INDONESIAN JOURNAL OF MULTIDISCIPLINARY SCIENCE

Volume 5 No. 10 July 2026
P-ISSN: 2808-5957 E-ISSN: 2808-6724

From Babel to Nehemiah: A Theological Reading of Magnifica Humanitas and the Governance of Artificial Intelligence

Raymond R. Tjandrawinata
School of Bioscience, Technology and Innovation, Atma Jaya Catholic University of Indonesia, Indonesia
Email: raytjan@yahoo.com

Keywords
Magnifica Humanitas; Catholic Social Doctrine; artificial intelligence governance; Babel and Nehemiah; Trinitarian communion.

Abstract

This study reads the encyclical within its own theological context and questions what vision of systemic governance is implied in its argument. This reading is organized through the Systems Shift heuristic a seven-element framework drawn from socio-technical transition theory, leverage point analysis, lock-in studies, Responsible AI research, and AI governance literature but this framework serves to support the encyclical's argument, not to organize it. The study's central argument is that Magnifica Humanitas opposes two competing civilizational scripts. Babel names a social order dominated by concentrated technical power, the idolatry of profit, homogenization, and the functional reduction of human beings to data and performance. Nehemiah names the opposite order: patient reconstruction, distributed responsibility, prayerful wisdom, and the protection of the vulnerable. Leo XIV himself framed the AI era through these biblical images, and this paper shows that this framing carries the full weight of the encyclical's theological anthropology: the measure of the technological order is the Incarnation the Word made flesh who reveals the human person as irreducibly embodied, relational, free, and called to communion. Five new theological dimensions are developed and integrated: (1) Trinitarian communion as the theological foundation of the Nehemiah paradigm; (2) structural sin as the proper theological name for the Babel configuration; (3) pneumatology as the basis of shared understanding; (4) incarnational biology as a deepening of Christological anthropology against AI reductionism; and (5) Marian eschatology the grammar of the Magnificat as the encyclical's concluding horizon for the human city.

INTRODUCTION

Magnifica Humanitas opens with a sentence that is both a theological claim and a governance argument: humanity faces a pivotal choice between constructing a new Tower of Babel and building the city in which God and humanity dwell together (MH 1). That sentence does not introduce a technical policy discussion. It introduces a theological contest about what kind of social world the digital age is building and what kind of social world it ought to build.

Pope Leo XIV signed the encyclical on 15 May 2026, the 135th anniversary of Rerum Novarum, and it was released on 25 May 2026. It is his first encyclical and the first in the tradition of Catholic Social Doctrine devoted principally to artificial intelligence. Its 245 paragraphs across five chapters range from a historical overview of Social Doctrine, through foundational principles of the common good, to specific chapters on truth, work, freedom, and the civilization of love in the digital age. Throughout, the encyclical insists that it is not a technical handbook but a process of shared discernment: nourished by dialogue with the

sciences, ordered to the dignity of persons and the good of all peoples, and always read in the light of the Gospel.

The key programmatic claim comes in Chapter One: AI is not merely yet another theme to be studied or a crisis to be managed, but a development that challenges the categories of Social Doctrine from within, requiring their further development in fidelity to the Gospel (MH 17). That claim announces the paper's central task. If the encyclical is right that AI challenges the categories of social theology from within, then reading it requires more than listing its ethical recommendations. It requires reading at the level of social architecture: asking what configuration of power, personhood, labour, and governance Leo XIV is diagnosing, and what alternative configuration his theological vision entails.

This paper undertakes that reading. Its argument has three layers. First, it shows that the encyclical's Babel/Nehemiah contrast is not rhetorical decoration but a fully developed political theology of the digital age, grounded in biblical typology, Christological anthropology, and the tradition of Catholic Social Doctrine from Leo XIII to the present. Second, it employs a System Shift heuristic seven analytical elements drawn from socio-technical transitions, leverage-point, and AI governance scholarship to disclose the systemic structure of the encyclical's diagnosis. This heuristic is not imposed on the encyclical; it is drawn out from it. System Shift does not supply the norm of the argument; Catholic Social Doctrine does. The framework only helps disclose the systemic structure of the encyclical's own theological and social diagnosis. Third, it translates that diagnosis into five governance proposals that are anchored in the encyclical and mapped to existing international frameworks (Geels, 2019; Mäntymäki et al., 2022).

Previous research on AI governance from a theological perspective has largely focused on either general ethical principle (such as human dignity and the common good) or specific applications (such as autonomous weapons or algorithmic bias) (Lacsa & Arroyo, 2024; Jobin et al., 2019). However, there has been limited work that integrates a systemic analysis of the socio-technical dynamics of AI with a theological reading of a specific magisterial document. This paper addresses that gap by reading Magnifica Humanitas at the level of social architecture and by developing a framework that can be applied to other AI governance contexts. The urgency of this task is underscored by the nature article in which the author, who advises the Vatican and the UN on AI, argues that "the papal letter goes beyond a religious document and diagnoses a failure in AI governance that the scientific community should heed" . The novelty of this paper lies in its integration of socio-technical transitions scholarship with theological ethics and in its development of concrete governance proposals that are anchored in the encyclical's theological vision.

METHOD

A System Shift Reading of Magnifica Humanitas

The System Shift heuristic employed in this paper was a synthesized analytic lens, not an authority independent of the encyclical. Its seven elements are drawn from socio-technical transitions scholarship (Geels), leverage-point analysis (Meadows), lock-in and path dependency, Responsible AI research (Nabavi and Browne), AI impact assessment (Ryan & Stahl, 2021; Stahl et al., 2023), and AI risk governance (NIST; UNESCO; EU AI Act). These elements are brought to the encyclical not to govern its interpretation but to disclose the

systemic structure of its own diagnosis. Table 2 maps each element to specific passages and governance implications.

Table 1 System Shift Framework for Theological Analysis of Magnifica Humanitas and AI Governance

Element	Analytical Question	MH Diagnosis	Governance Implication
System Condition	What is the historical configuration of power?	A 'change of era': AI reconfigures labour, truth, politics, war, and personhood at socio-technical scale (MH 4–6, 17)	Governance must match the level of transformation not piecemeal ethics, but systemic redesign at architectural level
Domain Lock	What stabilizes an unjust configuration against reform?	Technocratic paradigm; profit idolatry; platform concentration; myth of algorithmic neutrality; transhumanist narratives (MH 9–10, 93, 104, 115–118)	Lock-ins must be named and structurally dismantled, not managed; requires normative intervention, not only alignment fixes
Actors	Who in practice governs the AI transition?	Private transnational actors whose resources exceed many governments; control over data, infrastructure, visibility rules (MH 5, 71–72)	Multi-stakeholder governance; Church as public-theological voice; distributed ecology of responsibility
Chokepoints	Where is power materially exercised?	Data ownership; computing resources; platform algorithms; supply chains; hidden labor; decision systems affecting rights (MH 101–109)	Registries, due diligence, supply-chain transparency, labor audits, impact assessment
Position	From where does the Church speak, and with what authority?	Not technical blueprint; voice of principles and shared discernment, in dialogue with sciences and all people of goodwill (MH 22–27)	Public theology, not technocracy; whole People of God as distributed moral actor, not clerical monopoly
Strategy	What institutional interventions translate principle into practice?	Assessment, transparency, participation, recourse, labour protections, digital literacy, stronger political oversight (MH 16, 71–72, 105–109)	AI HDIA; Digital Common Good Governance; Subsidiarity by Design; Solidarity Checkpoint; Freedom Safeguard
Feedback	How are correction and accountability structured?	Independent checks, avenues of appeal, no exclusion without recourse, accountability at every stage of design–deployment (MH 71, 103–106)	GOVERN/MEASURE cycle (NIST); auditability (UNESCO); human oversight (EU AI Act); public annual review

System Condition: The Change of Era

Magnifica Humanitas describes a change of era in which digitalization, AI, and robotics are transforming the world rapidly and profoundly. Crucially, the encyclical does not

characterize this transformation merely in terms of new tools or capabilities. It insists that these systems have become interwoven into the fabric of daily life, shaping decision-making processes and deeply affecting the collective imagination (MH 4) (Zuboff, 2019). AI is not simply a more powerful instrument; it is an environment a new condition of social life within which personhood, truth, work, and freedom are being reorganized (Geels, 2019; Geels, 2020).

This systemic diagnosis justifies the use of a transitions-level heuristic. If AI were merely a new tool, governance could be organized as a series of regulatory adjustments (Birkstedt et al., 2023). Because it is a socio-technical regime reshaping institutions, narratives, epistemic conditions, and material infrastructures simultaneously governance must operate at the same level of depth and breadth as the transformation it addresses.

Domain Lock: Structural Stabilization of the Babel Configuration

The encyclical identifies several overlapping mechanisms that stabilize the Babel configuration against reform mechanisms that transitions scholarship calls lock-ins. The technocratic paradigm, inherited from Laudato Si' and intensified in MH's account of the Babel syndrome, is the dominant cultural lock-in (Corrêa et al., 2023): the worldview that technical power is inherently progressive, that efficiency is the primary value, and that human weakness is a problem to be solved rather than a condition to be received (MH 9–10, 115–118).

Alongside the cultural lock-in, the encyclical identifies material and institutional lock-ins: platform concentration that controls visibility, access, and participation; profit-maximizing business models built on behavioural data extraction; and a governance vacuum created by the fact that the primary actors of development are private, transnational, and largely unaccountable to democratic institutions (Coveri et al., 2024; Radu, 2021) (MH 5, 71–72). These are not independent failures but a mutually reinforcing system: the cultural narrative justifies the institutional arrangements, which in turn produce the material concentrations that make alternative arrangements appear impractical.

In the language of Simoens and colleagues, these are discursive lock-ins: not merely material constraints on reform but cognitive and cultural frameworks that make the present configuration appear natural, necessary, and inevitable (Geels, 2020). Breaking them requires not only technical regulation but what the encyclical calls shared discernment: the recovery of alternative social imaginaries capable of making different configurations thinkable and desirable.

Actors, Chokepoints, Position, Strategy, and Feedback

The encyclical is unusually explicit about the actor map of the AI transition. The primary drivers of development are private, transnational actors whose resources and capacity to intervene surpass those of many governments (Radu, 2021; Birkstedt et al., 2023) (MH 5). This observation has decisive implications for governance architecture: any framework that imagines the state as the sole upper-level authority misunderstands the present distribution of digital power. The relevant actors are multiple and asymmetrically resourced: technology corporations, state regulators, multilateral bodies, civil society organizations, faith communities, research institutions, and affected populations.

The encyclical's chokepoints include data ownership, computational resources, platform algorithms, supply chains, and decision systems affecting rights. Its most striking chokepoint analysis insists that AI is never immaterial: it consumes energy and water, depends on physical infrastructure, and rests on the hidden labour of workers who sustain algorithmic systems and

are often exploited in the process (Regilme, 2024; Aloisi & De Stefano, 2022) (MH 101, 175–180). Governance that abstracts from this material reality reproduces the asymmetries it claims to address.

The Church's position in this actor landscape is carefully defined. Social Doctrine is not a technical blueprint, and the Church does not claim to possess a monopoly on truth (MH 25). Its contribution is one of principles for shared discernment, in dialogue with all people of goodwill. This is a public-theological, not technocratic, position: the Church speaks not as a technical authority but as a witness to truth, justice, and communion in history. (Lacsa & Arroyo, 2024)

The encyclical's strategy explicitly rejects two inadequate responses: naive technological enthusiasm, and alignment discourse controlled by those who already dominate the field (Ryan & Stahl, 2021; Mäntymäki et al., 2022) (MH 107). Instead, it calls for interventions that are political, legal, and participatory: assessment of human and social impact from the outset of design; transparent and accountable governance; equitable access to data and infrastructure; meaningful participation of affected communities; avenues for recourse; labour protections; and stronger political oversight of private actors.

The feedback element accountability structures capable of enabling challenge and correction is the institutional form of the encyclical's demand that no person be excluded from appeal, and that those responsible for design, development, and deployment be identifiable, challengeable, and correctable (MH 71, 103–106). This is not merely procedural; it is theological: a system that forecloses appeal denies the personhood of those it affects.

RESULTS AND DISCUSSION

Babel and Nehemiah as Rival Theological-Social Imaginaries

The Babel Syndrome as Structural Sin

Leo XIV's opening biblical contrast begins with a diagnostic that runs deeper than individual moral failure. Babel is not simply individual pride writ large by technology. It is the structural deformation of technical power: a social sin in which systems of profit, visibility, prediction, and control make the reduction of persons appear normal, efficient, and inevitable.

The language of structural sin belongs to the tradition of Catholic Social Doctrine. John Paul II developed it in Sollicitudo Rei Socialis to name configurations of institutional life that contradict God's will for human flourishing, not because any single actor is wholly malevolent but because the arrangement of incentives, rules, and power makes harm systematic. In Magnifica Humanitas, Leo XIV applies precisely this category to the digital order. The Babel syndrome is not primarily about bad intentions; it is about a social architecture in which the idolatry of profit, platform concentration, the elimination of diversity, and the reduction of the person to a data point have become structural features of how AI systems are designed, financed, and deployed (MH 9–10). The Babel configuration is, in this strict sense, a digital form of structural sin.

This naming matters. Without the category of structural sin, Babel remains a moral metaphor a useful warning against pride and greed, but one that can be addressed through individual conversion or corporate ethics commitments. With the category of structural sin, Babel becomes a theological diagnosis of a social condition that requires structural intervention: not only personal conversion, but the transformation of the institutions,

incentives, and power arrangements that make dehumanization systemic. This is why Leo XIV insists that a more moral AI is not enough if that morality is determined by a few (MH 107): alignment controlled by those who already dominate the field reproduces the Babel configuration under a new name.

The Nehemiah Paradigm as Trinitarian Communion

If Babel names a structural sin, Nehemiah names a social form that is theologically ordered. But the Nehemiah paradigm in Magnifica Humanitas is not merely a model of participatory governance. It is a social analogue of communion and its deepest ground is Trinitarian.

The encyclical grounds its vision of social life in the mystery of the Triune God. The human person is created in the image of God who is, in himself, a communion of persons: Father, Son, and Holy Spirit, not a unity achieved by suppressing difference but a unity that is the overflowing fullness of distinct relations (MH 48–50). From this Trinitarian anthropology follows the encyclical's social vision: authentic unity is communion, not uniformity. Plurality is not a threat to order but the form through which shared life becomes possible. This is why Babel fails: it confuses unity with homogenization, and discovers that the elimination of difference produces not communion but dispersion.

Nehemiah models the contrary. The rebuilt city of Jerusalem is not a uniform project but a distributed labor of many people men, women, priests, artisans, heads of households, young people each assigned a section of the wall, each contributing their distinct role to a shared work (MH 8). This is communion in the Trinitarian sense: distinct persons, distinct vocations, distinct responsibilities, ordered to a single good that none could achieve alone. The Nehemiah paradigm resists the Babel-like fantasy of unity through uniformity and proposes instead a communion in which plurality is not a threat to order but the very form through which the human city becomes a place of safe return.

For AI governance, the Trinitarian ground of the Nehemiah model has a specific implication: governance cannot be adequately designed around a single actor, a single standard-setter, or a single decision procedure. Any governance architecture that concentrates authority whether in a state, a corporation, or an intergovernmental body reproduces the Babel structure even if its intentions are good. Legitimate governance of the digital age must be genuinely plural: distributed across actors, sectors, traditions, and communities, accountable to many rather than few, capable of translating diversity into a resource rather than a problem.

Table 2. The Babel/Nehemiah Contrast

Dimension	The Babel Paradigm	The Nehemiah Paradigm
Ontological orientation	Reality as finished object: the system, the tower, the archive to be completed and possessed	Reality as ongoing process: the city is always being rebuilt; governance is always shared labour
Social form	Uniformity, centralization, single algorithm, homogenization of difference	Distributed responsibility, plural voices, coordination without uniformity
Anthropology	Person reduced to data exhaust, cognitive performance, optimization target, predictive profile	Person as embodied, relational, vulnerable, irreducible, called to love and communion

Dimension	The Babel Paradigm	The Nehemiah Paradigm
Political economy	Private transnational actors monopolizing data, infrastructure, and visibility; profit-idolatry	Subsidiarity, universal access, shared stewardship of digital commons, labor protected
Epistemology	One code, one metric, one language of value; mystery translated into performance	Discernment, plural sciences, truth pursued in common, Spirit active in history
Labor	Hidden, exploited workers sustaining algorithmic systems; new forms of digital slavery	Dignity of work; fair wages; no person reduced to productivity or cognitive performance
Theological horizon	Self-sufficiency without God; person as technical substrate; weakness as error to be corrected	Incarnation as measure of all things; history as place of communion; weakness received, not erased

The table synthesizes the encyclical's development of the Babel/Nehemiah typology across its five chapters, together with the anti-technocratic horizon of Laudato Si' and the politics of social friendship in Fratelli Tutti. Its most important row is the first: the ontological contrast between reality as static object and reality as ongoing process. That contrast is not only social but metaphysical. The Babel configuration treats the digital system as something to be completed and possessed; the Nehemiah configuration treats governance as something to be continually rebuilt, corrected, and renewed. The human city is always in process. This is the ontological warrant for every institutional proposal that follows.

The Incarnation, Embodied Life, and the Measure of Technological Order

The Incarnation as the Normative Centre

Any theologically serious reading of Magnifica Humanitas must ask what vision of the human person governs the argument. The encyclical's answer is given with characteristic directness in its conclusion: at the heart of everything is the mystery of the Incarnation (MH 231–234). Gaudium et Spes had established the Christological anthropology that runs through all subsequent Catholic Social Doctrine: it is only in the mystery of the Word made flesh that the mystery of humanity truly becomes clear; the human person does not fully find himself except through sincere self-gift (GS 22, 24). Magnifica Humanitas resumes and deepens that anthropology in explicit contrast to the logic of AI.

The Christological norm is precise and consequential. Dignitas Infinita (2024) argues that dignity is universal, intrinsic, and not contingent on qualities or circumstances; Christ's Incarnation reveals that dignity can never be lost, particularly in the weak and despised (Dignitas Infinita 17–19). Antiqua et Nova (2025) applies that theological anthropology to AI directly, insisting that AI applications must be judged by their embedded vision of the person and the common good (Antiqua et Nova 43, 50). Magnifica Humanitas completes this trajectory: the measure of any technological order is not its computational capacity or its economic efficiency, but whether the persons who live within it are recognized in their embodied, relational, vulnerable, and irreducible dignity as persons called to relationship and communion, not projects to be optimized.

Incarnational Biology and the Irreducibility of Embodied Life

The Christological anthropology of Magnifica Humanitas can be deepened by attending to what may be called an incarnational-biological perspective: the recognition that the Word did not assume a neutral computational form but human flesh biological life, with all its vulnerability, metabolic dependence, affective resonance, and historical situatedness. This deepening is not extraneous to the encyclical's argument; it makes explicit what the encyclical implies when it refuses to identify human intelligence with machine processing.

The human person is not an abstract bearer of cognition, nor merely an information-processing system capable of being modelled, predicted, and optimized. The person is embodied life: vulnerable, metabolic, affective, relational, historical, and open to God. In the Incarnation, the Word assumes not a mind but a body a body that is born, grows, hungers, is touched, suffers, ages, and dies. This means that the body is not a disposable substrate beneath intelligence; it is the concrete place where dignity, dependence, freedom, suffering, labour, love, and communion are lived.

For this reason, the theological critique of AI reductionism cannot be limited to the claim that machines lack consciousness or soul. It must also insist that AI lacks the biological and existential condition through which human personhood is actually lived: birth, fatigue, illness, grief, touch, aging, mortality, and self-giving love. An AI system may process discourse about vulnerability, but it does not inhabit vulnerable flesh. Magnifica Humanitas should therefore be read not only as a defence of rational or moral agency, but as a defence of embodied human life against its translation into data, prediction, and performance.

The practical implication for governance is concrete. AI systems that make decisions about employment, creditworthiness, healthcare access, or criminal justice affect not abstract cognizes but embodied persons whose vulnerabilities, histories, and needs exceed any model's capacity for representation. Governance frameworks that treat affected persons only as bearers of data points, rather than as embodied, mortal, and relational beings, reproduce the Babel reduction even within ethical frameworks. The AI Human Dignity Impact Assessment proposed below is, in this sense, an incarnational instrument: it asks not only whether a system is accurate or fair in the abstract, but whether it receives the person in their embodied singularity.

The Common Good, Subsidiarity, and Solidarity

From dignity follows the common good. Magnifica Humanitas offers a relational account: the common good is not a mere aggregate of private interests but a relational surplus generated by interdependence, reciprocity, and shared responsibility (MH 59–64). The implication for digital governance is structural. Data, platforms, algorithms, and computational infrastructures are not merely objects of private property or market advantage. Their effects are constitutively public, and their governance must reflect that public character. This is why the encyclical describes data as something to be managed as a common or shared good: not through nationalization, but by refusing the absolutist private-property logic that treats socially constituted goods as private resources (MH 65–67).

Subsidiarity and solidarity supply the institutional grammar of the Nehemiah order. Subsidiarity requires that governance remain as close as possible to those it affects, protecting the capacity of communities to decide and correct. In the digital order, where power has migrated from states to private transnational actors, subsidiarity does not mean less governance

but differently structured governance: independent checks, algorithmic transparency, equitable access, participation, and recourse (MH 68–72). Solidarity prevents subsidiarity from collapsing into localism by insisting that the future of each is bound to the future of all that the hidden workers, the marginalized communities, the poorer states, and the children are not residual casualties of digital progress but its co-bearers and co-inheritors (MH 73–74).

Governing for the Common Good: Five Proposals

The governance proposals that follow are not external supplements to Magnifica Humanitas. Leo XIV himself calls for responsible planning, assessment of human and social impact, inclusion of the vulnerable, digital literacy, due diligence, effective safeguards, meaningful participation, transparency, and recourse (MH 16, 71–72, 105–109, 178–180). Contemporary international governance frameworks provide the institutional vocabulary by which those requirements can be implemented. Table 3 maps each proposal to its doctrinal grounding, its governance analogue, and its core requirement.

Table 3 Governance Proposals Derived from Magnifica Humanitas for Ethical Artificial Intelligence

Proposal	Doctrinal Grounding (MH)	Global Analogue	Core Requirement
AI Human Dignity Impact Assessment	Dignity as design criterion; assessment of human and social impact from outset; no person reducible to data or performance (MH 16, 104–109)	UNESCO EIA; NIST MAP/MEASURE; Stahl et al.	Ex ante, mid-cycle, and ex post review of rights, labour, bias, environment, recourse
Digital Common Good Governance	Data as shared good; anti-monopoly on infrastructure; universal destination of goods; data not merely private property (MH 65–67, 108–109)	OECD AI Principles; EU AI Act public-rights orientation	Algorithmic registries; public-interest oversight; transparent procurement; reporting duties
Subsidiarity by Design	Local agency; participation; recourse; no paternalistic management; standards not set elsewhere without community voice (MH 68–72, 105–109)	NIST GOVERN; EU human oversight clauses	Community consultation; contestability; documented human-override pathways; explainability proportionate to context
Solidarity Checkpoint	Hidden workers; colonial data extraction; global justice; no new forms of slavery; gains not to accrue to few (MH 73–74, 109, 175–180)	UNESCO fairness and due diligence; supply-chain review frameworks	Labor audits; Global South equity; child protections; anti-trafficking safeguards; cross-border distributional review

Proposal	Doctrinal Grounding (MH)	Global Analogue	Core Requirement
Freedom Safeguard	Against manipulation, attention capture, behavioural profiling; freedom as capacity for truth and love, not only absence of coercion (MH 141–146, 170–171)	UNESCO transparency/oversight; EU AI Act prohibited practices	Anti-addictive design rules; transparency labels; proportional limits on intrusive systems; robust complaint and remedy

AI Human Dignity Impact Assessment

The AI Human Dignity Impact Assessment is the central instrument of the proposed architecture. Its task is not compliance checking but prospective moral reasoning (Stahl et al., 2023; Choung et al., 2023): does this system's design and deployment risk reducing persons to less-worthy lives, foreclosing appeal, eroding work, exploiting unseen labour, harming children, intensifying discrimination, or commercializing vulnerability? The incarnational-biological perspective developed in Section 6 grounds the assessment instrument: it asks not only whether a system is accurate or fair in the abstract, but whether it receives the embodied, relational, singular person in their full humanity.

UNESCO's Ethical Impact Assessment provides the closest existing analogue: a structured instrument for assessing positive and negative impacts ex ante, requiring public-facing information and supporting mitigation strategies (Stahl et al., 2023). Stahl and colleagues show that AI impact assessments are emerging as a convergent governance practice because they stimulate structured reflection across AI ecosystems, bridging technical, legal, ethical, and social dimensions that standard risk frameworks tend to isolate. The NIST AI Risk Management Framework's MAP and MEASURE functions provide complementary tools for categorizing AI context and documenting impacts through the deployment lifecycle.

Digital Common Good Governance

Digital Common Good Governance follows from the encyclical's claim that patents, algorithms, digital platforms, technological infrastructures, and data must now be included among the goods whose accumulation in a few hands creates unjust imbalance (Dencik et al., 2019) (MH 65–67). The proposal does not require abolishing private ownership of data; it requires refusing the absolutist property logic in domains whose social effects are constitutive of public life. Practically, this supports algorithmic registries for high-impact systems, public-data stewardship models, transparent procurement standards for government use of AI, and reporting duties for systems affecting rights. The OECD AI Principles' values-based approach and the EU AI Act's public-rights orientation provide workable regulatory analogues. (Kusche, 2024)

Subsidiarity by Design

Subsidiarity by Design means that communities are not merely passive recipients of standards set elsewhere. Institutions deploying AI must build procedures for consultation, explainability proportionate to context, rights of contestation, documented human roles, and meaningful override (Choung et al., 2023) when fundamental rights or access to essential goods are at stake. The encyclical states that subsidiarity requires protecting the ability of

communities to decide and correct, without relegating their intervention to monitoring conducted after standards have been established elsewhere (MH 68). The NIST GOVERN function's requirements for documented roles, executive accountability, diverse teams, and structured incorporation of external feedback provide an implementation template.

Solidarity Checkpoint

The Solidarity Checkpoint is the distributive conscience of the architecture. It asks: who bears the hidden burdens of this system's success? The encyclical insists that solidarity obliges recognition of the hidden, often exploited workers who sustain algorithmic systems (MH 109), and names new forms of digital slavery, colonial data extraction, and the danger that AI's gains accrue to a few while its harms are displaced onto the vulnerable. The Checkpoint requires supply-chain transparency, labour due diligence, fair data practices across borders, child and anti-trafficking safeguards (Regilme, 2024), and attention to cross-border inequities. It concretizes a solidarity that is institutional, not merely sentimental.

Freedom Safeguard

The Freedom Safeguard translates the encyclical's account of digital dependence and manipulation into policy. But it does so with an anthropology that is distinctively Catholic: freedom in Christian theology is not merely the absence of external manipulation. It is the capacity of the person to respond to truth and love. AI systems that capture attention, predict behaviour, and subtly narrow the horizon of desire therefore threaten freedom not only politically but spiritually (Susser et al., 2019) they reshape the conditions under which persons learn to choose the good (MH 141–146, 170–171).

This proposal includes design restrictions on manipulative interfaces, strict safeguards for minors, disclosure requirements for AI-mediated interaction, limits on invasive profiling, and robust complaint and remedy mechanisms. The EU AI Act's prohibited practices and transparency obligations, and UNESCO's principles on oversight, literacy, and fairness, provide workable legal analogues. The encyclical's Christological warrant for this proposal is the insistence that the person bears an interiority and a vocation to love and worship that no architecture of behavioural capture can legitimately colonize.

Counterarguments and Limitations

The first counterargument is that the Babel/Nehemiah framing is too allegorical to do serious governance work. That objection can be answered on textual grounds: the symbolism is Leo XIV's own chosen frame, joined to concrete discussions of data, platforms, algorithms, labour, education, war, and political responsibility. More importantly, the objection misunderstands what social imaginaries do. As Section 3 argued, the Babel/Nehemiah contrast functions as a normative social imaginary: it names the total architecture of a social configuration, not merely its surface features. That is precisely why it is analytically useful for governance it shows what is at stake at the level of social design, not only at the level of specific policy choices.

The second counterargument is that theology is not competent to address AI governance. Magnifica Humanitas anticipates this objection by denying that the Church offers definitive technical solutions while insisting that it contributes principles for shared discernment in dialogue with the sciences (MH 23–27). This paper accepts that framing entirely. Its purpose is not to replace computer science, law, or public administration, but to ask what kind of human being, community, and future those disciplines ought to serve. That is where theology is both

competent and necessary: no technical framework can answer that question from within its own resources, because the answer requires an account of the good that exceeds the measurable (Carabantes, 2020; Laine et al., 2024).

The third counterargument concerns the System Shift framework itself. A theology reviewer might object that System Shift controls the reading of the encyclical, subordinating theological argument to social-scientific analysis. The response is given explicitly in the paper's framing: System Shift does not supply the norm; Catholic Social Doctrine does. The seven elements are heuristic servants, not epistemic governors. Their function is to prevent the encyclical's systemic claims from being dissolved into a list of ethical platitudes. They draw out the structural depth of what the encyclical is already doing. A reviewer unconvinced by the framework's terminology can strip it away and find the theological argument standing on its own.

The fourth and most serious counterargument is that the paper places too much hope in governance architecture. Catholic theology does not assume that better procedures alone can redeem distorted power. Sin is not solved by frameworks. The answer is that the governance architecture proposed here is not presented as salvation by governance, but as a disciplined form of social responsibility under grace, always aware of sin, institutional capture, failure, and the constant need for conversion. The Nehemiah model is instructive here: Nehemiah does not announce a perfect constitution. He organizes a community of imperfect, vulnerable people to rebuild, section by section, under God, knowing that opposition will come and that the work will never be finally complete. Governance, in this theological grammar, is not redemption but faithfulness.

A fifth limitation should be acknowledged candidly: the paper interprets a text published only weeks before its own composition. The secondary literature on Magnifica Humanitas is necessarily thin. Future scholarship will reveal dimensions of the encyclical that this reading does not adequately capture. The paper should be understood as a first theological-systematic reading rather than a definitive interpretation.

CONCLUSION

This article demonstrates that Magnifica Humanitas is not a technical guide to AI governance, but rather a theological work within the tradition of Catholic Social Teaching that discerns new configurations of power in the digital age. Analysis using the System Shift heuristic reveals two opposing paradigms: Babel, as a manifestation of structural sin that emphasizes homogenization, profit-idolatry, and the reduction of humans to data; and Nehemiah, as a paradigm of Trinitarian communion that emphasizes the responsibility of distribution, participation, and protection of the vulnerable. From a theological perspective, this document emphasizes that technological progress must be measured against an acceptance of the incarnate, relational, and irreplaceable dignity of the human person. The article identifies five AI governance proposals: AI Human Dignity Impact Assessment, Digital Common Good Governance, Subsidiarity by Design, Solidarity Checkpoint, and Freedom Safeguard, designed to translate theological principles into ethical and inclusive digital governance practices. Thus, the article emphasizes that ethical AI governance requires systemic analysis, plural governance, and pneumatologically attention. Technology is not just a tool, but a social environment that influences personhood, work, and freedom. The challenge of the future is not simply to build

intelligent machines, but to rebuild a humane city that respects dignity, plurality, and community in line with the incarnational and communion vision outlined by Magnifica Humanitas.

REFERENCES

Aloisi, A., & De Stefano, V. (2022). Essential jobs, remote work and digital surveillance: Addressing the COVID-19 pandemic panopticon. International Labour Review, 161(2), 289–314. https://doi.org/10.1111/ilr.12219

Birkstedt, T., Minkkinen, M., Tandon, A., & Mäntymäki, M. (2023). AI governance: Themes, knowledge gaps and future agendas. Internet Research, 33(7), 133–167. https://doi.org/10.1108/INTR-01-2022-0042

Carabantes, M. (2020). Black-box artificial intelligence: An epistemological and critical analysis. AI & Society, 35(2), 309–317. https://doi.org/10.1007/s00146-019-00888-w

Choung, H., David, P., & Ross, A. (2023). Trust in AI and its role in the acceptance of AI technologies. International Journal of Human–Computer Studies, 174, 103007. https://doi.org/10.1016/j.ijhcs.2023.103007

Corrêa, N. K., Galvão, C., Santos, J. W., Del Pino, C., Pinto, E. P., Barbosa, C., Massmann, D., Mambrini, R., Galvão, L., Terem, E., & de Oliveira, N. (2023). Worldwide AI ethics: A review of 200 guidelines and recommendations for AI governance. Patterns, 4(10), 100857. https://doi.org/10.1016/j.patter.2023.100857

Coveri, A., Cozza, C., Guarascio, D., & Zanfei, A. (2024). Platform power: Monopolisation and financialisation in the era of big tech. Structural Change and Economic Dynamics, 70, 394–407. https://doi.org/10.1016/j.strueco.2024.03.016

Dencik, L., Hintz, A., Redden, J., & Trèrè, E. (2019). Exploring data justice: Conceptions, applications and directions. Information, Communication & Society, 22(7), 873–881. https://doi.org/10.1080/1369118X.2019.1606268

Geels, F. W. (2019). Socio-technical transitions to sustainability: A review of criticisms and elaborations of the Multi-Level Perspective. Current Opinion in Environmental Sustainability, 39, 187–201. https://doi.org/10.1016/j.cosust.2019.06.009

Geels, F. W. (2020). Micro-foundations of the multi-level perspective on socio-technical transitions: Developing a multi-dimensional model of agency through crossovers between social constructivism, evolutionary economics and neo-institutional theory. Technological Forecasting and Social Change, 152, 119894. https://doi.org/10.1016/j.techfore.2019.119894

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. Nature Machine Intelligence, 1(9), 389–399. https://doi.org/10.1038/s42256-019-0088-2

Kusche, I. (2024). Possible harms of artificial intelligence and the EU AI Act: Fundamental rights and risk. Journal of Risk Research, 27(9), 1143–1159. https://doi.org/10.1080/13669877.2024.2350720

Lacsa, J. E., & Arroyo, R. C. (2024). Integrating Catholic social teaching with AI ethics to address inequity in AI healthcare. Journal of Religion and Health, 63(5), 3435–3454. https://doi.org/10.1007/s10943-024-02140-2

Laine, J., Minkkinen, M., & Mäntymäki, M. (2024). Ethics-based AI auditing: A systematic literature review on conceptualizations of ethical principles and knowledge contributions

to stakeholders. Information & Management, 61(5), 103963. https://doi.org/10.1016/j.im.2024.103963

Mäntymäki, M., Minkkinen, M., Birkstedt, T., & Viljanen, M. (2022). Defining organizational AI governance. AI & Society, 37(4), 1452–1462. https://doi.org/10.1007/s00146-022-01477-5

Radu, R. (2021). Steering the governance of artificial intelligence: National strategies in perspective. Policy and Society, 40(2), 178–193. https://doi.org/10.1080/14494035.2021.1899866

Regilme, S. S. (2024). Artificial intelligence colonialism: Environmental damage, labor exploitation, and human rights crises in the Global South. SAIS Review of International Affairs, 44(2), 75–92. https://doi.org/10.1353/sais.2024.a950958

Ryan, M., & Stahl, B. C. (2021). Artificial intelligence ethics guidelines for developers and users: Clarifying their content and normative implications. Journal of Information, Communication and Ethics in Society, 19(1), 61–86. https://doi.org/10.1108/JICES-12-2019-0138

Stahl, B. C., Antoniou, J., Bhalla, N., Brooks, L., Jansen, P., Lindqvist, B., Kirichenko, A., Marchal, S., Rodrigues, R., Santiago, N., Warso, Z., & Wright, D. (2023). A systematic review of artificial intelligence impact assessments. Artificial Intelligence Review, 56(11), 12799–12831. https://doi.org/10.1007/s10462-023-10420-8

Susser, D., Roessler, B., & Nissenbaum, H. (2019). Technology, autonomy, and manipulation. Internet Policy Review, 8(2). https://doi.org/10.14763/2019.2.1410

Zuboff, S. (2019). Surveillance capitalism and the challenge of collective action. New Labor Forum, 28(1), 10–29. https://doi.org/10.1177/1095796018819461