Framework

Artificial intelligence (AI) is increasingly shaping humanitarian decision-making, influencing how agencies forecast crises, allocate resources, and define vulnerability. Yet as predictive systems gain authority, questions arise about **who defines hunger, whose data are trusted, and whose knowledge is excluded**. This chapter positions the BTT framework as a bridge between ethical theory and policy practice. It demonstrates how the framework can guide institutional change, inform international law, and advance theoretical understanding of AI governance.

The starting point is the recognition that **humanitarian AI systems operate within global asymmetries of power**. Major agencies such as the FAO, WFP, and World Bank often depend on donor-driven data, proprietary algorithms, and limited local participation. Ethical governance therefore requires more than procedural compliance - it demands **relational transformation**.

The BTT framework responds by treating ethics as a *living infrastructure* that connects algorithmic fairness, institutional transparency, and social trust. In doing so, it translates moral principles into actionable mechanisms for inclusive governance.

At the **policy level**, BTT provides a coherent ethical architecture that aligns with emerging global instruments for AI governance and digital rights. The **OECD AI Principles (2019)** emphasize transparency, fairness, and accountability; BTT operationalizes these through relational ethics, replacing static codes with continuous dialogue among data producers, policymakers, and communities. The **UNESCO Recommendation (2021)** calls for AI that “benefits humanity and the environment.” BTT gives this call practical meaning by embedding ethical foresight into data and decision systems. Within the **FAO Digital Council** and **WFP Data Responsibility Policy**, the framework can function as an ethical monitoring tool that balances predictive efficiency with human-rights obligations. By aligning with such instruments, the BTT model transforms theoretical ethics into a **governance toolkit** for humanitarian policy. To make these principles operational, institutional structures must adapt across multiple levels - from international organizations to local communities.

Table 8 Institutional Levels of Application of the BTT Framework

Institutional Level	Application Focus	Key Policy Actions	Ethical Outcomes
Global (FAO, WFP, UNDP)	Governance architecture and inter-agency accountability	Establish an <i>Ethical AI Charter</i> ; create cross-agency data-ethics committees; align predictive models with	Strengthened global coherence and legitimacy.

		UNESCO and OECD principles.	
National (Ministries, Statistical Offices)	Policy integration and ethical adaptation	Develop national predictive-ethics guidelines; implement transparency dashboards; conduct periodic algorithmic bias audits.	Increased local ownership and adaptive transparency.
Community (NGOs, Local Councils)	Participatory validation and grassroots feedback	Create citizen review panels; co-design interpretive forecasts; translate alerts into local languages and cultural idioms.	Deepened trust, inclusivity, and epistemic justice.

Table 8 outlines how the BTT framework can be operationalized at three institutional tiers - global, national, and community - to achieve balanced ethical legitimacy in AI-based food-security governance. As Table 8 illustrates, embedding BTT across institutional scales transforms ethical governance from a top-down mandate into a participatory ecosystem. Global agencies gain coherence, national systems enhance accountability, and communities become co-producers of foresight rather than passive recipients of prediction.

Implementing BTT thus moves organizations:

- **from secrecy to structured openness**, through participatory transparency mechanisms;
- **from algorithmic dependence to interpretive partnership**, by involving local advisory panels;
- **from donor accountability to shared accountability**, emphasizing community validation; and
- **from static ethics to iterative learning**, institutionalizing moral reflection at every update cycle.

Together, these transformations redefine success as *relational legitimacy* rather than procedural compliance.

Beyond immediate policy application, the BTT framework enriches **theoretical understanding** of ethical AI. It introduces a **relational model of legitimacy**, showing that ethical credibility arises from dynamic interaction among bias reduction, transparency, and trust — not from isolated principles. Drawing on **systems theory** (Meadows, 2008) and **communicative ethics** (Habermas, 1984), it reframes the ethical question from “*What principles should we follow?*” to “*How do principles coexist and reinforce one another in practice?*” The framework also bridges **technical and social ethics**, connecting algorithmic precision with participatory accountability. Transparency acts as the mediating force linking technical fairness to social legitimacy. In uniting insights from **data justice** (Heeks & Renken, 2022) and **anticipatory governance** (Boyd & Holmes, 2023 2015), BTT establishes that ethical foresight requires just data, and just data require transparent foresight. Trust becomes a **process of ethical co-production** rather than a static trait. This interpretation deepens organizational and applied-ethics theory (Lewicki et al., 1998), portraying trust as relational, evolving, and essential to predictive legitimacy. Most distinctively, the BTT framework articulates **predictive legitimacy** - the moral credibility of systems that act on the future. It extends classical legitimacy theory beyond retrospective justification (“why we acted”) to anticipatory ethics (“how we foresee responsibly”). Predictive legitimacy depends on whether predictive

algorithms empower collective resilience or reinforce dependency. Within food-security forecasting, legitimacy is achieved only when predictions are **transparent, inclusive, and dialogically validated**.

The framework also outlines **future research directions**;

- Empirical studies could test measurable indicators of bias, transparency, and trust across contexts.
- Cross-sectoral research could adapt the model to health, climate, or disaster-risk governance.
- Philosophically, further work could refine predictive legitimacy through the lenses of epistemic responsibility and collective ethics.
- Development organizations might employ BTT conceptually to design ethical-evaluation tools, transparency-training curricula, and adaptive-governance prototypes - illustrating how theoretical ethics can inform practice while maintaining normative depth.

In conclusion, the **BTT framework** contributes simultaneously to policy and theory. On the policy side, it provides actionable methods for ethical foresight - bias audits, deliberative transparency, and participatory trust-building. Theoretically, it introduces **predictive legitimacy** as a unifying concept linking ethics, governance, and future-oriented decision-making. By framing AI not merely as a technical instrument but as a **moral infrastructure**, the BTT framework transforms predictive ethics from a reactive posture into a **proactive system of justice** - aligning foresight with fairness and grounding innovation in shared trust and accountability.

Chapter 9: Conclusion

This study has developed a **theoretical and normative framework** for understanding the ethical dimensions of AI-based hunger prediction in developing countries. At its core lies the **BTT** model, which redefines ethical AI not as a static checklist of principles, but as a *relational equilibrium* among three interdependent dimensions:

Bias, representing the technical and epistemic distortions that emerge from unbalanced data or inequitable knowledge systems;

Transparency, embodying the communicative openness through which predictive systems become understandable, contestable, and accountable;

Trust, the relational confidence that sustains collective belief in predictive systems and their legitimacy.

Through this triadic model, the paper has argued that the **ethical legitimacy** of predictive food-security systems emerges only when these three elements coexist in balanced interaction. Bias reduction enhances factual credibility; transparency converts technical processes into shared understanding; and trust transforms knowledge into social cooperation. The central proposition advanced throughout this research is therefore that:

Predictive AI systems achieve moral legitimacy when transparency effectively mediates between algorithmic bias and stakeholder trust, allowing foresight to become participatory rather than technocratic. This proposition provides both an **analytical lens** for academic inquiry and a **normative compass** for policymakers and humanitarian organizations.

- **9.1 Theoretical Contributions and Key Policy Insights**

The study contributes to the growing literature on AI ethics by moving beyond rule-based or principle-based models. Instead of listing values such as fairness, accountability, and transparency as discrete goals, the BTT model **maps the relationships among them**, revealing how ethical outcomes depend on their balance. This relational approach enriches both moral philosophy and systems theory, showing that ethical AI is not achieved through optimization but through *dynamic adjustment and dialogue*.

A key theoretical innovation introduced in this study is **predictive legitimacy** - the moral credibility of anticipatory governance systems. Unlike traditional legitimacy, which evaluates outcomes after decisions, predictive legitimacy assesses *the process of foresight itself*. AI-driven forecasts command legitimacy only when they are transparent, inclusive, and reflexive. This concept extends classical theories of legitimacy (Habermas, 1984; O’Neill, 2002) to the domain of artificial prediction and humanitarian governance.

By synthesizing **data justice** (Heeks & Renken, 2022) and **anticipatory governance** (Boyd & Holmes, 2015), the study unites two previously separate literatures. Data justice focuses on fairness in information systems, while anticipatory governance explores institutional preparedness for uncertain futures. The BTT framework bridges these by demonstrating that ethical anticipation requires just data practices, and that just data practices require anticipatory transparency.

Trust is reconceptualized here as *co-produced*, not bestowed. The BTT model posits that trust arises through iterative transparency and participatory communication. This understanding contributes to a more social and institutional theory of trust in AI, emphasizing human relationships rather than technological reliability alone.

The theoretical arguments yield several **policy-relevant lessons** for global organizations and governments deploying AI-based food-security models. **Bias** cannot be eliminated but must be *governed*. The BTT model encourages organizations to institutionalize bias audits, participatory data validation, and inclusion metrics. Ethical governance thus becomes proactive, not reactive—addressing distortion before it undermines legitimacy. **Transparency** must go beyond code publication or technical explainability. It should be **communicative**, enabling non-experts to understand, question, and contextualize AI-generated insights. FAO and WFP could adopt interpretive dashboards, multilingual data narratives, and community review panels as core features of transparency practice. **Trust** is earned through shared ownership of knowledge. Participatory anticipation—where communities validate and interpret predictive outputs—ensures that early-warning systems serve as *public goods* rather than *bureaucratic instruments*.

Together, these insights establish a foundation for ethical foresight-governance that anticipates crises while respecting dignity, diversity, and agency.

- **9.2 Theoretical Reflection: AI as Ethical Infrastructure**

The analysis presented in this study supports a larger philosophical claim: **AI in humanitarian governance is not merely a tool and it is an infrastructure of moral choice**. Predictive algorithms structure how societies imagine their futures, prioritize interventions, and distribute care. As such, their design is never value-neutral.

In this sense, ethics must be embedded not as external oversight but as an **intrinsic design logic**. The BTT model treats ethics as architecture: bias auditing is its foundation, transparency its structure, and trust its sustaining framework. Only when all three are harmonized can predictive systems serve as instruments of justice rather than control.

This interpretation aligns with **Floridi's (2019)** philosophy of information, which envisions ethics as a form of “conceptual design”—creating systems that embody values rather than merely conforming to them. The BTT model thus contributes to a shift from *compliance ethics* to *design ethics* in the governance of AI.

• 9.3 Limitations

No theoretical framework is without boundaries. The BTT model, by focusing on relational ethics, risks **normative idealization** - presuming that actors will collaborate in good faith. In reality, humanitarian systems are shaped by power asymmetries, donor politics, and institutional inertia. Implementation therefore requires not only ethical insight but political will.

Moreover, while the model identifies bias, transparency, and trust as universal dimensions, their manifestations are context-specific. Cultural conceptions of trust or fairness vary across societies; thus, empirical localization remains essential for practical adoption.

Finally, as a conceptual study, this work cannot claim predictive certainty; its goal is **theoretical clarity**, not empirical generalization. It offers a moral compass, not a mathematical equation.

• 9.7 Final Message

The message of this study is a transformative one: **anticipation must be democratized**. AI-driven hunger forecasting should not replace human judgment but enrich it through collective intelligence. When algorithms are transparent and participatory, they cease to be instruments of surveillance and become instruments of solidarity. The challenge for the coming decade is to ensure that **the future is co-created**, not pre-coded. The BTT model provides a pathway toward this vision—a framework through which humanitarian AI can evolve from *seeing hunger* to *understanding humanity*.

In sum, this paper argues that predictive AI in food-security governance stands at a moral crossroads. If left unchecked, algorithmic systems risk entrenching new hierarchies of data and power; yet, guided by ethics, they hold the potential to transform humanitarian foresight into an inclusive practice of justice. The **Bias-Transparency-Trust (BTT)** model offers a conceptual compass for navigating this terrain. By harmonizing technical precision with social legitimacy, it enables predictive governance that is not only intelligent but also *wise*.

The future of hunger prediction, therefore, should not be measured merely by its accuracy, but by its **ethics of anticipation**—the degree to which it reflects collective values, empowers vulnerable communities, and sustains public trust in a shared humanity.

References

1. Boyd, E., Nykvist, B., Borgström, S., & Stacewicz, I. (2015). *Anticipatory governance for social-ecological resilience*. *Ambio*, 44(Suppl 1), S149–S161. <https://centaur.reading.ac.uk/54062/1/art%253A10.1007%252Fs13280-014-0604-x.pdf>
2. Chaves-González, J., Milano, L., Omtzigt, D.-J., Pfister, D., & Poirier, J. (2022). *Anticipatory action: Lessons for the future*. *Frontiers in Climate*, 4. <https://doi.org/10.3389/fclim.2022.932336>
3. Dignum, V. (2020). *Responsible Artificial Intelligence: How to Develop and Use AI in a Responsible Way*. ResearchGate https://www.researchgate.net/publication/337014098_Responsible_Artificial_Intelligence_How_to_Develop_and_Use_AI_in_a_Responsible_Way

4. Duffield, M. (2016). The resilience of the ruins: Towards a critique of digital humanitarianism. *Resilience*, 4(3), 147–165. <https://doi.org/10.1080/21693293.2016.1153772>
5. Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press. <https://lthj.qut.edu.au/article/view/1386>
6. Floridi, L. (2019). *The Logic of Information: A Theory of Philosophy as Conceptual Design*. Oxford University Press. https://www.researchgate.net/publication/340125134_LUCIANO_FLORIDI_The_Logic_of_Information_A_Theory_of_Philosophy_as_Conceptual_Design_Oxford_Oxford_University_Press_2019_272_s
7. Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., & Chazerand, P. (2018). AI4People: An ethical framework for a good AI society. *Minds and Machines*, 28(4), 689–707. <https://doi.org/10.1007/s11023-018-9482-5>
8. Guston, D. H. (2014). Understanding 'anticipatory governance'. *Social Studies of Science*, 44(2), 218–242. <https://doi.org/10.1177/0306312713508669>
9. Habermas, J. (1984). *The Theory of Communicative Action, Volume 1*. Beacon Press. <https://teddykw2.wordpress.com/wp-content/uploads/2012/07/jurgen-habermas-theory-of-communicative-action-volume-1.pdf>
10. Heeks, R., & Renken, J. (2022). Data justice for development: What, why, and how? *Information Technology for Development*, 28(1), 1–22. <https://journals.sagepub.com/doi/abs/10.1177/0266666916678282>
11. Hirschman, A. O. (1970). *Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States*. Harvard University Press. <https://pages.ucsd.edu/~bslantchev/courses/ps240/05%20Cooperation%20with%20States%20as%20Unitary%20Actors/Hirschman%20-%20Exit,%20voice,%20and%20loyalty%20%5BCh%201-5%5D.pdf>
12. Jaccard, J., & Jacoby, J. (2020). *Theory Construction and Model-Building Skills: A Practical Guide for Social Scientists* (2nd ed.). Guilford Press. <https://e.itg.be/MTM/qmm/2.pdf>
13. Jerven, M. (2013). *Poor Numbers: How We Are Misled by African Development Statistics and What to Do About It*. Cornell University Press. <https://dokumen.pub/poor-numbers-how-we-are-misled-by-african-development-statistics-and-what-to-do-about-it-9780801467615.html>
14. Jones, A. D., Ngunjiri, F. M., Pelto, G., & Young, S. L. (2013). "What are we assessing when we measure food security?" *Food Policy*, 36(1), 1–7. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3771133/>
15. Lewicki, R. J., McAllister, D. J., & Bies, R. J. (1998). Trust and distrust: New relationships and realities. *Academy of Management Review*, 23(3), 438–458. <https://www.jstor.org/stable/259288>
16. Madianou, M. (2019). Technocolonialism: Digital innovation and data practices in the humanitarian sphere. *Social Media + Society*, 5(3), 1–13. <https://doi.org/10.1177/2056305119863146>
17. Martin-Shields, C., & Kudejira, D. (2021). Digital humanitarianism, data governance, and the politics of anticipatory action. *Global Policy*, 12(S1), 19–29. <https://doi.org/10.1111/1758-5899.12950>
18. Meadows, D. H. (2008). *Thinking in Systems: A Primer*. Chelsea Green Publishing. <https://research.fit.edu/media/site-specific/researchfitedu/coast-climate-adaptation-library/climate-communications/psychology-amp-behavior/Meadows-2008.-Thinking-in-Systems.pdf>

19. Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on bias and fairness in machine learning. *ACM Computing Surveys*, 54(6), 1–35. <https://doi.org/10.1145/3457607>
20. Miller, H. G. (2019). The datafied state: Numbers, algorithms, and the politics of anticipatory governance. *International Political Sociology*, 13(2), 94–111. <https://doi.org/10.1093/ips/oly026>
21. OECD. (2019). *OECD Principles on Artificial Intelligence*. Organisation for Economic Co-operation and Development. <https://oecd.ai/en/ai-principles>
22. O’Neill, O. (2002). *A Question of Trust*. Cambridge University Press. <https://assets.cambridge.org/97805215/29969/sample/9780521529969ws.pdf>
23. O’Neil, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown. https://www.researchgate.net/publication/314165204_Cathy_O'Neil_Weapons_of_Math_Destruction_How_Big_Data_Increases_Inequality_and_Threatens_Democracy_New_York_Crown_Publishers_2016_272p_Hardcover_26_ISBN_978-0553418811
24. Rawls, J. (1971). *A Theory of Justice*. Harvard University Press. <https://doi.org/10.4159/9780674042605>
25. Taylor, L. (2017). What is data justice? The case for connecting digital rights and freedoms globally. *Big Data & Society*, 4(2). <https://doi.org/10.1177/2053951717736335>
26. UNESCO. (2021). *Recommendation on the Ethics of Artificial Intelligence*. United Nations Educational, Scientific and Cultural Organization. <https://unesdoc.unesco.org/ark:/48223/pf0000380455>
27. Watmough, G. R. et al. (2019). *Socioecologically informed use of remote sensing data to predict rural poverty*. PNAS, 116(4), 1213–1219. <https://www.pnas.org/doi/10.1073/pnas.1812969116>
28. Whetten, D. A. (1989). What constitutes a theoretical contribution? *Academy of Management Review*, 14(4), 490–495. <https://www.jstor.org/stable/258554>
29. World Food Programme (WFP). (2023). *Innovation Accelerator Annual Report 2023*. WFP. <https://www.uninnovation.network/innovation-library/wfp-innovation-accelerator---year-in-review-2023>
- World Food Programme (WFP). (2023). *HungerMap LIVE: Global Food Security Monitoring System*. <https://hungermap.wfp.org>