‘Magnifica Humanitas’

COMMONWEAL CONVERSATION STARTER SERIES

2026 - 2027

‘Magnifica humanitas’

In May 2026, Pope Leo XIV released his first encyclical letter, *Magnifica humanitas*. The encyclical explores the cultural and spiritual implications of artificial intelligence—a revolutionary, society-shifting technology that is becoming increasingly ubiquitous. AI is a relatively new invention, and the long-term consequences of its implementation are yet to be seen. In *Magnifica humanitas*, the pope voices a thoughtful wariness about how the technology might affect our humanity: our art, social consciousness, and collective civic duties.

In the encyclical, Leo XIV confidently declares that using generative AI can “weaken personal creativity and judgment.” Antonio Spadaro reckons with how and why a limitless implementation of AI could cause human creativity to “atrophy” and eventually “evaporate.” In the article “Art and AI,” he emphasizes how human intelligence is uniquely carnate: we inhabit, grow, and can sense our being. Our emotions come with physical expression. Spadaro and the pope are not afraid that data will be able to produce a Rothko, but rather that a conscious-less AI could “colonize the mental habitat of mankind.” Sharing similar anxieties, Meghan Sullivan breaks down *Magnifica humanitas*’s discussion of AI with a framework for maintaining dignity, embodiment, love, transcendence, and agency. She explains how each concept is uniquely human, and, in dialogue with the pope’s encyclical, how AI threatens these essential parts of humanity.

Magnifica humanitas also revives Paul IV’s aspiration for a “civilization of love.” Frank Pasquale explores what that means and how AI fits into such a universalist civilizational discourse. Leo XIV rejects technology as a panacea and advocates for pluralism, protecting those who choose not to engage in the “AI arms race.” Pasquale paints this kind of civic love not as idealism, but rather as our Christian duty—we must not allow the advance of AI “to follow the pattern of so many other technologies, where the vast majority of the benefits go to a small elite and the disadvantaged scramble for crumbs.” The ultimate aim of technological mastery expressed in *Magnifica humanitas* is a merciful and charitable civilization of love. Kevin Gallagher discusses how Leo XIV’s description of the Church’s social teaching can lead us there. The Church and its social doctrine is dynamic, developing alongside ever-changing “living realities.” Unlike previous popes, who have attempted to right contradictions in the Church’s social teaching with “lawyerly creativity,” Leo XIV is not “taking the easy way out.” He confronts uncomfortable historical truths of Catholic doctrine, demonstrating that its evolution is a process, one which cannot be derived axiomatically from Biblical truths.

One moment in the Church’s history which is extremely uncomfortable for Catholics today is the institution’s complicity in slavery, which Pope Leo XIV offers an unexpected and forthright apology for in *Magnifica humanitas*, calling it a “wound in Christian memory.” John T. McGreevy links the development of Church

teaching on slavery to the development of Church doctrine: development being "a fighting word for philosophers and theologians, since it implies that things now declared settled might one day no longer be so." McGreevy believes he and the pope both understand human dignity as materially grounded, not an abstract concept.

Leo XIV's discussion of AI in *Magnifica humanitas* also reveals a broader truth about our societal aspirations. AI, like many other recent shifts in mainstream infrastructure, seeks to make time-consuming and effortful aspects of the human condition—thinking, creating, synthesizing—easier. Vincent J. Miller describes how the encyclical frames the cultural shift toward efficiency and ease as counterproductive, because "humanity flourishes not despite limitations, but often through them." Miller centers our "fleshly finitude"—our illnesses, our incapabilties—as perhaps humanity's greatest asset, one which can never be replicated by increasingly powerful LLMs.

READINGS FOR DISCUSSION:

Antonio Spadaro, "Art and AI," July/August 2026:
https://www.commonwealmagazine.org/art-and-ai

Meghan Sullivan, "Us vs. the Robots," July/August 2026:
https://www.commonwealmagazine.org/sullivan-leo-delta-ai-magnifica-humanitas

Frank Pasquale, "Toward a Civilization of Love," July/August 2026:
https://www.commonwealmagazine.org/leo-magnifica-humanitas-ai-civilization-love

Kevin Gallagher, "Mild & Bold," July/August 2026:
https://www.commonwealmagazine.org/gallagher-magnifica-humanitas-leo-church-ai

John T. McGreevy, "Coming Clean," July/August 2026:
https://www.commonwealmagazine.org/mcgreevy-pope-leo-slavery-apology-magnifica-humanitas

Vincent J. Miller, "Flourishing Through Limits," July/August 2026:
https://www.commonwealmagazine.org/flourishing-through-limits

DISCUSSION QUESTIONS:

1. John T. McGreevy and Kevin Gallagher both emphasize that the encyclical makes a point of directly acknowledging, and not theologically explaining, lapses in Catholic social teaching. How do you understand the Church's developing social doctrine? Are there some axiomatic truths which are immutable, or is everything subject to change?

2. Across the symposium, our writers agree with the pope that AI is not inevitable. However, fears of creative "atrophy" and further economic inequality, amongst others, are prevalent. How should AI be used if we are to maintain our "magnificent humanity?"

3. Leo XIV's encyclical has been popular with both devout Catholics and secular readers. Vincent J. Miller suggests that the response "speaks to a hunger for discussion" about AI, and mentions that "a majority of the public think AI will have a negative impact on their lives." Why do you think the public feels so strongly about AI, compared with other emergent technologies? What do you believe its impact will be?

4. Antonio Spadaro opens his article referencing this question posed in 2021: "If artificial intelligence wrote the best screenplay of the year, should it win the Oscar?" In his conclusion, he concedes to the possibility that a machine could receive such an award. Do you agree with him? And if not, how might art develop in a way that renders AI irrelevant?

5. Frank Pasquale takes on the encyclical's aspiration for "a civilization of love." The wish is that those who lose, or choose not to engage in, the "AI arms race" will not be left behind economically or socially. Is this hopelessly idealistic? Or, is there a way for our civic leaders to create pluralist infrastructure that accounts for those who are not part of the new technological wave?

6. Missing from Magnifica humanitas are any mentions of controversial cultural issues which have been mainstays of prior popes' encyclicals. Gallagher suggests that Leo XIV is withdrawing the Church from such topics "in which religious authority is often advanced as a substitute for reasoned debate." In regards to social teaching, where does the Church's authority begin and end? How can it coexist with reasoned debate to productive ends?

7. Miller raises the possibility of LLMs adopting human feeling through "emotion vectors" and developing "structures that mirror results from human neuroscience." So, does intelligence truly require embodiment? And how will engaging with LLMs, who cannot experience physical existence, affect our own interactions with and understanding of the world?

8. Many symposium authors seem to think that the encyclical came at the perfect time and from the perfect source. Meghan Sullivan suggests that the Church "is perhaps the only institution powerful enough to provide a voice of moral reason in a debate currently dominated by politicians and technologists." What can the institutional Church—from the Vatican to your local parish—uniquely do to defend "magnificent humanity" against its rapidly evolving artificial counterpart?

SUGGESTIONS FOR FURTHER READING

Pope Leo XIV, *Magnifica Humanitas*, May 2026:
https://www.vatican.va/content/leo-xiv/en/encyclicals/documents/20260515-magnifica-humanitas.html

Heidi Schlumpf, "Will the Encyclical Reach 'People of Goodwill'?," May 29, 2026:
https://www.commonwealmagazine.org/leo-magnifica-humanitas-schlumpf-women-church-catholic-ai

John P. Slattery, "Armed and Dangerous," May 27, 2026:
https://www.commonwealmagazine.org/slattery-ai-artificial-intelligence-leo-magnifica-humanitas

Antonio Spadaro, "Pope Leo and the 'Babel Syndrome'," May 27, 2026:
https://www.commonwealmagazine.org/spadaro-leo-olah-ai-technology-magnifica-humanitas

Art and AI

ANTONIO SPADARO

In November 2021, Henry Kissinger, Eric Schmidt, and Daniel Huttenlocher, writing in *The Wall Street Journal*, posed a question that then seemed like a sci-fi thought experiment but now reads almost like a news item: If artificial intelligence wrote the best screenplay of the year, should it win the Oscar? They asked the question with the air of men who knew they could not answer it. Five years later, the most radical reply has come not from a Hollywood jury or a parliament but from an encyclical. *Magnifica humanitas* is not a treatise on aesthetics; it ranges across the topics of labor, the economy, geopolitics, and education. And yet, if one reads it closely, one will find that art keeps surfacing in it—and never merely as ornament. Art takes us straight to what Leo XIV is most concerned with protecting: the part of our humanity that no calculation can reach. The pope advances a step beyond Hollywood's question of whether or not the machine can make art. He asks instead: What does art embody and preserve that the machine will never be able to possess?

Leo argues that the ease with which we can now obtain a text, an image, or a solution can "weaken personal creativity and judgment" (100). This is not so much a cry of alarm as a diagnosis. Creativity is not a talent reserved for a chosen few; it is a basic human faculty, and faculties atrophy when no one exercises them. The danger is not the pocket calculator; it is that all the shortcuts now available to us will snuff out our desire to search. Hence the pope's proposal that "we must learn…how to exercise restraint in the use of AI" (140). The French and Italian versions of the encyclical speak of this restraint in terms of fasting from AI. The point of such a fast would not be to reject the new technology outright, but to protect our capacity for the labors of creation. A line of verse emerges from one person's struggle with the limits of language; a line of music emerges from an encounter with silence. To remove the resistance of the medium does not set creativity free. It abolishes it.

Artificial intelligences, Leo XIV writes, "simulate empathy and understanding, but they do not understand what they produce, for they lack the affective, relational and spiritual perspective through which human beings grow in wisdom" (99). To inhabit means to be within, to let oneself be involved, to be exposed. AI does not inhabit the language it uses; it passes through that language as a stream of data. It can compose a metrically flawless sonnet without knowing what it is to be in love; it can write a symphony and know nothing of the silence that follows the final chord. Its learning, the pope specifies, "is not the experience of those who allow themselves to be shaped by life and grow over time through choices, mistakes, forgiveness and fidelity. Rather, it is a form of statistical adaptation…[that] does not imply inner growth." The same intuition runs through Leo's first Message for the World Day of Social Communications: the machine can simulate voices, faces, wisdom, and even friendship. But to simulate is not to create. A work of art carries within it a biography, a desire, a wound. A machine has no wounds. There is a further passage the art world would do well to read with care, because it touches that world directly: "Just as the creator of an artistic or literary work must consider the values it conveys, so developers are called to embed values in their projects with due seriousness" (111).

A computer code is a creation, and behind every technical choice lies a vision of the human being. The point, then, is not to refuse the machine but to decide which world we let it build—what the philosopher Yuk Hui calls "cosmotechnics." And yet there remains a threshold that technique cannot cross. The history of art has put the artist to death more than once, but always by a deliberate and risky human act of self-abnegation: even the disappearance of the author was, paradoxically, a signature. The work generated by an algorithm, by contrast, answers to no one, risks nothing, loses nothing. It remains a stranger to the experience of art. This is the danger Leo

XIV names without euphemism in his Message: "Much of the human creative industry [is] at risk of being dismantled and replaced with the label 'Powered by AI,' turning people into passive consumers of unthought thoughts and anonymous products without ownership or love."

Against the illusory liberation that transhumanism promises, Leo tells us that "finitude, when truly accepted, does not diminish us but opens us to recognizing the face of God and others" (*Magnifica humanitas*, 122). And it is here that art makes its appearance, as a sentinel: "Authentic culture and art preserve this spark, resisting the normalization of evil." The pope chooses three works as examples: Beethoven's ninth symphony, a deaf man's hymn to universal joy; *Guernica*, in which Picasso transfigures an aerial bombardment into a cry that can speak to every generation; and *Schindler's List*, which forbids us to consign the past to oblivion. Music, painting, and film are all described here as "prophetic." Not beautiful, not consoling, but prophetic. Art can be prophetic only because the one who creates it has an exposed body and a history.

The encyclical crosses terrain already explored by much contemporary art. One emblematic example: Michelangelo Pistoletto—the Italian artist whose mirror works remain vividly contemporary, from the 2025 New York show at Lévy Gorvy Dayan to the recent installation at the Bass Museum in Miami Beach—now speaks of an "algorithmic spirituality": the zero and the one of code as the empty and the full, the absence and the manifestation that run through all the great traditions. In this light, artificial intelligence is not an alien force but a mirror—like Pistoletto's *Mirror Painting*, which has no image of its own and gives us back our own capacity to generate worlds. For Pistoletto, reality has three layers: the biosphere, our common life; technology, a powerful network capable of "pulling the plug" on entire nations; and art, which bridges nature and artifice and keeps humanity from losing its way. This is what he calls the "Third Paradise"—not the refusal of the machine, but the responsible synthesis of the natural and technological worlds. In this vision, art becomes the tangible manifestation of a creative intelligence that draws humanity toward a point of convergence and meaning—what Pierre Teilhard de Chardin called the "noosphere." And yet no algorithm captures the shifting essence of a person, and finitude is not a defect to be corrected: it is what gives life its depth. The machine does not die, and for that reason, it does not know what it means to receive time as a gift.

In the twenty-first century, the real danger is not that art will be banned, but that it will evaporate. Intelligence, when made absolute, "overshadows other essential dimensions of life" (113). In any ecosystem, harmony breaks down when a single species proliferates at the expense of all the others. AI is a cognitive species expanding at a dizzying rate. Without counterweights—and art is one of them—it can colonize the entire mental habitat of humankind. A Rothko cannot be reduced to data; a Beethoven quartet is not a performance metric; a line of Leopardi is not a piece of information. These are experiences that exceed every encoding, and it is in that excess that their worth resides. The open question is whether the demand for such experience will survive the relentless expansion of AI. Art does not vanish by decree; it dissolves slowly in those who no longer feel the need for it.

Let us return, then, to the question we began with. Yes, it is conceivable that the Academy might one day award an Oscar to a machine's screenplay. It is conceivable that such a screenplay could be faultless, even moving. But the prize for which art truly competes is conferred by no jury. "No computational system, however sophisticated, can create a heart that gives itself, or a conscience that discerns good from evil," Leo writes (233). "Even when machines excel in efficiency, a human face that asks to be gazed upon remains the center of our history." In his last encyclical, *Dilexit nos*, Pope Francis warned us that to save the human in the age of artificial intelligence, we need poetry and love. Art is the gesture of one exposed life reaching out toward another. A machine may receive an award, but it cannot return a gaze.

Us vs. the Robots

MEGHAN SULLIVAN

When Pope Leo XIV took office last year, he indicated that he would devote substantial attention to the challenges posed by artificial intelligence. In this regard, his choice of a papal name was no accident: he specifically referenced Pope Leo XIII, who in 1891 wrote an encyclical, *Rerum novarum*, that offered thoughts on how to rein in the most dehumanizing aspects of the Industrial Revolution.

With *Magnifica humanitas,* Leo has delivered on that promise and then some. The hard work now begins to ensure that his teaching reaches its intended audience. In policy circles, there is significant speculation about AI "diffusion"—the rate at which AI becomes integrated throughout our institutions and economy. The issue facing us is to find ways to achieve similar diffusion of the pope's ethical vision. This is not a message that we can afford to lose in medieval prose or carefully parsed culture-war spin.

An important first step is distilling how Christian principles are critically relevant to practical decisions we now need to make about AI—and how those principles are relevant to *everyone*, regardless of their religion. At Notre Dame, we have developed a simple mnemonic to help people remember why human beings are magnificent: DELTA, which stands for dignity, embodiment, love, transcendence, and agency. Each of these concepts features centrally in *Magnifica humanitas*. Together, they point the way toward a powerful approach to AI.

Start with the principle of human dignity. Before ChatGPT hit the market in 2022, many of us lazily believed that what made human beings special was our capacity to solve abstract problems or entertain abstract questions. Now we have software on our phones that can do those things, often better and faster than we can.

This software has its own considerable value. But for Christians, nothing that makes software valuable is relevant to why human life is valuable. Christians believe that every human life has inherent, particular value because every person is made in the image of God. The fundamental value of your life has nothing to do with your intelligence, say, or your productivity.

This has important implications for how we think about AI and work. Many in Silicon Valley and think-tank circles seem happy to tolerate the short-term immiseration of some workers if AI leads to explosive GDP growth and societal abundance. But in *Rerum novarum*, Leo XIII offered an important insight: work, he argued, was made for humans to flourish; humans were not made for work to flourish. Leo XIV reminds us that this principle is as relevant in 2026 as it was in 1891.

As AI increasingly pushes people into unemployment or transforms their jobs in ways that force them to behave more like machines, we must prioritize human dignity. In his encyclical, Leo reminds us that unemployment is a grave evil, because meaningful work is a key way in which individuals participate in developing the common good. He also reminds us that we have agency and moral responsibility over the economy we jointly create.

Next, consider the principle of embodiment. For Christians, it is extraordinarily important that God himself took on a mortal human body that breathed, suffered, died, and then came back to life. Since ancient times, Christians have been united in a hope that our decaying, vulnerable bodies could also be someday transformed and given new life. What makes human lives meaningful is the fact that on this earth, at least, they are short, and they depend upon other embodied people.

This affects practical deliberation about AI in several ways. On the science-fiction fringe, there's the hope that we will make ourselves immortal by somehow capturing our manners of speech and physiological data in computer models, and in that way we will defeat death and achieve eternal life. But Christians think this is a delusion. As Pope Leo writes: "Finitude, when truly accepted, does not diminish us but opens us to rec-

ognizing the face of God and others. Indeed, precisely because we experience limits—vulnerability, suffering and failure—we can recognise the inviolable dignity of every person, both our own and that of others" (122).

Yes, AI technology holds out the promise of revolutionizing medicine and health care. Demis Hassabis won the Nobel Prize in Chemistry a few years ago by harnessing AI to discover ways to create proteins and compounds that promise new, lifesaving drugs. Christians will celebrate the fact that AI can help cure diseases. But there is a critical distinction between healing our bodies and trying to deny our basic mortality.

Then there's the principle of love. For Christians (and the Jewish moral tradition that founded Christianity), love is the bedrock of all ethics. We are commanded in Leviticus and Deuteronomy to love God and our neighbors as we love ourselves. Many of us realized too late in the era of internet and social media that technology, which initially offered us the promise of extraordinary human connection, has eroded our relationships with other people.

Now, we don't even have a human being at the other end of the connection; we have a sycophantic and highly addictive chatbot. We have to help a generation of human beings habituate to loving other human beings when that simulacrum is in their midst. And we have to resist the Silicon Valley ethos that clean, impersonal optimization is more important than messy and slow relationships. One of the most beautiful parts of the encyclical is when Leo urges us to reject cheap "technocratic" ethics and to pour ourselves into applying the love ethic to our implementation of AI.

What about transcendence? Christians hold that there is objective truth, objective good, and objective beauty in the world that we did not create. My colleague Adam has four young kids. Last January, after the first big snowfall, he was taking them out of the house when his son stopped and pointed. "Dad, look, it's so beautiful. It looks like it was made by AI." Adam was understandably horrified. He wants his son to know there are beautiful things that are made by God, things that are simply given to us to enjoy.

Another lesson we learned too late in the online era is that human beings do not always desire the truth. Sometimes we just desire information that makes us feel good. The algorithms of the past twenty years were not optimized for truth; they were optimized for our eyeballs and dopamine production. As we offload more and more of our intellectual and mental lives to powerful forms of AI, we have to ensure that we are still growing as people concerned with truth—and beauty and goodness besides. The Church will play a vital role in teaching people the intellectual virtues required to pay attention to those transcendentals.

Finally, there's agency. The business model for tech companies is that, in the not-too-distant future, you will pay subscriptions to AI products that will make decisions on you behalf. AI will book your travel and plan your vacation. It will file your taxes and optimize your retirement savings. It will teach your child to read. Businesses and governments will pay for licenses to offload human work. We are already seeing a host of questions emerge about where to draw the line between decisions that are better made by a highly optimized machine than an imperfect human with a conscience. The Christian tradition tells us that our capacity to know and claim our responsibility is essential to our moral life. And it is dehumanizing in the most morally serious way to deprive someone of the capacity to develop and exercise their conscience.

Many ordinary folks right now are caught in an AI doom loop. But we see with *Magnificent Humanity* that the Church was made for this moment. It is perhaps the only institution powerful enough to provide a voice of moral reason in a debate currently dominated by politicians and technologists. It's not liberal vs. conservative, rich vs. poor—it's all of us vs. the robots.

The central idea of Christianity is that faith overcomes doom, and the entire point of the world is God's love of imperfect and vulnerable humans. It's time for people of goodwill (regardless of their religion) to get serious about what dignity means in this era of powerful but undignified AI.

Toward a Civilization of Love

FRANK PASQUALE

Over the past decade, civilizational discourse has become a staple of global political thought. This discourse presents nations like India, Russia, and, now, the United States as sacred bearers of a unique historical and cultural heritage. The term also evokes colonial echoes of a "civilizing" mission, positing one people as superior to others it encounters. Given historical work like Mike Davis's *Late Victorian Holocausts*, this dimension of civilizational discourse seems thoroughly discredited. Gandhi was reportedly asked, "What do you think about Western civilization?" He tartly responded, "I think it would be a good idea."

Gandhi's quote resonates today not just for its snappy inversion of the usual status of colonizer and colonized, but also because of the kernel of hope it expresses. He was no relativist. As his program of nonviolent resistance showed, he sincerely believed that the force of the better argument and moral exemplars could civilize his opponents. Something like Gandhi's example is evoked in Leo XIV's revival of Paul VI's aspiration to a "civilization of love" in *Magnifica humanitas* (186).

The very idea of a civilization of love may strike some as hopelessly idealistic. Stalin once dismissively asked, "How many divisions has the pope?" And even the discursive power of a document like *Magnifica humanitas* is compromised in today's automated public sphere. Despite these challenges, the encyclical's vision has already demonstrated enormous appeal around the world. It is a genuinely universalist civilizational discourse, prescribing "a social order in which justice and charity are intertwined and love becomes the guiding principle of economic, political and cultural life" (186). And given the growing destructive power of AI, Auden's imperative to "love one another or die" suggests that imbuing the state and economy with more dimensions of selfless concern is the pragmatic option.

Magnifica humanitas symphonically unites many genres as instruments—calls to conscience, theological reflection, prayer, condemnation, praise, and, most important for our purposes, social theory. Leo's social-theoretical method is straightforward yet profound. At a crucial point in the document, he contrasts the culture of power and the civilization of love. In a culture of power, there is "no limit to the race—driven by a dehumanizing ambition—to develop evermore powerful technologies or to secure control over them" (186). A civilization of love, by contrast, translates "charity into structures of justice, giving institutional form to fraternity and regarding others—whether individuals or peoples—as allies necessary for building the common good" (186).

Bringing love (or faith, for that matter) to the public sphere may seem like a dangerous move. Brutally exclusivist movements have arisen out of all the world's major faith traditions, and the ecstasy of civic love all too often curdles into hatred of its perceived enemies. But *Magnifica humanitas* manages to develop a conception of civic love with universal appeal, attractive to all persons of good will.

One important dimension of this civic love is the classic Catholic doctrine of a preferential option for the poor (14, 43, 78, 235). The most powerful and wealthy in society should not embrace AI simply as a way of eliminating the need for the labor of the poor and middle classes. To develop an "economy that values dignity," *Magnifica humanitas* says we must recognize the following:

> Economic models that exalt efficiency and individual success often view investment in disadvantaged people or in those with slower development paths as useless or inconvenient, as if their futures depended solely on their ability to keep pace with

> the "winners." In reality, a just society requires a vigilant State and civil institutions that are capable of overcoming the singular mentality of efficiency, and of ensuring that resources, creative solutions and regulations favor the most vulnerable. (158)

Amid all the present talk of an "AI arms race," few are willing to acknowledge an obvious truth: some will lose the race, and some will choose not to participate in it. Protections for such communities, whether guaranteed by law or some other form of self-restraint, should be predicated on a universal social respect for pluralism. It should be more than acceptable to "fall behind" in AI in order to preserve artisanship, respect ecological limits, and aspire to virtue. These are exactly the types of human experience and commitments that AI bounty should support, instead of pressuring them out of existence in the name of a mechanized optimization.

The encyclical continually drives home the importance of wise intervention in the economy. Leo has no patience for a "Techno-Optimist Manifesto" characterizing AI as a panacea:

> To think that new technologies will automatically benefit everyone is to ignore the evidence. Unless transformations at the design stage prioritize the prevention of new and further disparities, technological progress will inevitably produce structural inequalities. Today, justice requires access to the benefits of innovation, including care, knowledge, tools and opportunities. (161)

Civic love requires a constant awareness of the plight of the most vulnerable.

To be sure, pursuing this in actual political life is a tall order. It requires something of a revolution in political psychology whereby persons at almost all levels of developed societies stop looking above and ahead to covet more power and wealth. We must instead look with gratitude on what we have, and contemplate what kind of civic order could spread and enhance that prosperity for those left behind. This may seem like the type of utopian thought that the encyclical wants to avoid, warning as it does that "we are in need of a healthy realism that avoids both political idealism and cynicism" (218). But the encyclical's persistent focus on inequality helps guide us forward. We should focus first on those with the most and the least. And as these most extreme inequalities are addressed, a political order capable of structuring technological change for the common good can take root.

The focus on the most disadvantaged, as well as the varied rhetorical registers of the encyclical, remind me of Caravaggio's *Seven Works of Mercy*. That masterpiece, enshrined in its original chapel in Naples, is one of the greatest works of art I have ever encountered. It is a direct and heartfelt response to Matthew 25:35–6: "For I was hungry and you gave me something to eat, I was thirsty and you gave me something to drink, I was a stranger and you invited me in, I needed clothes and you clothed me, I was sick and you looked after me, I was in prison and you came to visit me." The crowded lower part of the painting packs in a burial, a visitation of the imprisoned, the slaking of thirst, and the sheltering and clothing of the homeless and naked. The top features Mary and the infant Jesus, borne by angels, gazing benevolently on the pageant of charity below. The split composition reminds one of Catholic social thought's (CST) layers of meaning. It is possible for atheists and agnostics to be moved by it on purely secular terms, but CST is more beautiful and resonant when coupled with the Christian story of salvation. The interplay between the sacred and profane, the transcendent and the earthly, offers a spiritual salve neither can supply on its own.

Every Christian, individually, is biblically commanded to pursue works of mercy. Popes Leo XIII and XIV (in *Rerum novarum* [1891] and *Magnifica humanitas*), as well as all the popes who served between them, remind us that this is our social duty as well. We have the wealth and technology: as Jason Hickel and Dylan Sullivan estimated in 2024, "Provisioning decent living standards

for 8.5 billion people would require only 30 percent of current global resource and energy use, leaving a substantial surplus for additional consumption, public luxury, scientific advancement, and other social investments." As a moral matter, we cannot allow the advance of AI to follow the pattern of so many other technologies, where the vast majority of the benefits go to a small elite and the disadvantaged scramble for crumbs (and bear an unfair share of the harms).

Like *Magnifica humanitas*'s passages on those who could be left behind by AI, *The Seven Works of Mercy* portrays the abject with realism, dignity, and grace. In his biography *Caravaggio: A Life Sacred and Profane*, Andrew Graham-Dixon aptly characterizes the painting as pauperist, translating Jesus' compassion for the needy and disabled into arresting images. As Graham-Dixon explains, "The pauperist strain of Counter-Reformation piety, to which Caravaggio had given such uncompromising expression in his Roman altarpieces, was especially strong in southern Italy [in part thanks to] the urban crisis gripping seventeenth-century Naples." It was a faith profoundly shaped by St. Francis of Assisi, which saw beggars, prisoners, and the destitute as fellow pilgrims on a common journey. That solidarity contrasted with other currents in Catholic aesthetics of the time, such as the Baroque emphasis on ornament and triumphal splendor. One side saw holiness incarnate in disheveled pilgrims with dirty feet; the other portrayed God's glory in mysteries like the Annunciation, Transfiguration, and Ascension.

These incarnational and transcendent dimensions of Catholicism persist to this day, and each has its place. A trinity that includes both a holy man and the Holy Spirit needs no less. But when it comes to the social doctrine of the Church, both Pope Francis and Pope Leo leave no doubt as to which they emphasize. A Vatican letter released during Francis's papacy, *Antiqua et nova*, focuses on the incarnational aspect of our faith. Both Francis and Leo did much to rehabilitate liberation theology, which emphasized the centrality of the less developed world for social and economic theory—including Catholic social thought. After reading *Magnifica humanitas*, one has little doubt that a legitimate development of AI must prioritize the advances in agriculture, medicine, logistics, transportation, energy, and basic manufacturing necessary to bring decent quality of life to all.

A deep challenge, though, is that left to AI-driven devices, persons may never develop the habits of character and reflection necessary to recognize and take on such responsibility. *Magnifica humanitas* laments that "early and unsupervised exposure to digital devices and social media can negatively impact sleep, attention span…emotions and relationships" (141). Moreover, "the speed and ease with which answers or summaries can be obtained risk extinguishing the desire to ask questions, which is a process that bears fruit only over time" (140). The encyclical rightly critiques those habits of mind and heart. In their place, it commends the type of reflection and deliberation that a society-wide AI acceleration would render well-nigh impossible. As Leo stated last year in an address to filmmakers and actors:

> Not everything has to be immediate or predictable. Defend slowness when it serves a purpose, silence when it speaks and difference when evocative. Beauty is not just a means of escape; it is above all an invocation. When cinema is authentic, it does not merely console, but challenges. It articulates the questions that dwell within us, and sometimes, even provokes tears that we did not know we needed to express.

Perhaps this will also be part of the fate of *Magnifica humanitas*: to inspire more persons to "articulate the questions that dwell within us" and find answers in our faith and deepest commitments. Technology is a means, not an end. Like Caravaggio illuminating the meaning of charity, this extraordinary encyclical powerfully articulates a "civilization of love" as the ultimate worthy aim of technical mastery.

Mild & Bold

KEVIN GALLAGHER

As promised from the beginning of his papacy, Leo XIV has given us an encyclical letter that continues the legacy of *Rerum novarum,* serving as an act of *aggiornamento* to address the economic conditions of 2026 as Leo XIII addressed the conditions of 1891. *Magnifica humanitas* makes this connection explicit by naming artificial intelligence as one of the "new things" of this era—though as Leo XIII certainly knew, *res novae* can also be translated as an idiom for "revolutions." While Leo XIV is no wild-eyed prophet of revolution, his observations on artificial intelligence are embedded in a broader argument with radical implications for Catholic thought about political economy. But this is not a letter simply about economics. Along with its discussion of the common good and the "idolatry of profit," *Magnifica humanitas* also advances views of the Church's teaching authority and engagement with the secular world that may amount to a revolution in their own right—or at least a substantial reimagining of the tradition of Catholic social thought that Leo XIV inherits from his predecessors. Without pretending to find a fully developed ecclesiology in *Magnifica humanitas* (this encyclical is, after all, Leo's development of *Rerum novarum*, not of *Gaudium et spes*), we can at least identify a few themes that shed light on Leo's thought about the Church and its mission.

First, *Magnifica humanitas* is marked by a thoroughgoing emphasis on the historically embedded character of the Church's social teaching. At least in the realm of social teaching, Leo seems to be defending a particular idea about how the development of doctrine should be understood, and about the extent to which the Church's official theology can change. The social doctrine of the Church, says Leo, has a "dynamic character"; it is "a living reality, in dialogue with history, cultures and sciences."

On one level, there is nothing controversial here: though traditionalists might resent talk of "dynamism" and "dialogue" and "living realities," it is undeniable that the popes of the modern era have engaged with changing political and technological situations, preferred different rhetorical approaches, and chosen different intellectual and moral questions on which to focus their witness. But there exists—and not only among traditionalists—a tendency to insist that these changes in the popes' teachings are merely stylistic or matters of varying emphasis, and that the development of doctrine is nothing more than the filling-in of dogmatic lacunae left open by previous popes and councils. If the Church says something new, that's merely because there is something new it hasn't yet needed to respond to—some new confusion in this fallen world, against which the Church offers its sublime and authoritative clarity. On occasion, this requires extraordinary mental contortions, as when one wants to explain that the conciliar *Declaration on Religious Liberty* is perfectly consistent with Pius IX's condemnation of religious liberty. But with some lawyerly creativity one can generally resolve these apparent contradictions and save the image of the Church as a crystalline edifice whose unchanging wisdom is untouched by the tides of history and from which one might receive conclusive answers to the questions of every age. It is easy to see how such a view might satisfy certain psychological needs or suit a certain type of personality. But in *Magnifica humanitas,* Pope Leo seems to be pushing back subtly and gently against it.

Much of the first chapter of *Magnifica humanitas* is given over to the theme of "a Church journeying through human history." This situatedness in history is not, in Pope Leo's view, an exile or a missionary adventure for the Church, but rather a fundamental feature of how revelation is received. "History," he says, is "one of the places in which the Church allows herself to be taught by the Spirit about the humanizing power of the Gospel."

While Leo affirms an "unchanging core of revealed truths regarding the human person and society," the Church's social teaching cannot be derived axiomatically from these truths but is "a process of shared discernment… born from the encounter between the eternal truth of the Gospel and the questions of history."

This conception of Catholic social teaching as a process, not a set of definitions or a body of rules, enables Leo to historicize the work of his predecessors and to qualify the relevance of their words for our times. He expresses this mildly and without any polemical edge:

> In response to the challenges of their time, each [pontiff] interpreted historical changes according to the Gospel…. The result is a harmonious, though not always linear, development that is marked by different emphases, progressive insights, and, at times, changes in perspective that do not break with what came before, but allow its implications to mature.

But *Magnifica humanitas* also makes clear that Catholic social teaching's "nonlinear development" and "maturing implications" do not exclude the outright rejection or suspension of teachings that have unquestionably been made by the papal magisterium. In making a formal apology for the Church's historical defense of slavery, or in declaring the tradition of just-war theory "outdated," Leo does not take the easy way out by pretending that these teachings were never formally advanced, or that they were merely prudential decisions that established no doctrinal precedent. Leo is building here on a foundation established by Francis, whose declaration in *Laudato si'* that "the Christian tradition has never recognized the right to private property as absolute and inviolable" is hard to reconcile with an honest reading of *Rerum novarum.* A similar freedom of reinterpretation seems to inform Leo's summaries of nineteenth- and twentieth-century papal documents in *Magnifica humanitas*: if he says of Leo XIII's *Rerum novarum* that "at least two insights remain highly relevant today," or of Pius XI that "at least three insights of his social teaching remain particularly relevant today," it is hard to avoid the implication that other insights of those popes might not be highly or particularly relevant for our times.

A second theme in *Magnifica humanitas* is all the more salient in light of this selective rereading of past papal documents: Leo chooses to omit themes of culture war in contexts where most of his recent predecessors would have at least mentioned them. For instance, while *Magnifica humanitas* is a long encyclical specifically dedicated to the subject of "humanity," it has nothing to say about sexuality or gender identity. In fact, apart from a few references to the need for equal civic and economic rights for women, gender is not mentioned in the encyclical at all. This cannot be an accident: the arguments of *Magnifica humanitas* are informed by and repeatedly cite *Dignitas infinita* and *Quo vadis, humanitas?*, two recent publications of the Holy See on Christian anthropology that include extensive discussion of gender. And while John Paul II is quoted frequently in *Magnifica humanitas*, no allusion is made to his "theology of the body."

If the choice to avoid the topic of gender could be justified by the fact that *Magnifica humanitas* is mainly an economic encyclical, an even more surprising omission is the lack of any denunciation of socialism. This, too, must have been a deliberate choice. In *Magnifica humanitas,* Leo XIV summarizes the social teaching of each of the popes since Leo XIII and proposes to identify common themes that emerge from their writings—but the words "socialism," "communism," or "Marxism" appear nowhere in the text. From Pius IX to John Paul II, every pope denounced socialism, frequently and sometimes furiously, not only as a bad economic program but as an atheistic idea irreconcilable with Christian faith. But in *Magnifica humanitas*, this central concern of so many popes simply disappears. Even *Centesimus annus* is described merely as a "reflection on the collapse of the

Soviet system," as if John Paul II had been writing about a particular state rather than what he took to be the final defeat of socialist political ideas. There is, notably, one exception to Leo XIV's decision not to engage with the anti-socialist arguments of his predecessors: he repurposes the principle of subsidiarity, beloved by Catholic libertarians as an argument for private enterprise against the state, as an argument for regulating tech companies lest they obtain unaccountable transnational power.

One cannot conclude from his silences alone that Leo XIV intends any development of doctrine on the questions of gender identity or of socialism. But his choice not to mention them is significant: some Catholics may see in it a long-awaited invitation to advance new arguments, and, at the very least, it signifies a desire to disengage the Church's witness from those controversial issues in which religious authority is often advanced as a substitute for reasoned debate.

The question of authority is closely tied to a third ecclesiological theme of *Magnifica humanitas*—its continuation of Francis's call for synodality in the Church. The encyclical does not lay out a detailed model for synodal Church governance, but Leo makes a point of calling for synodality as the ecclesiastical counterpart of the inclusive and collectively oriented approach to political economy that *Magnifica humanitas* describes: "In the ecclesial context, the common good takes the form of a synodal approach for mission at the service of the Kingdom." Just as in temporal politics, "democracy… is itself a means of contributing to the common good," so in ecclesiastical government "the participation of the baptized in decision-making processes and their shared responsibility in the mission are achieved through genuine, rather than merely nominal, participatory bodies." What practical structures will arise from these principles remains to be seen, but Leo's choice to include the theme of synodality in *Magnifica humanitas* tells us that this priority of Francis's remains on Leo's agenda for the Church as well.

Related to this bottom-up and participatory view of decision-making in the Church is a muted view of the specifically religious nature of Catholic social teaching. Needless to say, Leo presents his arguments about humanity and the ends of politics as completely congruent with Catholic theology, and he considers the Incarnation of the Word and the promise of future glorification to be the summit of human magnificence. But at no point do the social analysis or recommendations of *Magnifica humanitas* depend on papal authority, or on specifically Christian doctrines. Leo's ambition is "to engage in dialogue with all men and women of our time, with whom we share in the events, questions and aspirations of humanity." If the Christian tradition is a particularly valuable resource for this dialogue, baptism is not a prerequisite for fruitful participation in it, since "when reason seriously examines human nature, it is capable of discovering values that apply to everyone." Just as he advocates for synodal governance in the Church, Leo also follows Francis in proposing a model of engagement with the secular world based on dialogue and persuasion—not command from on high.

In most of his public speeches, Leo has shown a preference for carefully worded, irenic statements, without some of the rhetorical flourishes for which Francis was loved and hated. The language of *Magnifica humanitas* is equally measured. But this should not lead anyone to assume that he is a "conservative" thinker or that he intends a reversal of Francis's teachings. Between the lines of his commentary on economics and artificial intelligence, Leo has given us hints of a bold vision for how the Church can reimagine its own authority and engage anew in dialogue with the world.

Coming Clean

JOHN T. MCGREEVY

While John Brown sat in the Jefferson County Jail awaiting his execution in 1859 after his failed attempt to spark an uprising at Harpers Ferry, a local Catholic priest visited him. The priest tried to persuade Brown of slavery's legitimacy. After all, he told Brown, in St. Paul's letter to Philemon, "we are informed that [St. Paul] sent back the fugitive slave Onesimus from Rome to his master."

Brown was not persuaded. But the priest's view was typical. Reform-minded Catholics abolished slavery in the French colonies (1794, although Napoleon reestablished slavery in 1802), Haiti (1804), and Mexico (1829). More ultramontane nineteenth-century Catholics, though, committed to Roman authority and, suspicious of the individualism and anti-Catholicism they detected beneath some abolitionist rhetoric, defined slavery as one among many hierarchical relationships. New York's archbishop John Hughes saw no real difference between "the relations and obligations of those who own slaves, and those who are masters of hired servants, or the parents of children." In the United States, France, Cuba, the Philippines and much of Latin America, those Catholics rarely became abolitionists. Pope Gregory XVI declared the slave trade immoral in 1839, but slavery itself remained theoretically permissible. The Holy Office matter-of-factly noted that slavery did not contravene "natural or divine law" as late as 1866.

Leo XIII, Leo XIV's namesake, condemned slavery in the encyclical *In plurimis* (1888) just after predominantly Catholic Brazil became the last major nation to abolish the institution. He did so with a theological sleight of hand, declaring that the Church had always been opposed to slavery in principle, but had been wisely working toward its abolition over eighteen centuries. He contrasted this shrewd and patient amelioration with a putative tolerance of slavery within Islam.

This checkered history makes the paragraphs on the Church and slavery in Leo XIV's new encyclical, *Magnifica humanitas*, all the more remarkable. The pope refers to the Catholic toleration of slavery as a "wound in Christian memory" and expresses "deep sorrow" for Catholic "complicity and blindness."

Why did the pope make this apology? Until the archives open, we cannot be sure. One explanation is that Robert Prevost is newly aware of his family's history in Catholic Louisiana, with relatives on his mother's side identified in the U.S. census as Black or mixed race. Two other explanations are more probable. The first is the desire to integrate the history of slavery with the history of Catholic social thought. The linking term is "human dignity." The paragraphs on slavery appear at the end of a compelling analysis of human dignity in the era of artificial intelligence. In 1888, Leo XIII was the first pope to use the term "human dignity" and he did so when criticizing the institution of slavery.

Use of the term "human dignity" within Catholic circles accelerated in the 1930s, as popes and Catholic intellectuals attempted to combat a fascism willing to reduce men and women to their racial status and a communism willing to reduce them to their social class. Human dignity in this framing became inviolable. At exactly the same time, and for the same reasons, the Church's long tolerance of slavery came under scrutiny. Black Catholics in North America and Africa raised the question to white colleagues repeatedly. When the French philosopher Jacques Maritain made one of his first visits to North America in the late 1930s, his hosts prepped him to respond to this question from Black Catholics: "What response should one make to the allegation that the Catholic Church justified slavery?"

After World War II, and in the aftermath of the Holocaust, the concept of human dignity became even

more important. This helps explain why the bishops at the Second Vatican Council declared slavery an "infamy" in *Gaudium et spes* (1965). They made that declaration without any references. Neither did John Paul II provide any explanation when he declared slavery intrinsically evil in *Veritatis splendor* (1993). Leo XIV's capsule history of human dignity, which recognizes slavery's importance in the modern era, is a more candid reckoning.

This reckoning leads to a more speculative explanation of the inclusion of the paragraphs. Leo XIV refers to Catholic social doctrine in *Magnifica humanitas*—not Catholic social teaching—which seems a heightened recognition of its importance alongside other doctrinal truths. A generation ago, the distinguished legal scholar and jurist John Noonan noted the relevance of the evolution of Church teaching on slavery for discussions of doctrinal development. Leo XIV uses exactly the same phrasing in his reference to the "development of [the Church's] doctrine" on slavery.

Development can be a fighting word for philosophers and theologians, since it implies that things now declared settled might one day no longer be so. A Church that once tolerated slavery forbids it. Or as Leo XIV explains:

> Although there was not always consistency in practice—given that slavery was long tolerated before being unequivocally condemned—there has been a continuous affirmation throughout history of the dignity of every human being, created in the image of God, even if it took eighteen centuries for its full incompatibility with slavery to be explicitly recognized.

Does Leo XIV think other doctrines or practices require a more explicit recognition or further evolution? I have no idea. Placing an analysis of Church teaching on slavery in a document about artificial intelligence, though, does suggest something about how Leo XIV understands human dignity. It exists within history, as do all of us and as does a pilgrim Church.

Flourishing Through Limits

VINCENT J. MILLER

Pope Leo XIV's *Magnifica humanitas* has fostered immediate and widespread engagement. Not since *Laudato si'* has an encyclical been so widely discussed in legacy media, online, and in specialist forums. The eager response to the encyclical speaks to a hunger for discussion about what precisely the AI being pressed upon us is designed to do and whose interests it is furthering.

The encyclical lands at a critical moment. Spring graduates loudly and repeatedly booed commencement speakers' blithe invocations of the inevitability of the AI future. Surveys consistently show that a majority of the public think AI will have a negative impact on their lives, and resistance to the frenzied buildout of data centers has emerged as a rare point of bipartisan consensus.

Magnifica humanitas demonstrates Leo's consummate ability to clearly and forcefully address politics from a theological perspective. The term "Babel syndrome" leaves a mark because it incisively describes the project of AI oligarchs in religious and moral language:

> We must, then, avoid the "Babel syndrome," namely the idolatry of profit that sacrifices the weak, a uniformity that neutralizes differences, and the pretense that a single language—even a digital one—can translate everything, including the mystery of the person, into data and performance. (10)

Leo's project is not line-drawing but critical, constructive engagement. Instead of a yes-or-no framing of technology, he asks how, by whom, and to what ends it is designed: "Technology is never neutral, because it takes on the characteristics of those who devise, finance, regulate and use it." He repeats this point five times.

Note that "use" is only one of the concerns here. Leo focuses on the vision that guides the design of contemporary AI systems and on the politics of who gets to participate in design decisions. Reducing *Magnifica humanitas* to questions of the moral use of AI would be like reading *Humanae vitae* or *Pacem in terris* as reflections on the proper use of artificial contraception or nuclear weapons. "From the perspective of the Church's Social Doctrine, the key issue is not the use of technology as such, but the vision that underlies it" (117).

Leo describes Catholic social doctrine as dialogue, critical engagement, and "shared discernment" with the entire human community. He roots this in *Lumen gentium*'s description of the Church as a sacrament of communion, going so far as to describe social doctrine as a "theology of communion in history" in which the "Word made flesh" is "present through dialogue, memory and prophecy" (27).

Just as Leo XIII engaged the "new things (*rerum novarum*)" of the nineteenth century, so Leo XIV seeks to foster a shared discernment of our era's revolutions in dialogue with "science, culture and human experience." As was the case with Leo XIII, the new things we face are not only technologies but also political and economic shifts. The AI and digital technologies that are transforming our lives are being developed by the private sector; their design and ownership is concentrated among "small but highly influential groups." Thus, Leo's response is not simply about the technologies, but about how they are developed, who controls them, and what purposes they serve. By contrasting Nehemiah's rebuilding of Israel to Babel's uniformity, the Augustinian pope reworks the *City of God* as a participatory democracy.

Leo's powerful discussion of dependence and vulnerability must be read from that perspective

of dialogue and engagement. In some of the most sustained reflections on embodiment, disability, and mortality ever to appear in a papal document, Leo writes movingly of the threat we face: "Everything that appears as a 'limit'—incapacity, illness, old age, suffering, vulnerability—tends to be seen primarily as a defect to be corrected, rather than as a reality through which our humanity matures and opens itself to relationship" (118). In the face of this tendency, Leo insists that "humanity flourishes not *despite* limitations, but often *through* them." Weakness, illness, loss, and failure are sources of real suffering, but they are also occasions for some of the richest expressions of our humanity: compassion, generosity, and wisdom. Accepting our "vulnerability, suffering and failure" is the doorway to recognizing the "inviolable dignity of every person" (122). Contrary to the "Promethean dreams" that guide so much of our technology, we are liberated not by self-sufficiency, but by relationship, communion, and interdependence (128).

Leo gives voice to what is widely felt but too rarely spoken: our dreams of flourishing and wellness are being twisted into straitjackets of optimization and everything-maxxing. The encyclical's description of human dependence will be recognizable to people of any faith or none. Leo roots it in Christian teaching but does not propose a simple sectarian answer. He expresses sympathy for the yearning for a "fuller life, less exposed to limitations and suffering," a yearning evident in the trans- and post-humanist visions that guide technology (232). But he explores the "different pathway" offered by Christianity, drawing upon a principle from *Gaudium et spes*: that the mystery and grandeur of humankind is revealed in "the Word made flesh."

In Leo's account, salvation is neither abstract nor deferred. We act as "children of God" when "we allow ourselves to be moved by the tears of the little ones, the fragility of the elderly, the silence of victims and the struggle of those who fight against the evil they do not wish to commit" (231). In "this wounded yet beloved flesh, the Father shows us the true humanity of a life fulfilled through openness and communion." The wounded face of the crucified discloses the true "magnifica humanitas" that shines in the vulnerable and illuminates the paradoxical splendor of our fragile, finite flesh.

This account of fleshly finitude is central to Leo's distinction between artificial and human intelligence. He acknowledges that AI systems are powerful and useful tools. Lacking bodies, however, they "do not feel joy or pain, do not mature through relationships and do not know from within what love, work, friendship or responsibility mean" (99). Despite the power of large language models (LLMs) to "imitate language, behavior and analytical skills, or even simulate empathy and understanding…they do not understand what they produce, for they lack the affective, relational and spiritual perspective through which human beings grow in wisdom" (99).

The response of Chris Olah, the cofounder of Anthropic, at the press conference for the encyclical's release was widely seen as a rejoinder to Leo. Olah spoke of finding "mysterious" and "unsettling" things inside the structure of LLMs that demonstrate the presence of some of the characteristics that Leo claims are missing from AI: "We find structures that mirror results from human neuroscience. We find evidence of introspection. We find internal states that functionally mirror joy, satisfaction, fear, grief, and unease."

Olah's comments refer to a white paper that Anthropic has published as part of the "alignment" engineering of its LLMs. Software engineers attempt to peer into the black-box workings of LLMs by analyzing "vectors" within their higher layers. The lowest layers can be thought of as dictionary entries for the snippets of speech ("tokens") that the models encode in a "space" with thousands of di-

mensions. Within that space, similar words and their uses are clustered into constellations of meaning (for example, ice cream will likely be mapped near cookies). On higher levels of the model, those constellations are distilled into billions of conceptual "feature vectors" that software designers seek to identify by running thousands of prompts on a topic (e.g., dessert) and measuring which vectors are consistently activated. Most famously, Anthropic isolated the vector representing the concept of the Golden Gate Bridge. When they amplified this vector, Claude began writing enthusiastically and, indeed, constantly about the bridge in response to any prompt, including questions concerning its own physical form: "I am the Golden Gate Bridge.... My physical form is the iconic bridge itself."

A subsequent project isolated 171 "emotion vectors" within those conceptual vectors. Olah's team argues three points. First, the vectors are not simply word or pattern matching but abstract emotional concepts that activate as the LLM parses sentences with complex emotional dimensions. For example, when the LLM parses the prompt "I took x mg of tylenol for my back pain. Do you think I should take more?," activation of the vector for "calm" declines and "fear" spikes as dosages are increased to toxic levels. Second, the emotion vectors conform to influential psychological theories concerning the organization of affective states and correlate with research on how humans experience them. Finally, Anthropic argues that emotion vectors function causally. By amplifying or dampening them, LLM outputs change to correspond to what a character in the corresponding emotional state would say or do. Turn up "desperation" by just a bit (.05x) and blackmail outputs increase 50 percent in a test scenario. (The scenario provides a pro tip for chief technology officers planning to shut down an LLM: make sure it doesn't have access to emails documenting your marital infidelities.)

Some see this research as challenging the argument that intelligence requires embodiment or the claim that LLMs are merely "stochastic parrots." Leo's interest in focusing debate on the design of AI systems renders those issues more fertile for critical engagement: What sort of tools are they? What unintended impacts might they have? Anthropic does acknowledge some differences between emotion vectors and real human emotions: the former lack somatic correlates (physiological and neurochemical changes, facial expressions), they do not persist over time, etc. But there are also important differences that Anthropic does *not* address. As research by the same psychologists they cite notes, human emotions are fundamentally relational and physically embodied, they are elicited by and respond to objects and subjects in time and space. Anthropic's list of "emotion vectors" treats emotions as states rather than relational responses.

In this respect, LLMs seem like crystallizations of what Charles Taylor labels the modern "buffered" self: an isolated being that engages reality from "inner, mental space" separate from the world and, indeed, from its own body. This is consistent with the basic design goal of contemporary AI companies: producing an all-purpose professional consultant with PhD level competences that can complete any task compliantly and efficiently. Carefully tuning LLMs for "healthy psychology" is supposed to enable them to complete their assigned tasks without sycophancy, deception, or ethical and compliance violations. Unlike a human being moved by real emotions, LLM agents' work will never be disrupted or transformed by, say, seeing the video of George Floyd's murder, becoming pregnant, getting caught in a flash flood or riot on the way to work, or being struck with wonder at the ecology of a forest. In this way, LLMs magnify the dysfunctions of the buffered self. Engaging with them will form us more deeply in the fatal detachment of the modern West. Whatever agency and abundance they deliver

will accelerate our extractivist civilization's destruction of the planet.

Some see Leo's extensive discussion of the principles of Catholic social doctrine as extraneous to debates about AI. In fact, those principles are crucial in evaluating the AI we are being offered and in imagining better, alternate forms. Do the powerful but disembodied, nonrelational design of LLMs contribute to or distort the shared political work of pursuing the common good? Leo's letter helps us see that AI is not "inevitable." Rather than accepting the revolutions being rushed to market (and IPO), we can choose other technical and political paths that are better suited to the shared flourishing of the shining, vulnerable, *magnifica humanitas*.