JOURNAL of CHRISTIAN LEGAL THOUGHT

LAW, FAITH, AND ARTIFICIAL INTELLIGENCE

VOLUME 14 2024 NUMBER 2

JOURNAL of CHRISTIAN LEGAL THOUGHT

VOL. 14, NO. 2 | 2024

PUBLISHED BY

Christian Legal Society (CLS), founded in 1961, seeks to glorify God by inspiring, encouraging, and equipping Christian attorneys and law students—both individually and in community—to proclaim, love, and serve Jesus Christ through the study and practice of law, through the provision of legal assistance to the poor and needy, and through the defense of the inalienable rights to life and religious freedom.

COVER EMBLEM

The inside design symbolizes the spirit of a builder in its dislocated features resembling the architecture of layered bricks and the four pillars representing the four ministries of CLS. The branches represent harvest and the ongoing mission of the Church to toil the land, water the seeds, and pray to God to send the increase. The circle represents completion—embodied in the incarnation and second coming of Christ as the proverbial Alpha and Omega.

"For we are co-workers in God's service; you are God's field, God's building." (1 Corinthians 3:9)

STATEMENT OF PURPOSE

The mission of the *Journal of Christian Legal Thought* is to equip and encourage legal professionals to seek and study biblical truth as it relates to law, the practice of law, and legal institutions.

Theological reflection on the law, a lawyer's work, and legal institutions is central to a lawyer's calling; therefore, all Christian lawyers and law students have an obligation to consider the nature and purpose of human law, its sources and development, and its relationship to the revealed will of God—as well as the practical implications of the Christian faith for their daily work. The *Journal* exists to help practicing lawyers, law students, judges, and legal scholars engage in this theological and practical reflection, both as a professional community and as individuals.

The *Journal* seeks, first, to provide practitioners and students a vehicle through which to engage Christian legal scholarship that will enhance this reflection as it relates to their daily work; and, second, to provide legal scholars a medium through which to explore the law in light of Scripture, under the broad influence of the doctrines and creeds of the Christian faith, and on the shoulders of the communion of saints across time.

While the *Journal* will maintain a relatively consistent point of contact with the concerns of practitioners and academics alike, it will also seek to engage outside its respective milieu by soliciting work that advances the conversation between law, religion, and public policy. Given the depth and sophistication of so much of the best Christian legal scholarship today, the *Journal* recognizes that sometimes these two purposes will be at odds.

EDITORIAL POLICY

The *Journal* seeks original scholarly articles addressing the integration of the Christian faith and legal study or practice, broadly understood, including the influence of Christianity on law, the relationship between law and Christianity, and the role of faith in the lawyer's calling. Articles should reflect a Christian perspective and consider Scripture an authoritative source of revealed truth. Protestant, Roman Catholic, and Orthodox perspectives are welcome as within the broad stream of Christianity.

However, articles and essays do not necessarily reflect the views of Christian Legal Society or any of the other sponsoring institutions or individuals.

To submit an article or offer feedback to Christian Legal Society, email us at CLSHQ@clsnet.org.

TABLE OF CONTENTS

BOARD OF ADVISORS

Defining Humanity Down: The Irony of AI and Human Anthropology

*by Jason Thacker**

Introduction

Amid the ongoing conversations surrounding artificial intelligence (AI) today, there seems to be a common thread that permeates so much of the current discourse of where we are heading as a society and the role AI ought to play in our lives. No matter what side of the debate one may fall on with these tools, it seems that most people recognize that things are changing at lightning speed and that our society doesn't seem to be ready for what is taking place. It seems like every day we hear about another innovation in this space, how yet another company is rebranding itself as an AI company, and how AI is already radically shaping our society both for good and ill. From wall-to-wall media coverage to bold predictions of what will take place just in the next few years with dream of human-level AI, it is nearly impossible to keep these discussions at arm's length any longer. Many now wonder how best, if at all, to use these tools in the academy, healthcare, business, industry, government, warfare, and even the church itself with recent controversy surrounding a defrocked Catholic AI "priest."[1] AI is everywhere and, despite much of the cultural hype, its use in society is growing at an exponential rate. But there seems to be little widespread consensus of what it is, where we are headed, and what we ought to use these tools for.

In our age centered on efficiency and convivence exacerbated by our technological innovations, we rarely slow down enough to ask the hard questions and think holistically about the power these tools have over our lives. Wisdom flowing from a distinctly Christian worldview calls us to think deeply about these tools and how they are shaping our pursuit of loving God and loving our neighbor as ourselves. We must ask ourselves: Is technology merely a neutral, valueless tool we simply use for convenience and efficiency's sake, or does it represent something more that is deeply shaping every aspect of our lives for both good and ill? Is it possible that we think we are simply using these tools, but that they are actually using us instead? Could these AI tools (and all technologies for that matter) be shaping our perception of the world and our values including some of the fundamental ideas we hold about what it means to be human?

AI is not a neutral tool, but rather one that is radically altering how we perceive reality, especially the value of humanity even for Christians who rightly understand that humanity is uniquely made in the *imago Dei*. We must first recognize the non-neutrality of technology before understanding the nature of AI and how it affects our beliefs about the value of humanity, which is not found simply in what we *do*, but rather in who we *are*. As Christians engaging these conversations and consequential decisions about the role of AI in society, we must keep human dignity at the center of our ethic and ask the ever-prevalent question of *should* we do something, rather than simply the question of *can* we. In a world that pushes us to go faster and be more efficient in every aspect of life, it is good and wise for us to take time to slow down and ask some of the hard questions about AI and how these machines are shaping our view of the world. We must reconsider how we often define humanity down in an age of advanced machines.

The Non-Neutrality of AI

Most are familiar with the adage "when you have a hammer, everything looks like a nail." This saying rings true whether you are a child with a toy

* Assistant Professor of Philosophy and Ethics, Boyce College; Senior Fellow, The Ethics and Religious Liberty Commission.

1 For more on the AI "priest" from Catholic Answers, *see* Gina Christian, *AI 'Priest' Sparks More Backlash than Belief*, Nat'l Catholic Rep. (Apr. 25, 2024), https://www.ncronline.org/news/ai-priest-sparks-more-backlash-belief.

hammer or a grown adult. When someone picks up a hammer, we all instantly know that it is designed to hit things—whether those things are actual nails or not. All tools have a particular purpose, design, and telos by design. Media theorist Neil Postman extends this truism and notes that "[t]o a person with a pencil, everything looks like a sentence. To a person with a TV camera, everything looks like an image. To a person with a computer, everything looks like data."[2] Theologian and ethicist Jacob Shatzer adds here that "when you've got a smartphone with a camera and the ability to post something online, everything looks like a status update."[3] And we can extend that again to say that to a person with access to powerful AI tools, humanity itself begins to look like a mere machine. Postman argues that those truisms call to our attention the fact that every technology has a prejudice, purpose, or design both with intended and unintended consequences. He explains, "embedded in every tool is an ideological bias, a predisposition to construct the world as one thing rather than another, to value one thing over another, to amplify one sense or skill or attitude more loudly than another."[4] He goes on to state that "[n]ew technologies alter the structure of our interests: the things we think *about*. They alter the character of our symbols: the things we think *with*. And they alter the nature of community: the arena in which thoughts develop."[5] We are often so blinded by the formative power of technology that we fail to clearly see many of these values and prejudices.

As opposed to a simple tool-based, neutral view of technology, these tools are shaping every aspect of our society and culture toward convenience, greater efficiency, and speed—at any cost.[6] As computer scientist Derek Schuurman highlights, technology is thus *value-laden* and never neutral.[7] No technologies exist simply as neutral, isolated tools, but rather as tools that represent the values of our larger culture and in particular an often truncated and bastardized view of the human person based on what we *do* rather than who we *are*. This push of efficiency often leads to an incomplete view of the human person, toward an instrumentalizing of humanity. These tools are deeply altering and shaping every aspect of our lives including our view of God, ourselves as human beings, and the world around us. This is especially true in how these tools fool us into thinking we are more powerful than we really are and how they shape our view of our neighbors who are made in the very image of God.

Defining AI

One of the ways we can see how AI is shaping us is through the language we use to even define it and our visions of where we are heading as a society given the ever-expanding access to and development of AI. From referring to our AI assistants as she/her or he/him to our dreams of conscious, human-like machines, we tend to anthropomorphize these machines in ways that are deeply concerning for human anthropology. Artificial intelligence can be defined as non-biological intelligence, where a machine can perform various tasks that were once reserved for human beings. AI represents an aspect of the broader field of computer science, which comprises several foci such as machine learning, deep learning, natural language processing, robotics, machine vision, speech recognition, and much more. AI has in recent years become a major topic of discussion across industries given how it can automate, streamline, and augment various aspects of our lives—and is increasingly being explored for use in medical applications, where the challenge

2 Neil Postman, Five Things We Need to Know About Technolgoical Change (Mar. 28, 1998), https://web.cs.ucdavis.edu/~rogaway/classes/188/materials/postman.pdf.

3 Jacob Shatzer, Transhumanism and the Image of God: Today's Technology and the Future of Christian Discipleship 7 (2019).

4 Neil Postman, Technopoly: The Surrender of Culture to Technology 13 (1993).

5 *Id.* at 20 (emphasis original).

6 For more on a Christian philosophy of technology, *see* Jason Thacker, *Simply a Tool? Toward a Christian Philosophy of Technology, in* The Digital Public Square: Christian Ethics in a Technological Society (Jason Thacker ed., 2023).

7 Derek C. Schuurman, Shaping a Digital World: Faith, Culture and Computer Technology 22 (2013).

of an instrumentalizing view of the human person is all too common.[8]

Most of the public discussions about AI tends to focus on the future of AI and what might be possible in the coming years, including debates over the possibility of human level or superhuman AI systems.[9] But the only form of AI that has ever been created—and that many believe is possible—is called narrow AI. These narrow AI tools have specific use cases and applications. These tools are already revolutionizing every aspect of our society and growing more advanced each day, often outperforming humanity in narrowly focused tasks such as a recommendation algorithm on social media, online shopping, or entertainment. They also often control various aspects of our home and work life through automation like smart devices, communications, and even banking. While the use of these tools is becoming quite ubiquitous as they often operate behind the scenes in our personalized digital experiences, these tools are mere objects that do not *know* or *understand* what they are doing despite us giving them names, faces, and wondering if they will become like us one day. As Catholic philosopher Robert Spaemann notes, "Even today [computers] are in many respects ahead of the intellectual feats that humans perform. Yet it is not pointless to say, 'the computer does not think.' It means that it does not know it is thinking. And it means that it does not experience thinking. There is no ghost in the machine."[10]

The next type of AI is highly debated and likely not even possible given the complexities and unique nature of humanity as more than a simple material being. Broad or general AI is often described as human level intelligence, where a machine not only is able to meet or surpass humans in narrow ways but in a much broader sense. Many have long debated in both computer science and philosophy if achieving human level AI is even possible given that humanity is not simply a material being, but also a spiritual one. Recently, tech leaders like Elon Musk joined the fray stating that within the next year or two, we will have AI that is smarter than humans or what is often deemed artificial general intelligence (AGI).[11] AI companies like OpenAI already have public plans and states goals for AGI systems despite widespread disagreement if these tools are even possible.[12] Some predict that humanity will even be able to achieve a superhuman type of intelligence, also known as artificial super intelligence (ASI) or a God-like intelligence.[13] This type of AI not only outperforms humanity in all aspects of life and gains consciousness, but also takes on a transcendent role in human affairs. Often these debates over the future of AI and what is possible are rooted in a naturalist/materialistic philosophy that is completely at odds with a Christian vision of reality, truth, humanity, and the good life that recognizes that humans are not mere material beings or the sum of our parts.

While we have always had various forms of technology, today's advancements in narrow AI can seem quite daunting given how complex and powerful they are becoming. These systems are performing a wide array of tasks that were once solely reserved for humans and pose an entirely different set of ethical questions for us to consider. But at the core of these seemingly novel questions is the reality that technology isn't really causing us to ask new questions of life per se, but rather to ask perennial questions that we have long asked as humanity in light of new opportunities. These tools are challenging long-held assumptions of human anthropology and are expanding our moral horizons. So, if these tools and some of the questions we are asking now aren't all that new per se, why the alarm over AI and why now?

---

8 For more on the instrumentalizing of humanity in medicine, *see* Jeffrey P. Bishop, The Anticipatory Corpse: Medicine, Power, and the Care of the Dying (2011).

9 I write more about these types of AI in chapter 8 of my book, The Age of AI: Artificial Intelligence and the Future of Humanity (2020).

10 Robert Spaemann, Persons: The Difference Between 'Someone' and 'Something' 42 (Oliver O'Donovan trans., 2017).

11 Elon Musk (@elonmusk), X (Mar. 12, 2024, 10:25 PM), https://x.com/elonmusk/status/1767738797276451090.

12 For more on plans for AGI from companies like Open AI, *see Planning for AGI and Beyond*, OpenAI (Feb. 24, 2023), https://openai.com/index/planning-for-agi-and-beyond/.

13 Thacker, *supra* note 9, at 174-75.

The Dehumanizing Irony of AI

One of most ironic things about this age of AI is that we tend to humanize our machines and dehumanize ourselves in the process. Humanity tends to develop and use these tools in ways that cause us to ask questions about what these tools might become, seeking to humanize our machines through anthropomorphic language, and even treat these machines as if they are our companions or worse: soon-to-be gods with dreams (nightmares) of artificial super intelligence.

A recent example of this phenomenon can be seen in the announcement of the Friend AI-empowered pendant worn around your neck that promises, not to help you be more productive per se, but to keep you company as a close companion and friend. This tool is always listening and communicates with the wearers through text messages and push notifications to a smartphone.[14] Yet while we humanize these machines with names, faces, and even misconceived ideas of real companionship, we also dehumanize ourselves, seeing each other as merely the sum of our parts and capacities in a materialistic framework devoid of human uniqueness and exceptionalism. It seems that one of the main reasons many in society find themselves both amazed and fearful about these tools is that AI is fundamentally challenging what we have long held of what it meant to be human. These advanced AI systems have fundamentally challenged much of what we have assumed about the uniqueness of humanity because for generations humanity has often assumed that what it meant to be human was simply a capacity or attribute including the ability to think, create, use language, make weighty decisions, and perform certain complex tasks. But AI systems are performing many of those tasks that in the past were solely reserved for humans, thus forcing us to question some of our anthropological assumptions. In truth, these tools are *imitating* and *mimicking* human behaviors that we have long assumed were only possible for other human beings to emulate.

Christians often rightly employ the language of the image of God, *imago Dei*, when speaking about the unique nature of humanity and for good reason. As Genesis 1:26-27 states,

> Then God said, "Let us make man in our image, after our likeness. And let them have dominion over the fish of the sea and over the birds of the heavens and over the livestock and over all the earth and over every creeping thing that creeps on the earth." So God created man in his own image, in the image of God he created him; male and female he created them.

Definitions of the image of God have long been debated within Christianity, with each generation often defining the image based on the contemporary challenges they face in society. A robust biblical anthropology is central to all of life, especially in Christian ethics, and is vital for navigating the complex opportunities and challenges before us in this age of AI. Better understanding what it means to be human can function as a robust moral apologetic as we seek to give a defense for the hope within us and to engage others with both gentleness and respect amid today's challenging technological questions.

Traditionally, the church has recognized three prevailing views of the image that emphasize a particular capacity or attributes as the defining factor of humanity. The first view is known as the substantive view of the image, which focuses on a capacity or attribute related to reason, rationality, creativity, or even the use of language. This has been a dominant view throughout church history, and a structural understanding of human value is widely held in philosophical anthropologies as well. In a Christian perspective, we see that God created us with the unique ability to reason or have the capacity for rational thought that is different and more advanced than other aspects of creation. Humanity does indeed often manifest a higher level of rationality, as opposed to other forms of life in creation. Though, theologian Ryan S. Peterson notes that the image of God "should not be interpreted by comparing humanity to other creatures, identifying

14 For more on the Friend AI pendant, *see* Boone Ashworth, *Wear This AI Friend Around Your Neck*, Wired (June 30, 2024), https://www.wired.com/story/friend-ai-pendant/.

the differences, and positing the differences as the definition of the image."[15] There are a number of challenges to this traditional view of the human person, but the most consequential one seems to be that if reason is seen as the defining factor of the image and of human dignity, then supporters of this view must address questions of the dignity of human beings who may not exhibit such common rational capacities or are incapacitated for various reasons. To counter these challenges, some philosophers have developed concepts such as the "privilege of the normal" or even identified reason as a "range property," which was famously articulated by political philosopher John Rawls in his *A Theory of Justice*.[16]

A second view tends to focus on the capacity or attribute of social interaction and relationships with God and others. The relational view tends to focus on our ability to form and maintain relationships with God and our fellow image bearers. Again, this view, like the one above, rightly sees this capacity and attribute as being uniquely manifested in humanity, but then one must again ask if this is the defining factor of the image and human dignity. This approach to human value and equality may fall prey to similar critique as the earlier substantive interpretation of the *imago Dei* because it tends to reduce the image down to a mere capacity for relationships—both with God and others—which are obvious implications of the text and the whole canon itself. But this view seems to fail to account for human dignity for those human beings who do not exhibit this attribute or do so at lower levels than traditionally associated with being human.

Lastly, the third view is centered on how humanity functions as representatives of God through our roles and responsibility. This vice-regency or representative view of the image sees the image primarily manifested in the capacity to perform certain functions or jobs and to represent God as His image bearers in this world. The royal interpretation is the idea that because we are created according to God's image, humanity functions as His representatives or vice-regents on earth by exercising dominion and stewardship over all of creation on behalf of God. This view rightly emphasizes humanity's shared moral responsibility and moral agency as God's image bearers. As Carl F.H. Henry correctly notes, this emphasis on taking dominion as a representative of God made in His image and likeness is "clearly an aspect of the Genesis teaching."[17] While being a main thrust of Genesis 1, Peterson notes this view tends to equate human identity and function, which are not "necessarily identical even if they are mutually dependent."[18] Similar to the critiques of the previous two views, this view tends to equate one's dignity with what one *does* rather than who one *is* by nature of being biologically human.

Not all human beings have high levels of intellect, emotional and relational IQ, or function in particular ways that we often see manifest in and associate with being human. This can be due to cognitive or physical disabilities, age, or even stage of development in the case of our preborn neighbors and the most vulnerable in our society, including young children. How we define what it means to be human has vast implications for all of life, including how we think about emerging technologies that are beginning to mimic or imitate attributes that we have long thought were exclusively human. In short, it is far too easy today to assume that our value and dignity—and that of our neighbor—is simply based on what we *do* or on what we can *contribute* to our society. But the Christian ethic reminds us—especially in an age of emerging technologies like AI—that the value and dignity of humans isn't rooted in what we *do*, but rather in who we *are* as unique image bearers of our creator. God made us in His very image and nothing—not even the most advanced AI systems—will be able to change that unique status given to us by our Creator.

While all three of the previously mentioned views are clearly implications of the image and are manifested properly in most human beings in varying degrees, there is a fourth and all-encompassing view of the image that isn't direct-

15 Ryan S. Peterson, The Imago Dei as Human Identity: A Theological Interpretation 33 (2016).

16 John Rawls, A Theory of Justice 444-46 (Belknap Press rev. ed. 1999).

17 Carl F. H. Henry, God, Revelation, and Authority 139-40 (1976).

18 Peterson, *supra* note 15, at 41.

ly tied to a particular attribute or capacity, but rather is seen as an ontological status as a biological human being made in God's image. This status is inalterable, unchanging, and something bestowed upon us uniquely by our Creator. As theologian Richard Lints rightfully points out, the *imago Dei* "does not appear as a place marker for an otherwise long list of human traits and qualities," meaning that while the image of God obviously has several implications and external manifestations in the lives of humanity, it isn't reduceable down to our traits, attributes, or qualities alone.[19] Spaemann notes something similar by stating that "human beings have certain definite properties that license us to call them 'persons'; but it is not the properties we call persons, but the human being who possess the properties."[20] He later notes that "there are, in fact, no potential persons. Persons possess capacities i.e. potentialities, and so personhood may develop. But nothing develops into a person."[21] A person in Spaemann's framework is *someone* (a subject) rather than *something* (an object), meaning regardless of one's capacities or attributes they are persons by simply being a member of the human species. He writes that "there can, and must, be one criterion for personality, and one only; that is biological membership of the human race."[22] Human beings are a specific kind of creature who do indeed exhibit certain characteristics and attributes in unique ways, but our dignity is based upon the mere presence of those attributes or capacities.

Critics of this view may be dissatisfied with speaking of the value of humanity as a kind of ontological status instead of a specific attribute or capacity to be identified, but, as C.S Lewis reminds us, "you cannot go on 'seeing through' things forever. The whole point of seeing through something is to see something through it."[23] Questions of the ethical development and use of AI should be centralized on an understanding of human dignity as an unchangeable status rather than simply what one does, especially with powerful machines that can now imitate or mimic particular human attributes in varying degrees. Thus, one of the ways we love God is by loving our neighbors as ourselves,[24] recognizing their dignity and value as image bearers of the Almighty God is not tied to what they *do* but who they *are*. These realties will refocus our approach to AI development and use—centering them on human dignity.

Moving Forward

There is a massive push to adopt new technologies like AI, often without adequate reflection on how these tools inevitably shape our view of the human person and the world around us. While advanced AI tools may mimic or imitate certain human characteristics, they are mere objects and machines, never subjects like you or me. It may sound trite given the challenges we face with AI today in society, but we must ask ourselves: Are these tools helping us to love God and love our neighbor as ourselves or are we sacrificing those things in the pursuit of increasing the bottom line or building ourselves up at the expense of another's dignity? We must consider how these tools are affecting other people who are made in the very image of God, not just what is profitable or productive. Christians developing and interacting with these tools need to ask the question of *should we* vs. *can we*, which is at the very core of wisdom and the path forward for thoughtful Christians in this age of AI. It is far too easy for us to become enamored with these tools—to give into worldly hype and simply assume we should use something just because we can. Instead of just adopting new technologies because everyone is talking about them or trying to sell them to us, we must begin thinking about their potential use, how they are shaping our perspective of the world, and the possible risks and dangers associated with these tools—especially as it relates to the dignity of our fellow human beings.

As Postman points out, we must enter with our eyes wide open and recognize that new tech-

19 Richard Lints, Identity and Idolatry: The Image of God and Its Inversion 60 (2015).

20 Spaemann, *supra* note 10, at 236.

21 *Id.* at 245.

22 *Id.* at 247.

23 C. S. Lewis, The Abolition of Man 81 (Harper Collins 2001) (1943).

24 *Matthew* 22:37-39.

nologies are not additive but ecological as they change everything about an environment when used. He employs the illustration of adding a drop of red dye to a bowl of clear water, noting that you don't end up with a bowl of clear water plus a drop of red food coloring. You end up with bowl of pink water as everything is changed by the inclusion of the drop.[25] Similarly, technology, especially AI, radically alters everything in our lives, including a view of God, ourselves as human beings, and the world around us—whether we realize it or not. One of the greatest temptations when faced with complex or challenging ethical questions with technology is the rush to a position of uncritical adoption or rejection of these tools. Wisdom, which is at the core of the Christian moral tradition, calls us to slow down and to think deeply about the nature of these tools, as well as the myriad of its uses.

As the influence of technology continues to increase throughout our society, Christians need to be reminded that we have a robust ethic centered on loving God and neighbor in the daily engagement of current issues from a place of hope and faith rather than debilitating pessimism or even unbridled optimism. AI isn't going anywhere, and Christians need to think deeply about how these tools are shaping our perspective of God, ourselves, and the world around us. As AI continues to fundamentally challenge what we have long assumed it means to be human, that does not alter how God made us in His image, as well as His deep and abiding love for us. As the church has historically done amid challenges, we must articulate even more precisely what we believe and "contend for the faith once for all delivered to the saints"[26]—no matter what challenges AI may bring. We have a steadfast hope, even in the midst of an uncertain future, because we know that God is, above all, sovereign over history and all of humanity, and that nothing will ever supplant how God made us in his very image, not even the most advanced AI system.

---

25 Postman, *supra* note 2.

26 *Jude* 3.

The Multifaceted Impact of Generative AI on Lawyers and Legal Services

*by Jordan Furlong**

Introduction

The legal services sector is about to enter a transformative era, driven towards an uncertain future by several sociological, environmental, and technological forces—most dramatically, by stunning recent advancements in Generative AI (Gen AI). This extraordinary technology has the potential to revolutionize legal services, reshaping not only the efficiency with which legal tasks are performed, but also the very nature of lawyers, law firms, and legal services provision.

Historically, law firms have thrived on a business model that values billable hours and the continuous labor of junior lawyers. But Gen AI enables lawyers to complete tasks with unprecedented speed and growing accuracy, not just boosting productivity, but also liberating lawyers to engage more deeply in creative and strategic thinking, thereby enhancing their service to clients. The implications of this shift extend beyond operational efficiency, however. As law firms gradually integrate Gen AI, they will be compelled to re-evaluate their business models, evolving from a time-based profit and compensation system to one that prioritizes outcomes and client experiences. This will trigger a fundamental rethinking of how law firms operate—from client relationships to internal collaboration to professional development—and what purpose lawyers ultimately serve.

Furthermore, the adoption of Gen AI has profound implications for the accessibility of justice. By potentially lowering the costs and increasing the availability of legal services, Gen AI could democratize access to legal assistance, addressing longstanding inequities in the legal system. For Christians in particular, these developments carry weighty spiritual and ethical significance. The integration of AI into the legal sector can be seen as an opportunity to fulfill the Christian mandate of promoting justice and human dignity, ensuring that technological advancements serve to enhance, rather than undermine, the cause of righteousness and good works in society.

This article will explore the multifaceted impact of Gen AI on the legal profession, including its transformative effects on law firms, its redefining impact on lawyers' purpose, its acceleration of lawyer development, and its potential for broader access to justice. Finally, it will consider the Christian implications of all these changes, reflecting on how AI might be harnessed to promote a more just and humane legal system, as well as a more meaningful and effective legal profession.

The Transformation of Law Firms

Gen AI enables lawyers to carry out numerous legal tasks much faster than has been possible in the past. It can best be understood as an incredibly inventive legal intern that supercharges a lawyer's "productivity"—not as law firms define that term, as simply the number of hours billed, but as the rest of the world does, as the ability to generate more outcomes more quickly with less effort and fewer resources. But Gen AI can also enhance a lawyer's creativity, helping to produce new ideas and perspectives when employed as a suggestion generator, sounding board, and role-playing assistant. As a result, Gen AI can also help unburden the lawyer of grinding, repetitive, low-value tasks and free the lawyer to think more deeply and to greater strategic effect.

In other words, Gen AI creates both an "efficiency play" and an "effectiveness play" for law firms. Each of these impacts of Gen AI, howev-

* Jordan Furlong is a legal market analyst, consultant, and author who helps law firms and legal organizations adapt to the rapidly changing legal services environment. An accomplished speaker, he has addressed thousands of lawyers and legal professionals at dozens of events throughout North America, Europe, and Australia over the past 15 years. Jordan is based in Ottawa, Canada, where he and his family attend Our Lady of Good Counsel Catholic Church.

er, has serious implications for the traditional law firm business model. On the efficiency side, most businesses would welcome a technology that reduces the time and cost required to produce output for customers because that would lower their costs and enhance their profits. But law firms aren't most businesses. They make money by having lawyers perform legal tasks thoroughly and painstakingly and billing the time and effort involved. It is particularly important to note that law firms' profits hinge on the ability to "leverage" the work of junior lawyers carrying out low-to-mid-level tasks, billing that work for a higher amount than the junior lawyer is paid, and pocketing the difference. Law firms' profit margins reside in the billed hours of their lawyer employees. A technology that turns those hours into minutes or seconds is fundamentally incompatible with that profit model.

But Gen AI's ability to supercharge creativity also presents a problem: law firms don't make money by being creative. Their *lawyers* often solve problems or create opportunities by being creative, but this happens relatively rarely. Much legal work is rote and does not require a great deal of ingenuity or imagination. Senior lawyers use their creativity more often, but their minute of creative insight is worth 1/60th of their hour of travelling to a deposition. By integrating into their business a new technology that improves both efficiency and creativity, therefore, law firms are playing against their own strengths. Law firms sell human effort; Gen AI vastly reduces it. Law firms neither reward nor make much use of human creativity; Gen AI kicks it into a higher gear.

The more use law firms make of Gen AI, the less like "law firms" they will become. They will turn themselves into *productivity-based businesses*—highly effective enterprises that deliver value through lawyer creativity and that profit from systemic efficiency. Gen AI, if it continues to evolve in its current direction, thus will render obsolete the organizing principles and business rationales of traditional law firms and oblige lawyers to come up with new value propositions to the market and to the workforce better suited to post-AI realities.

This transition will not be painless. There is no instruction manual for developing an AI-powered law firm that generates profits from the creative, efficient productivity of its highly skilled workers. But whenever and however that law firm evolves, these will be its salient features:

- Relationships with clients will be long-term and continuous, not episodic and irregular.
- Relationships with workers will be empowering and cultivating, not extractive and exploitative.
- Internal collaboration will be essential to client-focused productivity, not a noisy distraction from hours-focused "productivity."
- Fees will be based on client relationships and strongly influenced by shared risks and jointly sought-after outcomes.
- Customized training and professional development of workers will start on day one and continue until the worker leaves the firm.

Most of all, a law firm that allows itself to be transformed by Gen AI will become more deeply attuned to the professional and ethical fulfillment of its clients' goals. When lawyers can no longer sell the time it takes to achieve a client's outcome, then they must sell something new: the outcome itself, and the client's experience of reaching it.

The Reinvention of Lawyer Activity

What is the least that Gen AI will be able to do in the law? The jury has only begun to deliberate on this question. Few observers are willing to wager that Gen AI's application to legal matters has already peaked, or that it ultimately will prove too error-prone and insufficiently specialized to have any impact on how legal work is done. But fewer still are betting on a full-scale replacement by AI of the entire legal profession. The likeliest outcome is that Gen AI's effect on the legal sector will land somewhere on the spectrum between these two extremes of changing nothing and changing everything. But not all landing points on that spectrum are weighted equally. As we move towards the maximum-impact end, the risks to the legal sector increase exponentially.

Suppose that Gen AI does trigger shattering changes in the legal landscape. If so, then legal institutions rooted in that landscape (law firms, law schools, courts) will face serious if not exis-

tential challenges to their structure and viability. Even if this outcome is very unlikely, the payload of its risk is still high because the impact would be so severe. We must therefore seriously consider a scenario in which Gen AI proves capable of performing a breathtaking range of legal activities—writing legal documents, conducting legal research, handling contract negotiations, monitoring regulatory compliance, rendering legal opinions, overseeing complex litigation. Every legal entity mentioned above should be devoting time and energy to war-gaming that possibility.

But even if Gen AI achieves those stratospheric heights, there is a case to be made that lawyers will still be needed to fulfill three essential roles that will remain beyond the capacity of AI and be invaluable to clients. Lawyers will *advocate, advise,* and *accompany.*

1. *Advocate*

A lawyer will personally represent a client and truthfully advance or protect their interests before a legal arbiter or in the public square, using knowledge of the law and human society and the skills of reasoning, persuasion, and rhetoric. The lawyer will advocate on the client's behalf, "standing in" for the client without personally joining to the client's position.

Only humans can advocate. The very act of advocacy, of rising to say, "I speak for this person," is a personal commitment that signals to the audience and the community that the person and their position deserve to be heard. A computer could be programmed to argue a position, but only a human can make that meaningful commitment.

2. *Advise*

A lawyer will provide a client with informed guidance regarding difficult or important choices. The lawyer will gather relevant information and recommend one or more courses of action that the lawyer believes will advance or protect the client's interests. Using the human qualities of judgment, wisdom, and good character, the lawyer will identify options and suggest which one(s) seem best.

To advise is to guide with assurance, counsel with confidence, offer direction with perspective. For years, clients facing hard choices have told their lawyers, "I know what the law says; but what do you think I should do?" In the future, "the AI" will substitute for "the law," but the question will be the same. People will still need good advice from trusted advisors.

3. *Accompany*

A lawyer will join a client on their journey, providing steady support and companionship, for as long as the client wants and for whatever the client might need. To accompany someone is to make an ongoing personal commitment of presence, interest, and reliability—to *affirm* the person's worth and their journey's importance through the ongoing act of personal engagement.

Advice and advocacy are centuries old in the legal profession; yet for lawyers, accompaniment might be their essential service: to maintain an ongoing relationship of supportive presence, offering insights or conversation or even just quiet presence while the client journeys towards their life or business objectives. In some ways, it is the most human thing a lawyer can do.

Gen AI, in the most advanced development scenario, poses an enormous threat to lawyers' traditional livelihood. But perhaps that livelihood represents a profoundly shortsighted view of lawyers' real capacity to provide transformative help for their clients. If Gen AI does become all that it might be, then perhaps lawyers will be forced to become all *they* might be.

The Acceleration of Lawyer Development

The role of the law firm associate has never looked more tenuous. As noted above, Gen AI is likely to become more widely used in law firms in the years ahead. As law firms learn to leverage high-proficiency technology rather than low-proficiency associates, they will naturally hire fewer such associates to carry out basic work that the technology can perform.

At the same time, many law firms are prioritizing the lateral hiring of senior partners with strong client relationships over the recruitment and development of new lawyers.[1] Firms are

1 Roy Strom, *Davis Polk Leaps Into Lateral Hiring Seen as Re-Shaping Industry*, Bloomberg L. (May 6, 2024), https://news.bloomberglaw.com/business-and-practice/davis-polk-leaps-into-lateral-hiring-seen-as-re-shaping-industry.

showing a preference for growing their external business development capacity over expanding their ranks of low-skilled lawyers. It now takes twice as long for homegrown associates to become partners in their law firms as it does lateral hires.[2]

The combination of these two trends means that most law firms will have fewer associates in the future than they do today. This has enormous implications for the life-cycle of lawyers and the numerous participants in the lawyer development industry. Law schools will likely fare the worst from this turn of events. Law firms are the primary "buyer" of law schools' inventory (new law graduates). If firms hire fewer new lawyers, that would lead to a significant decrease in law schools' post-graduation employment levels, making the law degree a less desirable commodity. Moreover, demand is likely to remain strong for graduates from the most highly regarded schools, meaning that the remainder will feel the impact of lower hiring rates disproportionately. A sustained drop in new lawyer hiring would create significant hardship for many law schools; a permanent drop to a new normal hiring rate could force a number of them to close. That could perhaps lead to a long-overdue decrease in the cost of a legal education. But an even better outcome would be a re-evaluation of the proper role of legal education and a reconfiguration of how the profession prepares its future members in their formative years.

It is becoming more widely accepted that neither a law degree nor bar examinations make new lawyers competent to practice law.[3] Five states have recently adopted or are currently considering new pathways to licensure that rely on supervised practice in a law firm rather than passage of the bar exam as the key post-graduate element of bar admission; more could follow.[4] A greater emphasis on practice readiness at the point of licensure is likely to result.

Similar shifts are likely to occur in law firms. The advance of Gen AI not only will prompt firms to hire fewer new lawyers, but also will oblige firms to change their approach to the lawyers they do hire. Increasingly, new lawyers will enter law firms to find that technology has taken over most of the work they are qualified to perform.[5] But they will also find fewer competitors for partner positions and a more favorable partner-associate ratio for training and mentoring. Gen AI-era law firms, perhaps paradoxically, will be better positioned to develop their young lawyers than they have been in decades. Entry-level lawyers have only ever been expected to perform and bill entry-level work, but Gen AI is taking that work away. Inexperienced lawyers will not be able to contribute value to the firm. Law firms, therefore, will start to make them more experienced lawyers earlier in their careers.[6]

The opportunity ahead for law firms is to start turning "associates" into "partners" from day one. They could enroll new lawyers in intensive two- to three-year programs to accelerate their development into polished, proficient, and confident legal professionals who can provide real value to clients and even start bringing in business three to four years into their careers. Once law firms begin to leverage technology and systems, they will be motivated to develop their associates into high-value partners as fast as they can. In this way, Gen AI will actually bring about an unprecedented acceleration in the pace of lawyer development.

---

2 Anna Sanders, *As BigLaw Leaders Age, Path To Partnership Lengthens*, Modern Law. (July 27, 2023), https://www.law360.com/pulse/modern-lawyer/articles/1701205.

3 Christine Charnosky, *Legal Experts Weigh In on ABA's Support of Alternative Pathways to the Bar*, ALM (May 22, 2024), https://www.law.com/2024/05/22/legal-experts-weigh-in-on-abas-support-of-alternative-pathways-to-the-bar/.

4 Marin McCall, *Oregon Becomes Third State to Approve Alternative to the Bar Exam*, 2Civility (Nov. 20, 2023), https://www.2civility.org/oregon-becomes-third-state-to-approve-bar-exam-alternative; Rachel Riley, *Wash. Supreme Court Clears Way For Bar Exam Alternatives*, Law360 (Mar. 15, 2024), https://www.law360.com/articles/1814404; Julianna Hill, *Nevada Will Consider Three-Stage Process to Join Bar*, ABA J. (May 29, 2024), https://www.abajournal.com/web/article/nevada-to-consider-three-stage-process-to-join-bar.

5 Andrew Maloney, *AI Will Ramp Up Intense Big Law Battle for Talent, Client Share*, ALM (Apr. 30, 2024), https://www.law.com/americanlawyer/2024/04/30/ai-will-ramp-up-intense-big-law-battle-for-talent-client-share/.

6 D. Casery Flaherty, *Education in the Age of Gen AI: Experiential Training Is Essential for Success*, ALM (Mar. 29, 2024), https://www.law.com/legaltechnews/2024/03/29/education-in-the-age-of-gen-ai-experiential-training-is-essential-for-success/.

The Possibilities for Justice

Regulators are frequently saddled with two opposing mandates. On one hand, they must ensure regulated products and services meet consistent standards of quality and effectiveness. So, regulators impose demands on providers, including competence and conduct norms on service professionals, thereby raising the cost of market participation and driving some providers out of business.

On the other hand, however, regulators must also ensure that the public can actually *obtain* the products or services they regulate. It does no good to apply such stringent regulatory standards that only a wealthy few can find or afford what they need. The more something is regulated, the scarcer and more expensive the product or service becomes, and vice versa.

Every regulator must grapple with these two sliding scales: push one up, and the other gets pulled down. Trying to balance the opposing mandates of quality and availability often leaves regulators having to choose between accessibility and dependability:

- **Accessibility** is a combination of *affordability* (the price of the service is within the financial means of the average consumer) and *convenience* (the consumer can obtain the service with a reasonable degree of ease and simplicity).
- **Dependability** is a combination of *proficiency* (the service and/or its provider meets industry standards for competence and accuracy) and *trustworthiness* (the service and/or its provider can be regarded as safe, reliable, and effective).

In practice, regulators must compromise: they sacrifice some accessibility to ensure baseline dependability and trade some dependability to ensure baseline accessibility. In the legal market, however, regulators have chosen to go all-in on dependability. They have restricted the supply of legal service providers to those whom they consider the only unimpeachably reliable option (members of the legal profession), while taking few if any steps to ensure those providers are even moderately accessible. The public's options for legal assistance, accordingly, are (a) a lawyer, if one can be found and hired, or (b) nothing.

In theory, regulators could move to increase the number of lawyers and authorize new categories of para-professionals to assist people with simpler matters. But even if there were not numerous political barriers blocking such efforts, this would not resolve the accessibility problem. Lawyers and paraprofessionals are one-to-one solution providers: they help one client with one issue at one time. The access crisis is not linear like that. Unmet and unrecognized legal needs vastly outnumber the people who can meet those needs. Millions of unaddressed and unresolved law-related life and business issues need a one-to-many solution—something that somehow manages to be both accessible *and* dependable.

Into this previous intractable problem steps Gen AI, the first legitimate candidate for potentially resolving the accessibility/reliability dilemma in legal services. Recognizing that there are myriad questions today surrounding its accuracy, attainability, and viability, Gen AI's potential here is undeniable. Already, it is affordable and convenient, more so than almost any comparably powerful technology. ChatGPT-4, Claude, Gemini, and other frontier Gen AI models are available to the public either free or at low cost. For further convenience, they are also embedded in public search engines and in widely used desktop software. Generative AI is not, of course, acceptably dependable as a legal resource yet. But dozens of leading law firms are already deploying Gen AI programs to their employees for legal and administrative work, while high-quality, law-specific Gen AI is entering into the market from Thomson Reuters, Lexis-Nexis, and other providers.

There is, today, a realistic pathway towards an outcome where Gen AI provides the world with a scalable, accessible, and dependable legal information and solutions option that's never existed before. It is still a long shot. But the possibility should excite and energize those who wish to see better justice more widely available to more people.

The Christian Implications of Legal AI

Our laws and legal systems are fundamentally secular entities, grounded in human actions and institutions. Yet the law retains a divine lineage. Many religions and cultures trace the first appearance of "The Law" in the world back to a moment when a Supreme Being provided it

directly to His people, whether on stone tablets or by some other means. People of faith therefore are called to view secular law not merely as a means of establishing social order, but also as an echoing expression of divine intention still reverberating throughout the modern world.

Christians' approach to the law in particular should be shaped by Jesus's declaration in Matthew 5:17 that he did not come "to abolish [the Law and the Prophets], but to fulfill them." In this way, Jesus affirmed the righteousness of the Mosaic law; but more importantly, He also identified himself as the means by which the purpose of the Law and Prophets—God's plan for our salvation—would be accomplished.[7] By fulfilling the Law through His life, death, and resurrection, Jesus also imbued the law with divine grace and purpose.

Christians therefore ought to view the law through these two complementary lenses—both as a set of rules that can help bring about a better society and as an architecture of justice and righteousness that can help bring closer the Kingdom of Heaven. St. Paul wrote about the law in both the secular and sacred senses, telling the Romans in 13:1 to be "subject to the governing authorities," but also advising the Galatians in 6:2 that Christians are to "bear one another's burdens, and so fulfill the law of Christ."

Applying that reasoning here, Christians should regard the rapid advance of artificial intelligence in the legal sector as an opportunity to improve the justice system, not only so that people can obtain legal remedies to their problems, but also that we might promote human dignity and advance human redemption in a just and upright society.

Here are three of the myriad considerations and implications for Christians in the deployment of Gen AI in the legal field.

1. *Justice Accessibility*

It is widely understood that our legal system fails to make effective legal assistance accessible to most people.[8] If you are wealthy, or utterly poverty-stricken, lawyers will be made available to you. If you are a corporation or a government, you can afford to hire both your own lawyers and outside counsel. If you do not fall into any of these categories, however—and the great majority of people do not—then you can scarcely afford to retain the services of a lawyer. For those lucky few who can, it is likely that the costs of pursuing a legal matter for any length of time will quickly become unsustainable.

There are myriad reasons why legal remedies are unavailable to most people. The quantum and unpredictability of lawyers' fees is a familiar target. Less widely noted is the legal system's decaying infrastructure: inadequate legal system funding, ponderous court systems, erratic government benefits, and so on. The legal disempowerment of everyday people deprived of knowledge about their legal rights and remedies is another reason. One overlooked but important factor, however, is the legal profession's aversion to scaling its services through technology, so as to enable one-to-many legal solutions that are both accessible and dependable. There are no "automated lawyer machines" as there are "automated banking machines" that allow consumers to carry out basic transactions through technology. Or at least, there are no such machines today. In the future, thanks to Gen AI, that could change.

What are the Christian lawyer's considerations and obligations here? The church has long held that justice is a universal expectation and human flourishing requires the ability to avail oneself of all the protections and remedies that the law provides. Christians therefore should support the development of any means, including Gen AI, by which the number of people who can access trustworthy and reliable information, guidance, and assistance might grow. But there is an important caveat here. People who cannot afford to hire a lawyer have the same rights to trustworthy and reliable legal assistance as people who can. Socio-economic disadvantage should not invalidate that truth.

It would be wrong to support a Gen AI legal solution that is developed through "experimen-

7 David VanDrunen, *Jesus Came "Not to Abolish the Law but to Fulfill It": The Sermon on the Mount and Its Implications for Contemporary Law*, 47 Pepp. L. Rev. 523 (2020).

8 Rebecca L. Sandefur & Matthew Burnett, *Justice Futures: Access to Justice and the Future of Justice Work*, *in* Rethinking the Lawyer's Monopoly: Access to Justice and the Future of Legal Services (David Engstrom & Nora Freeman eds., 2024).

tation" on disadvantaged or marginalized people, on the grounds that they should be grateful for any assistance, even if flawed and incomplete.[9] People are not to be used as means to an end; they are ends in themselves. Christians therefore should ensure that Gen AI legal systems are as suitable for poor clients as they would be for rich ones and that people who help to train them give their fully informed consent to (and are duly compensated for) participating in that process. And should those systems ultimately develop to a point where one-to-many justice solutions are indeed widely available, Christians should bend their efforts to ensuring that these solutions remain available to everyone, not just to a select few.

2. *Murky Origins*

The promise of Gen AI for expanding the horizons of justice is real. But Gen AI is developed by humans and trained on human activity, and with human origins come complications that inevitably touch on moral and ethical concerns that Christians should be prepared to address.

Gen AI programs were assembled and trained using staggering amount of data, and it is generally acknowledged even by the companies that created them that this data was farmed without the consent of the people who first created it. As of May 2024, there were 24 different copyright lawsuits pending against AI companies on the basis that their products were trained on data illegally scraped from the internet,[10] the most famous of which pits *The New York Times* against Open AI and Microsoft.[11] The merits of all these claims and counterclaims are outside the scope of this article and, to some extent, ancillary to the subject.

Legal distinctions aside, Christians cannot ignore the possibility that a tool was developed through the use of resources that did not belong to the developer and to which it had no legal or moral claim. The Eighth Commandment still binds Christians, even if the theft in question was carried out indiscriminately on a worldwide scale. At the same time, pragmatism requires a recognition that many aspects of our daily lives in society—including, for most of us, the very land upon which we live and work—were developed or acquired without the consent of those whose labor contributed to their foundation or to whom the land originally belonged. It is not realistic to require Christians to abstain from the fruit of the countless poisoned trees in our world. But nor are we permitted to merely shrug our shoulders and continue on.

A fully formed Christian response to the origins and ongoing use of Gen AI has not been developed even by institutional authorities. But individual Christian users of this technology, if they are satisfied that it was developed through illegal or immoral means, should strongly consider ways in which they can make reparations or offer some form of recompense, perhaps through support of or donations to nonprofit entities that compensate or advocate for the rights of content creators. An imperfect sacrifice, perhaps, but better than leaving the altar bare of any sacrifice at all.

3. *The Accompanying Lawyer*

In an earlier section, I noted that one of the three activities suggested as core occupations for lawyers in the post-AI era is accompaniment—"to join a client on their journey, providing steady support and companionship, for as long as the client wants and for whatever the client might need." The Christian subtext here should not be difficult for readers to notice. Indeed, the inspiration for choosing "accompaniment" as a core lawyer function was the author's involvement with a missionary organization called Catholic Christian Outreach (CCO). CCO encourages its missionaries to carry out their activities through "Intentional Accompaniment," which it describes as "a model of evangelization that

9 Colleen V. Chien & Miriam Kim, *Generative AI and Legal Aid: Results from a Field Study and 100 Use Cases to Bridge the Access to Justice Gap,* Loy. L.A. L. Rev. (forthcoming 2024).

10 Cassandre Coyer, *The Debate on Data Scraping Was Almost Over—Until Generative AI Rekindled It*, ALM (May 16, 2024), https://www.law.com/legaltechnews/2024/05/16/the-debate-on-data-scraping-was-almost-over-until-generative-ai-rekindled-it/.

11 Michael M. Grynbaum & Ryan Mac, *The Times Sues Open AI and Microsoft Over A.I. Use of Copyrighted Work*, N.Y. Times (Dec. 27, 2023), https://www.nytimes.com/2023/12/27/business/media/new-york-times-open-ai-microsoft-lawsuit.html.

focuses on walking alongside others on their journey toward conversion, then remaining alongside them after conversion as they begin a life of missionary discipleship."[12] CCO drew its model of "accompaniment" from the two disciples whom Jesus joined on the road to Emmaus.

Now, to be clear, the use of missionary terminology here is not meant to suggest that Christian lawyers should attempt to actively evangelize their clients. The overriding ethical duty of Christian lawyers is to fulfill their professional obligation by serving their clients' interests with integrity. A commitment to accompany their clients on their legal journeys is a professional commitment first and foremost. But both a legal career and a faith commitment can be considered a "vocation," in that the person who undertakes one often does so in response to what they feel is a calling of service to others for a higher purpose.

Much has been written about the duty of the Christian lawyer to respond to God's call insofar as it relates to the lawyer's personal conduct and ethical choices. But relatively little has been said about the duty (and opportunity) of the Christian lawyer to express their faith and fulfill their vocation through *relationships with others*, most notably their clients. At a time when society is suffering a quiet crisis of loneliness and alienation, the personal commitment that a lawyer makes to their client—not just to assist them and act on their behalf, but also to listen to them, acknowledge them, and validate them through the accompaniment of a professional relationship—is also a repudiation of our culture of isolation culture.

Our clients are also our neighbors, and as C.S. Lewis once wrote: "Next to the Blessed Sacrament itself, your neighbor is the holiest object presented to your senses."[13] To accompany our clients on their journeys, which Gen AI might allow us to do more frequently, surely would be a deeply Christian act as much as the highest fulfillment of our professional duty.

Conclusion

Gen AI holds out the promise of a new era for the legal profession, enhancing the efficiency and effectiveness of legal services and allowing lawyers to focus on higher-value activities, greater creativity, and strategic thinking in service to clients. This transformation, however, will require us to reimagine the traditional architecture of legal services—shifting law firm business models from time-based to value-based and redefining lawyers to advocate, advise, and accompany clients through judgment, empathy, and ethical commitment. The potential of Gen AI to democratize access to legal services represents a significant step toward addressing long-standing inequities in the legal system.

For Christians, Gen AI offers both a unique opportunity to fulfill the moral imperative of promoting justice and a warning to respect ethical standards and human dignity. Lawyers must ensure this technology serves people's interests and enhances societal well-being. By embracing the potential of AI while upholding the core values of the legal profession, lawyers can contribute to a more just and equitable future.

---

12 *The Intentional Accompaniment Series*, Proclaim, https://weareproclaim.com/resources/intentional-accompaniment-series (last visited Aug. 4, 2024).

13 C.S. Lewis, The Weight of Glory and Other Addresses 46 (1949).

More Than Machines: The Ethical and Human Implications of Generative AI on Lawyering

*by Larry O. Natt Gantt, II**

Introduction

The release of OpenAI's ChatGPT on November 30, 2022, exploded the use and popularity of generative artificial intelligence in various sectors around the world. Although artificial intelligence has been in use for many years, the creative capabilities of generative AI captured the attention of many individuals in diverse sectors ranging from higher education to business to medicine.

Legal education and the legal profession were quickly caught up in this wave, as law school professors and legal profession leaders began opining about how generative AI would disrupt legal education and the practice of law. Scripture reminds us that "there is nothing new under the sun,"[1] yet we must recognize that technological advances and other societal changes affect the form—although not perhaps the nature—of the challenges we face in this world.

In this article, I discuss how generative AI is indeed impacting the form—although not perhaps the nature—of the ethical challenges attorneys face, and I then summarize those ethical challenges. I next discuss the more fundamental question of how generative AI has understandably caused a reexamination of what it means to be a lawyer—and particularly in this context, what it means to be a *Christian* lawyer.

Artificial Intelligence (AI) Defined

Although the legal definition of AI is evolving, the emerging definition comes from current regulatory frameworks, such as the European Union Artificial Intelligence (EU AI) Act, formally adopted by the European Council on May 21, 2024, and the Biden Executive Order on AI, issued on October 30, 2023. For instance, the Biden Executive Order defines AI as:

> A machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments. AI systems use machine- and human-based inputs to perceive real and virtual environments; abstract such perceptions into models through analysis in an automated manner; and use model inference to formulate options for information or action.[2]

The Order defines *generative* AI as "the class of AI models that emulate the structure and characteristics of input data in order to generate derived synthetic content. This can include images, videos, audio, text, and other digital content."[3] This technology therefore goes beyond earlier forms of AI to generate products, such as textual responses or images, in response to a user's request based on the large set of existing data on which the program has been trained. This technology specifically employs machine learning models called large language models (LLMs) that process the user's request and are designed to generate outputs that resemble human-created content.

Ethical Implications of Lawyers' Use of Generative AI

Given that lawyers are often called upon to produce written content for clients and others, it was only a matter of time before lawyers would em-

* Professor of Law and Associate Dean for Academic Affairs, High Point University Kenneth F. Kahn School of Law.

1 *Ecclesiastes* 1:9b. The full verse states: "What has been will be again, what has been done will be done again; there is nothing new under the sun." All Bible quotations are to the New International Version unless otherwise noted.

2 Exec. Order No. 14110, 88 Fed. Reg. 75,191 (Oct. 20, 2023).

3 *Id.*

ploy generative AI to help draft written products like court filings and client contracts. Such use quickly caught the attention of judges and bar regulators because generative AI programs notoriously produce content, usually termed "hallucinations," that is fabricated but is presented as real authority.[4] Another significant problem in using generative AI is AI bias, in which such programs have been shown to produce text and images that perpetuate biases relating to race, gender, political affiliation, and other factors.[5]

These problems, particularly hallucinations, have already led attorneys to be disciplined and have garnered significant media attention. In one of the first such cases, *Mata v. Avianca, Inc.,* two lawyers were fined $5,000 for submitting a brief that included fictitious legal research generated by the AI model ChatGPT.[6] The judge in the case ruled that the lawyers acted in bad faith by relying on the AI-generated research without verifying its accuracy.[7] In a later case, *People v. Zachariah C. Crabill,* the Colorado Supreme Court suspended Crabill for a year and a day, with ninety days to be served and the remainder to be stayed upon his successful completion of a two-year probation period, for filing a motion that included fictitious case law generated by ChatGPT.[8] He failed to verify the information before submission and initially blamed an intern for the error when questioned by the judge.[9]

Despite the novel context in which these attorney discipline cases arise, AI technology implicates many of the same ethical duties, such as competence, diligence, confidentiality, proper supervision, and independent professional judgment, that apply to lawyers' use of other forms of technology. Indeed, lawyers who use generative AI to complete legal tasks must consider the same ethical implications as if they were overseeing another nonlawyer completing those tasks; the ethical issues themselves again are not new, just the context. The relative straightforward nature of these legal ethics issues is highlighted in a 2023 article in the *North Carolina State Bar Journal* in which the author quoted ChatGPT's own response to the question "What are the ethical considerations for a lawyer's use of artificial intelligence in a law practice?" and then recognized that its answer generally acceptably summarized those issues.[10] Moreover, on July 29, 2024, the ABA Standing Committee on Ethics and Professional Responsibility issued Formal Opinion 512 in which it discussed the ethical implications of using generative AI.[11]

The primary legal ethics issues implicated by using generative AI include these issues summarized below.

1. *Competence & Diligence*

The ABA Model Rules and many state rules of professional conduct expressly include comments providing that lawyers' general duty of competence extends to their use of technology like artificial intelligence. According to these comments, lawyers who use AI must understand the "benefits and risks" associated with such technology.[12] Lawyers therefore risk violating their duty of competence if they avoid using AI when that use would benefit their provision of legal services.[13] Moreover, if they use the technology, they must have at least a "reasonable understanding" of how the technology functions

4 *See When AI Gets It Wrong: Addressing AI Hallucinations and Bias*, MIT Mgmt., https://mitsloanedtech.mit.edu/ai/basics/addressing-ai-hallucinations-and-bias/ (last visited June 22, 2024).

5 *Id.*

6 Mata v. Avianca, Inc., 22-cv-1461 (PKC) (June 22, 2023).

7 Sarah A. Emmerich, *Artificially Unintelligent: Attorneys Sanctioned for Misuse of ChatGPT*, Minding Your Bus. (June 20, 2023), https://www.mindingyourbusinesslitigation.com/2023/06/artificially-unintelligent-attorneyssanctioned-for-misuse-of-chatgpt/.

8 People v. Crabill, 2023 WL 8111898 (Colo. O.P.D.J. Nov. 22, 2023).

9 *Id.*

10 Brian Oten, *Artificial Intelligence, Real Practice*, 28 N.C. St. Bar J. 6, 6-7 (2023).

11 ABA Standing Committee on Ethics & Professional Responsibility, Formal Opinion 512: Generative Artificial Intelligence Tools (2024), https://www.americanbar.org/content/dam/aba/administrative/professional_responsibility/ethics-opinions/aba-formal-opinion-512.pdf.

12 Model Rules of Prof'l. Conduct r. 1.1 cmt. 8 (Am. Bar Ass'n 2024) [hereinafter MRPC].

13 ABA Standing Comm. on Ethics & Prof'l Responsibility, Formal Op. 512, *supra* note 11.

and how its use can result in ethical problems.[14] The recent ABA opinion stresses that, given the fast-paced nature of technological change, maintaining this level of understanding is "not a static undertaking."[15] Ethics opinions acknowledge that lawyers and law firms often will need to rely on consultants generally in the use of technology, but caution that in doing so the lawyers still have an independent responsibility to ensure ethical standards are satisfied.[16]

Related to competence, lawyers' ethical duty of diligence requires them to exercise "reasonable diligence" in representing a client.[17] Competence and reasonable diligence, in turn, require that lawyers do not so heavily rely on AI tools that they fail to provide the proper human oversight to review AI output and ensure adequate client representation.[18]

2. *Confidentiality*

The ABA Model Rules and many state rules also expressly provide that lawyers' duty of confidentiality requires they undertake "reasonable efforts" to prevent the disclosure of confidential client information.[19] This duty applies to lawyers' submission of confidential information to AI programs like ChatGPT, Claude, and Gemini for lawyering projects. Lawyers thus must undertake such efforts to ensure client information is not improperly disclosed in training the AI or through data breaches or through improper commingling with other data.[20]

The ABA Model Rules and many state rules outline in their comments several factors used in assessing the reasonableness of the lawyers' efforts and precautions, such as "the sensitivity of the information, the likelihood of disclosure if additional safeguards are not employed, [and] the cost of employing additional safeguards."[21] At a minimum, lawyers must understand the terms of use and privacy policy of any program they utilize. For instance, confidential client information should not be uploaded into ChatGPT, as Open AI's privacy policy provides several ways that it uses and discloses users' personal information and transaction history.[22] The confidentiality concerns surrounding self-learning generative AI tools led the recent ABA opinion to conclude that lawyers should obtain informed consent from their clients before inputting confidential client information into such tools.[23]

3. *Duty of Supervision*

With the rise of AI, much legal commentary has already been devoted to how AI systems, particularly generative AI, can perform certain legal tasks that junior lawyers and paraprofessionals, like paralegals, have traditionally performed. Lawyers have duties in ABA Model Rules 5.1 and 5.3, and their state counterparts, to supervise lawyers and nonlawyers properly; and the comments to Rule 5.3 clarify that the duties of proper nonlawyer supervision can apply to law-

14 *Id.*

15 *Id.*

16 *See id.*; ABA Standing Committee on Ethics & Professional Responsibility, Formal Opinion 495: Lawyers Working Remotely (2020), https://www.lawnext.com/wp-content/uploads/2021/09/aba-formal-opinion-495.pdf; ABA Standing Committee on Ethics & Professional Responsibility, Formal Opinion 498: Virtual Practice (2021).

17 MRPC r. 1.3 (Am. Bar Ass'n 2024).

18 ABA Standing Comm. on Ethics & Prof'l Responsibility, Formal Op. 512, *supra* note 11.

19 MRPC r. 1.6 (Am. Bar Ass'n 2024) (defining confidential information as "information relating to the representation of a client").

20 Note that these concerns also relate to potential implications on the attorney-client privilege.

21 MRPC r. 1.6 cmt. 18 (Am. Bar Ass'n 2024). The comments to Virginia Rule of Professional Conduct 1.6 also provide detailed information on steps lawyers should consider taking to protect confidential client information. In a context analogous to AI programs, ABA Formal Opinion 498 provides detailed guidance on the considerations necessary to ensure the protection of a client's files and communications when lawyers use vendors to provide cloud storage. *See* ABA Standing Comm. on Ethics & Prof'l Responsibility, Formal Op. 498, *supra* note 16. For instance, the opinion adds that the lawyer must take steps to ensure the vendor regularly backs up any client data stored with the vendor. *Id.*

22 *See Privacy Policy*, OpenAI, https://openai.com/policies/privacy-policy/ (Nov. 14, 2023).

23 *See* ABA Standing Comm. on Ethics & Prof'l Responsibility, Formal Op. 512, *supra* note 11.

yers' use of technology to facilitate the representation.[24]

Specifically, the relevant language in Rule 5.3 requires lawyers to "make reasonable efforts" to ensure that the actions in which technologies like AI engage are "compatible with the professional obligations of the lawyer."[25] This broad language underscores that lawyers remain independently responsible for their use of generative AI tools and cannot "blame" misconduct on the technology. Moreover, because lawyers cannot provide ethical direction to AI tools the same way they might communicate such direction to nonlawyer personnel, lawyers must be vigilant to understand how generative AI works in creating the content lawyers might use.

4. *Unauthorized Practice of Law*

Related to lawyers' duty to supervise nonlawyers properly is the duty that lawyers cannot ethically delegate certain tasks to a nonlawyer and cannot assist a nonlawyer in the unauthorized practice of law (UPL).[26] At the same time, as noted above, lawyers' duty of competence and diligence encourages lawyers not to "under-delegate" tasks to AI when such delegation would improve their provision of legal services.[27]

An interesting AI development related to UPL is legal chatbots. These are "AI-powered programs that interact with users who have legal issues by simulating a conversation or dialogue."[28] Consumers are accessing these chatbots to perform law-related tasks like "fight[ing] parking tickets, advis[ing] victims of crimes, or draft[ing] privacy policies or non-disclosure agreements."[29] When lawyers create or maintain these tools, the question arises whether the lawyers are assisting another, here AI-powered technology, in engaging in the unauthorized practice of law. Lawyers therefore must not improperly delegate certain tasks to such tools because the technology is unable to exercise the independent professional judgment and provide the nonlegal counseling needed in many legal situations.[30]

5. *Communication*

Another ethical issue relating to the use of AI concerns lawyers' duty to keep their clients "reasonably informed" about their matters and to "reasonably consult with the client about the means by which the client's objectives are to be accomplished."[31]

From this duty, the question arises whether lawyers must consult with their clients when they use AI, particularly generative AI, to conduct the tasks needed to represent the client. Lawyers, of course, do not generally need to consult with their clients when they use technology to assist in the representation because such assistance should reasonably be assumed. As potential uses for technology expand, however, and take on tasks traditionally performed by humans, such delegation to AI resembles outsourcing client work to nonlawyers. Some

---

24 *See* MRPC r. 5.3 cmt. (Am. Bar Ass'n 2024). Pursuant to Rules 5.1 and 5.3, lawyers' duty of supervision also includes the responsibility to train subordinate lawyers and nonlawyers on how to comply with lawyers' ethical obligations when using AI. *See* ABA Standing Comm. on Ethics & Prof'l Responsibility, Formal Op. 512, *supra* note 11.

25 MRPC r. 5.3(a-b) (Am. Bar Ass'n 2024).

26 MRPC r. 5.5 (Am. Bar Ass'n 2024).

27 *See* David Lat, *The Ethical Implications of Artificial Intelligence*, Above the L., https://abovethelaw.com/law2020/the-ethical-implications-of-artificial-intelligence/?rf=1 (last visited June 28, 2024).

28 *Id.*

29 *Id.*

30 MRPC r. 2.1 (Am. Bar Ass'n 2024) ("In representing a client, a lawyer shall exercise independent professional judgment and render candid advice. In rendering advice, a lawyer may refer not only to law but to other considerations such as moral, economic, social and political factors, that may be relevant to the client's situation."); *see also* ABA Standing Comm. on Ethics & Prof'l Responsibility, Formal Op. 512, *supra* note 11.

31 MRPC r. 1.4 (Am. Bar Ass'n 2024).

authorities have opined that such "outsourcing" to AI requires client consent when confidential client information is involved.[32] Moreover, if the lawyers' use of AI materially impacts the lawyers' fee, the general allocation of authority between clients and lawyers supports discussing these impacts with the client.[33] These general considerations led the recent ABA ethics opinion to conclude that lawyers should consult with their clients about their use of generative AI any time "its output will influence a significant decision in the representation."[34]

6. *Fees and Billing*

Finally, lawyers' use of generative AI potentially impacts their duty to ensure their fees are "not unreasonable."[35] With the capacity of AI to enable lawyers to complete certain legal tasks much more quickly, lawyers must ensure their fees remain ethically compliant. Lawyers who bill hourly therefore cannot charge for time they would have spent on a matter but no longer need to because of their use of AI. Lawyers, of course, remain free to charge clients through other billing methods, such as flat fees, as long as the fee remains "not unreasonable" and does not mispresent the time the lawyer spent on the matter.[36]

Implications of AI on the Role of Christian Lawyering

The ethical implications discussed above must undoubtedly be considered by lawyers as they examine how best they can use generative AI in their practice while satisfying their ethical responsibilities. As the Preamble to the ABA Model Rules outlines so well, lawyers must recognize that these responsibilities flow "to clients, to the legal system and to the lawyer's own interest in remaining an ethical person while earning a satisfactory living."[37] In addition to these ethical implications, however, the rise of generative AI fuels a more fundamental inquiry: As AI overtakes many of the technical aspects of lawyering,

32 For instance, the State Bar of California Standing Committee on Professional Responsibility and Conduct has issued guidance stating that lawyers "should consider" informing their clients if generative AI tools will be used as part of their representation. State Bar of California Standing Committee on Professional Responsibility and Conduct, Practical Guidance for the Use of Generative Artificial Intelligence in the Practice of Law (2023), https://www.calbar.ca.gov/Portals/0/documents/ethics/Generative-AI-Practical-Guidance.pdf; *see also Ethics Opinion: Opinion 24-1*, The Fl. Bar (Jan. 19, 2024), https://www.floridabar.org/etopinions/opinion-24-1/ ("[I]t is recommended that a lawyer obtain the affected client's informed consent prior to utilizing a third-party generative AI program if the utilization would involve the disclosure of any confidential information."); *2007 Formal Ethics Opinion 12*, N.C. State Bar (Apr. 25, 2008), https://www.ncbar.gov/for-lawyers/ethics/adopted-opinions/2007-formal-ethics-opinion-12/. (allowing outsourcing legal tasks to third parties "provided the lawyer properly selects and supervises the foreign assistants, ensures the preservation of client confidences, avoids conflicts of interests, discloses the outsourcing, and obtains the client's advanced informed consent").

33 *See* MRPC r. 1.2, cmt. 2 (Am. Bar Ass'n 2024) (noting that "lawyers usually defer" to clients regarding the means of representation when those means impact the expenses incurred).

34 *See* ABA Standing Comm. on Ethics & Prof'l Responsibility, Formal Op. 512, *supra* note 11 (providing several such instances, including when lawyers use generative AI "to evaluate potential litigation outcomes or jury selection").

35 *See* MRPC r. 1.5(a) (Am. Bar Ass'n 2024). Certain state rules have slightly different standards regarding fees than ABA Rule 1.5. For instance, North Carolina Rule 1.5 provides that lawyers shall not charge "clearly excessive" fees or expenses. N.C. Rules of Prof'l Conduct r. 1.5(a) (2024).

36 In addition to their base fee, lawyers generally may not charge clients an overhead or administrative fee in which they roll in the firm's general expenses for office-related costs, which could include costs associated with technology (like AI). Although decided well before the advent of AI technology in lawyering, a 1993 ABA Formal Ethics Opinion includes reasoning that clearly applies to billing for lawyers' use of such technology. Specifically, the opinion reasoned that lawyers cannot charge clients general office overhead absent disclosure to the client in advance of the engagement. The opinion also prohibits surcharges on expense disbursements above the amount actually incurred in directly representing the client, absent disclosure to the client. From this reasoning, lawyers who seek to pass along AI costs to their clients must not do so through a general administrative fee, unless they disclose this fee to the client prior to the engagement. Similarly, absent client consent, they cannot bill clients for AI services over the amount those services cost the lawyer (including any direct cost and cost for allocated expenses) to provide the specific work the lawyer dedicates to the client. *See* ABA Standing Comm. on Ethics & Prof'l Responsibility, Formal Op. 512, *supra* note 11. *Cf. 2022 Formal Ethics Opinion 4,* N.C. State Bar (Oct. 27, 2023), https://www.ncbar.gov/for-lawyers/ethics/adopted-opinions/2022-formal-ethics-opinion-4/ (addressing billing of expenses to clients).

37 MRPC Preamble (Am. Bar Ass'n 2024).

what are the *human* aspects of lawyering that will become the essence of lawyering in the future?[38]

This recognition of the importance of the human aspects of lawyering is not new. Indeed, the well-being crisis in the legal profession has underscored that developing lawyers must cultivate not only doctrinal knowledge and practical skills, but also their "professional identity" in which they discern how their personal values align with their professional calling.[39] Generative AI, however, has accelerated this discussion as focusing on the human aspects of lawyering becomes more than a question of educational best practices; it becomes central to the issue of lawyers' market relevance. Indeed, unlike other recent technological breakthroughs like the internet which caused lawyers to reexamine the practice of law, generative AI feels different. Previous technologies facilitated lawyers' ability to communicate with others and access information; generative AI, in contrast, *creates* content designed to mimic human-centered lawyering.

Legal scholars and futurists have opined in response that lawyers must become more "people-focused" and concentrate on "build[ing] professional relationships and trust in ways that machines may never be able to replace."[40] They also recommend that lawyers hone their "'soft' human skills like communication and creativity."[41]

These recommendations are certainly well-taken, as experts agree that current AI technology is not sentient to any meaningful degree.[42] These same experts, however, recognize that "rapid advances in AI technology could soon create AIs of plausibly debatable sentience and moral standing, at least by some relevant definitions."[43] Moreover, recent research indicates that scientists are developing new algorithms to help AI models minimize hallucinations and increase their reliability.[44] What then when AI becomes arguably sentient and sufficiently reliable? Might a client be able to develop an attorney-client relationship with a robot?

I understand that such questions seem fantastic and far-fetched, but as technological advances expand, Christian lawyers need to affirm a principled approach to lawyering that underscores the fundamental truth that humans are created in the image and likeness of God.[45] Machines will never be. From this core theological truth flows eternal principles about human-to-human relationships, many of which are beyond the scope of this article.[46] For Christian lawyers who must serve their clients and the legal system, certain of these truths become paramount. I discuss three of these below: advocacy, empathy, and wisdom.

1. *Advocacy*

I have often heard others reference 1 John 2:1 as support for the role of the Christian lawyer: "My dear children, I write this to you so that you will not sin. But if anybody does sin, we have an advocate with the Father—Jesus Christ, the Righteous One." Although Christ's advocacy for humankind does not mirror Christian lawyers' for their clients—after all, Jesus not only advocated for sinners, He also took our punishment—the principle of Christian advocacy certainly permeates Scripture. The Bible is replete with passages that call believers to advocate for those

38 *See* L.O. Natt Gantt, II, *Law Schools' Pivotal Role in Lawyer-Leader Formation*, Law Prac. Mag. (Feb. 5, 2024), https://www.americanbar.org/groups/law_practice/resources/law-practice-magazine/2023-november-december/lawschools-pivotal-role-in-lawyer-leader-formation/.

39 *See id.*

40 *See* Bernard Marr, *How Generative AI Will Change the Jobs of Lawyers*, Forbes (Mar. 14, 2024), https://www.forbes.com/sites/bernardmarr/2024/03/14/how-generative-ai-will-change-the-jobs-of-lawyers/.

41 *Id.*

42 Eric Schwitzgebel, *AI Systems Must Not Confuse Users About Their Sentience or Moral Status*, Patterns 4 (Aug. 11, 2023), https://www.cell.com/action/showPdf?pii=S2666-3899%2823%2900187-3.

43 *Id.*

44 *See* Billy Perrigo, *Scientists Develop New Algorithm to Spot AI "Hallucinations,"* Time (June 19, 2024), https://time.com/6989928/ai-artificial-intelligence-hallucinations-prevent/.

45 *Genesis* 1:26-27.

46 For a helpful discussion of theological implications of our being created in God's image, see Chapter 24 ("The Image of God in the Human") in Millard J. Erickson, Christian Theology (2d ed. 1998). See also the piece from Jason Thacker in this volume.

in need, perhaps most famously in Proverbs 31:8-9: "Speak up for those who cannot speak for themselves, for the rights of all who are destitute. Speak up and judge fairly; defend the rights of the poor and needy." Moreover, from Moses to Daniel to Esther to Paul, men and women in Scripture exemplify that calling of advocating for others in need.[47]

Generative AI is becoming increasingly adept at creating content that can be used to advocate for others, but it cannot—and will not—serve as the advocate. Lawyers uniquely serve that role. They are the ones to select the causes in which they contend, and Christian lawyers can select their causes with a focus on promoting justice, as defined by the moral law of Scripture.[48] Christian lawyers can also approach their role as advocate with a Christ-like spirit of excellence, service, and humility. Colossians 3:23 reminds Christian lawyers to represent their clients and engage in their work "with all your heart, as working for the Lord, not for human masters." Finally, Christian lawyers can conduct their advocacy with an appreciation that they are representing individuals who are image-bearers of God or entities that work and operate through such image-bearers.

2. *Empathy*

Just as Christian lawyers are able to advocate for their clients in ways machines cannot, this appreciation that clients and others are similarly created in God's image demands they be treated with dignity, respect, and love. My former Regent colleague Ben Madison and I have written much on how good lawyers are empathic lawyers.[49] Empathic lawyers are able to better understand how others are affected by their decisions and therefore are able to better evaluate competing options in their decision-making.[50] For Christian lawyers, seeing others as image-bearers of God deepens the meaning of empathy. C.S. Lewis powerfully recognizes the import of our status as image-bearers in his oft-quoted passage in *The Weight of Glory:*

> It is a serious thing to live in a society of possible gods and goddesses, to remember that the dullest most uninteresting person you can talk to may one day be a creature which, if you saw it now, you would be strongly tempted to worship, or else a horror and a corruption such as you now meet, if at all, only in a nightmare. All day long we are, in some degree helping each other to one or the other of these destinations. It is in the light of these overwhelming possibilities, it is with the awe and the circumspection proper to them, that we should conduct all of our dealings with one another, all friendships, all loves, all play, all politics. There are no ordinary people. You have never talked to a mere mortal. Nations, cultures, arts, civilizations—these are mortal, and their life is to ours as the life of a gnat. But it is immortals whom we joke with, work with, marry, snub, and exploit—immortal horrors or everlasting splendors.[51]

As Robert Cochran explores in his book *The Servant Lawyer*, Lewis' emphasis that no one is a "mere mortal" reminds us to treat others with due regard to their spiritual identity.[52] Cochran recognizes that Christian lawyers should show

---

47 *See* World Vision, *What Does the Bible Say About Advocacy?,* Nat'l Christian Found. (Jan. 30, 2019), https://www.ncfgiving.com/stories/what-does-the-bible-say-about-advocacy/.

48 *See Micah* 6:8 ("He has showed you, O man, what is good; and what does the LORD require of you but to do justice, and to love kindness, and to walk humbly with your God?") (RSV).

49 *See* Benjamin V. Madison, III & Larry O. Natt Gantt, II, *The Emperor Has No Clothes, But Does Anyone Really Care?* How *Law Schools are Failing to Develop Students' Professional Identity and Practical Judgment*, 27 Regent U. L. Rev. 339, 386-90 (2014-15).

50 *Id.* at 390.

51 C.S. Lewis, The Weight of Glory and Other Addresses 45-46 (2001).

52 Robert F. Cochran, Jr., The Servant Lawyer: Facing the Challenges of Christian Faith in Everyday Law Practice 8-20 (2024).

others respect and love, making sure that we listen to our clients and appreciate the difficulties they are facing.[53]

With the rise of generative AI, the lawyer's ability to understand and appreciate the client—and the ability to express in a gracious way that understanding and appreciation—become even more critical. Generative AI will become increasingly skilled at accessing legal information and creating legal documentation. It will never, however, be able to empathize with others as a fellow human. We as lawyers must cultivate a sense of presence so we can *abide with* others in crisis, including being present with our clients in the midst of the legal problems they face.

3. *Wisdom*

By its very name, artificial intelligence evokes concerns over machines that become so intelligent they outsmart humans. Think HAL 9000 in the classic 1968 movie *2001: A Space Odyssey.* These concerns conflate intelligence with wisdom. Artificial intelligence computing is defined as "the math-intensive process of calculating machine learning algorithms, typically using accelerated systems as software."[54] Wisdom, however, is much deeper and richer than math computations and algorithms. Theologians recognize that our status as image-bearers of God includes an "inner sense of right and wrong" and an ability to obtain wisdom that is inaccessible to other created things.[55]

In discussing the importance of wisdom for the good lawyer, much of the recent literature on lawyers' professional formation has underscored the classic Aristotelian virtue of *phronesis*, or practical wisdom. "Aristotle viewed practical wisdom as the cardinal virtue implicit in the other virtues, including courage, self-control, fairness, gentleness, loyalty, friendliness, and honesty. Significantly, Aristotle's emphasis was on one who knew how to employ these virtues 'practically' in human affairs in order to be wise."[56] Such wisdom combines intellectual understanding with moral and ethical sensitivity and facilitates an individual's ability to resolve dilemmas when competing virtues (such as honesty and loyalty) appear in conflict.[57] Moreover, such wisdom guides individuals not just to what to do in a situation of uncertainty, but also *how* to do it.[58]

Beyond Aristotle's practical wisdom, exercising *biblical* wisdom enables Christian lawyers to proceed on a fundamentally different level computers can never attain. Although Christians are not able to discern the complete wisdom of God,[59] theologians recognize that "God's wisdom is, of course, in part communicable to us."[60] It begins with reading and obeying His Word.[61] Scripture specifically adds that we discern wisdom from fearing God and seeking to please and depend on Him.[62] God promises that He will grant us wisdom when we ask: "If any of you lacks wisdom, you should ask God, who gives

---

53 *Id.* (referencing *James* 1:19b: "Everyone should be quick to listen, slow to speak and slow to become angry.").

54 Rick Merritt, *What is AI Computing?*, NVIDIA (Jan. 20, 2023), https://blogs.nvidia.com/blog/what-is-aicomputing/.

55 Wayne Grudem, Systematic Theology: An Introduction to Biblical Doctrine 445-46 (1994).

56 Madison & Gantt, *supra* note 49, at 346.

57 *Id.* at 346-47; *see also* Patrick Emery Longan et al., The Formation of Professional Identity: The Path from Student to Lawyer 16-17 (2d. ed. 2024).

58 Longan et al., *supra* note 57, at 17.

59 *See Isaiah* 55:8-9 ("'For my thoughts are not your thoughts, neither are your ways my ways,' declares the LORD. 'As the heavens are higher than the earth, so are my ways higher than your ways and my thoughts than your thoughts.'"); *Romans* 11:33 ("Oh, the depth of the riches of the wisdom and knowledge of God! How unsearchable his judgments, and his paths beyond tracing out!").

60 Grudem, *supra* note 55, at 194.

61 *See Psalm* 19:7 ("The law of the Lord is perfect, refreshing the soul. The statutes of the Lord are trustworthy, making wise the simple.").

62 *See Proverbs* 9:10 ("The fear of the Lord is the beginning of wisdom, and knowledge of the Holy One is understanding.").

generously to all without finding fault, and it will be given to you."[63] Moreover, the presence of the power of the Holy Spirit, which the Apostle Paul calls the "Spirit of wisdom,"[64] facilitates in the lives of believers their access to the wisdom of God. Finally, the fellowship of godly believers and appreciation of Christians' witness throughout history can enhance Christian lawyers' wisdom, for "[a]s iron sharpens iron, so one person sharpens another."[65] In the face of smarter and smarter computers, Christian lawyers must therefore renew their focus on appropriating godly *wisdom* in their practice.

Conclusion

Generative AI represents a groundbreaking technology that is set to disrupt legal education and the legal profession. Its ethical implications, although not fundamentally different from the implications of other technological advances, do lead lawyers to face these implications in new contexts where computers perform legal tasks previously done by lawyers or paraprofessionals. Lawyers must thus be vigilant in recognizing the potential ethical pitfalls when they use generative AI.

In the midst of this disruption, however, the more foundational issue concerns how generative AI will change what it means to be a lawyer. Understanding the uniquely human qualities of advocacy, empathy, and wisdom in the age of generative AI underscores that, now more than ever, Christian law students and lawyers must seek to cultivate these qualities in their lives. We must put ourselves in positions where we better appreciate and understand the needs of others,[66] and we must pray to have the heart of God that pursues justice and manifests the fruit of the Spirit.[67] In the end, this technological advancement may surprisingly lead to revival in our core mission as Christian lawyers to love God and love others through our vocation.

---

63 *James* 1:5.

64 *Ephesians* 1:17 ("I keep asking that the God of our Lord Jesus Christ, the glorious Father, may give you the Spirit of wisdom and revelation, so that you may know him better.").

65 *Proverbs* 27:17.

66 Putting ourselves in such positions may occur in a variety of ways, from representing indigent clients to serving the community through nonprofit ministries.

67 The fruit indeed are moral qualities that generative AI cannot embody: "But the fruit of the Spirit is love, joy, peace, forbearance, kindness, goodness, faithfulness, gentleness and self-control. Against such things there is no law." *Galatians* 5:22-23.

Instructing AI Through the Exercise of Labor's Solidarity: A Christian Perspective

*by Alvin Velazquez**

Introduction

To the average reader, what Gretchen Huizinga calls the "current 'Holy Trinity' of Artificial Intelligence (AI) innovation—large data sets, sophisticated algorithms, and unprecedented compute power"—appear to bring with it a new set of problems for those interested in the study of Christian theology to think about.[1] However, the Book of Ecclesiastes in the Old Testament teaches that there is nothing new under the sun.[2] While AI is a new development when examined in the arc of human history, the workplace and property law questions it implicates are not.[3] Depending on the intentions of the user, AI can facilitate the theft of another person's labor,[4] or it could help doctors diagnose illnesses.[5] This essay is limited to starting a conversation about the use of generative AI by employers in the workplace.[6]

Plenty has been written about how AI adoption presents existential questions for large amounts of the workforce, especially white collar and other creative workers. These groups have not been silent in the face of the threat that AI presents. For example, the Writers Guild of America (WGA) went on strike last year over Hollywood's use of AI. They did so over their employers' use of generative AI to write scripts for various sorts of productions and their waning residuals, amongst other things.[7] Ultimately, the WGA settled and obtained several concessions on the use of AI.[8] The WGA used its power in

* Associate Professor of Law, Indiana University Maurer School of Law. My thanks to Michael Oswalt, Asaf Lubin, David Skeel, and Christopher Hampson for their thoughtful comments.

1 Gretchen Huizing, Righteous AI: The Christian Voice in the Ethical AI Conversation 10 (2022) (University Libraries, University of Washington).

2 *Ecclesiastes* 1:9 (NIV).

3 Singularity, or the moment in which the capacity of AI supersedes human thinking, may completely undermine this statement, and the statement in Ecclesiastes 1:9, by inverting the focus away from humanity. My thanks to Asaf Lubin for making this observation. For an easily digestible discussion of the concept of "singularity," *see* David Streitfeld, *Silicon Valley Confronts the Idea That the 'Singularity' Is Here*, N.Y. Times (June 11, 2023), https://www.nytimes.com/2023/06/11/technology/silicon-valley-confronts-the-idea-that-the-singularity-is-here.html.

4 *See, e.g.,* Joe Pompeo, *Inside the Legal Tussle Between Authors and AI: "We've Got to Attack This From All Directions,"* Vanity Fair (Oct. 18, 2023), https://www.vanityfair.com/news/2023/10/legal-tussle-between-authors-ai.

5 *See, e.g.,* Jackie Snow, *The Algorithm Will See You Now: How AI is Helping Doctors Diagnose and Treat Patients*, PBS (Mar. 6, 2019), https://www.pbs.org/wgbh/nova/article/how-ai-is-helping-doctors-diagnose-and-treat-patients/.

6 One definition that Pauline T. Kim and Matthew T. Bodie uses loosely defines AI as "systems that leverage data-rich inputs and computational techniques to make predictions that either aid or replace human decision-making." Pauline T. Kim & Matthew T. Bodie, *Artificial Intelligence and the Challenges of Workplace Discrimination and Privacy*, 35 ABA J. Lab. & Emp. L. 289, 290 (2021). Those predictive elements (including large language models) have been a focus of scholarship in employment law literature. However, I will use Henry H. Perritt, Jr.'s definition. While admitting that it is difficult to derive one definition for AI, he notes that:

> A generative AI system is a computer system or collection of systems that: (a) engages in machine learning by application of neural network transformers to large databases with more than 100,000 data entries, and (b) uses the results of that machine learning to generate new expressive output according to relatively simple and short prompts by a user.

Henry H. Perritt, Jr., *Robot Regulations*, 75 S.C. L. Rev. 219, 248 (2023).

7 Simmone Shah, *The Writers Strike is Taking a Stand on AI*, Time Mag. (May 4, 2023), https://time.com/6277158/writers-strike-ai-wga-screenwriting/.

8 *WGA Negotiations—Tentative Agreement* 6-7 (Sept. 25, 2023), https://www.wgacontract2023.org/WGAContract/files/WGA-Negotiations-Tentative-Agreement.pdf.

collective bargaining to fight back against the existential risk that AI presented to its membership. The union sought to bar its members' materials from being used to train AI but did not obtain that, though the union did at least mitigate many of the harms AI can inflict on its members.[9] By negotiating to end the strike, the union acted as a vessel to obtain workplace justice for its members and vindicate biblically defined wrongs such as theft and wage oppression against their members.[10]

This claim might sound controversial at first, but organized labor unions are well situated for mitigating the harm AI can cause while serving as vehicles for vindicating the distributive and anti-wage theft norms contained in the Christian Bible. The aim of this short essay is to demonstrate that point and begin a revival of the historical conversation that Christian leaders had about labor rights and apply it to modern day problems such as AI.[11] As hard as it may be to believe given current political discourse, the labor movement and the church worked closely at various points in U.S. history. Christian churches and labor worked with each other and grounded that relationship in a shared understanding of the Christian Bible.[12] Even famed evangelist Billy Graham noted the importance of having evangelical Christians work with labor leaders.[13] Over the last 40 years or so, the gulf between evangelical Christians and labor has loosened.[14] Today, United Auto Workers president Shawn Fain is attempting to revive a tradition of biblically informed labor discourse in reviving the rhetoric of what at times has been called the "Christian left" for fusing causes that political commentators view as progressive with theologically based language.[15]

In this piece, I am asking this journal's readers to reimagine what theological inquiry from the "Christian left" can add to modern day discussions about labor law and AI. Perhaps such an exercise might eventually lead to the reimagination of the role of church as society grapples with AI governance and regulation, but for now I will only focus on organized labor's role within it.

---

9 *Id.*

10 *See Matthew* 5:9 ("Blessed are the peacemakers, for they will be called children of God"); *see also Exodus* 20:15 ("Thou shalt not steal.") (KJV); *James* 5:4 ("Look! The wages you failed to pay the workers who mowed your fields are crying out against you. The cries of the harvesters have reached the ears of the Lord Almighty."); *Leviticus* 19:13 ("Thou shalt not defraud thy neighbour, neither rob him: the wages of him that is hired shall not abide with thee all night until the morning."(KJV); *Deuteronomy* 24:14 ("Do not take advantage of a hired worker who is poor and needy, whether that worker is a fellow Israelite or a foreigner residing in one of your towns.") (NIV).

11 *See, e.g.,* Alvin Velazquez, *Drawing on the Christian Tradition as a Source for the Renewal of Labor Law Theory*, 69 St. Louis L.J. (forthcoming 2025) (calling for the reincorporation of theological inquiry on a non-privileged basis into the development of modern labor law theory).

12 *See, e.g.,* Judy Hilovsky, *Celebrating Labor: The Biblical Principles behind Labor Day*, Museum of the Bible, https://www.museumofthebible.org/magazine/impact/celebrating-labor-the-biblical-principles-behind-labor-day (last visited Aug. 16, 2024); *see also* Christopher D. Cantwell et al., *Between the Pew and the Picket Line, in* The Pew and the Picket Line: Christianity and the American Working Class 10 (Christopher D. Cantwell et al. eds., 2016) (asserting that Pentecostal churches were an important partner with the United Mine Workers).

13 Mainline denominations are not the only ones that have had uneven relationships with labor; evangelical leaders have historically tended to avoid engagements as well, even when labor was at the zenith of its influence. *See, e.g.,* Ken Estey, *Billy Graham and the Evangelical Origins of Organized Labor*, LAWCHA (Mar. 13, 2018) https://lawcha.org/2018/03/13/billy-graham-and-the-evangelical-origins-of-organized-labor/ (noting that Graham was disappointed that evangelical preachers were not reaching out to organized labor).

14 Lainey Newman & Theda Skocpol, Rust Belt Union Blues: Why Working-Class Americans Are Turning Away from the Democratic Party 15 (2023) (explaining that as "union halls closed and membership dwindled, other networks and community group influences, often propelling more conservative values and messages, have become more central to the daily lives of workers and residents," such as evangelical churches and gun clubs).

15 *See* Elizabeth Breunig, *Shawn Fain's Old-Time Religion*, The Atl. (Oct. 24, 2023), https://www.theatlantic.com/ideas/archive/2023/10/uaw-union-leadership-american-christian-culture/675741/. *But see* Michael M. Oswalt, *Steeply Solidarity: Mainline Church Renewal and the Union Corporate Campaign*, 50 J. Cath. L. Stu. 227, 229 (2011) (noting that "[t]he growing evangelical religious base also tends to support conservative political candidates who are sympathetic to labor's interests").

Christianity Provides a Foundation for Thinking About Collective Bargaining

In this part, I will explain how the Bible and various Christian denominations provide a foundation for understanding the work of labor unions as a vehicle for vindicating workplace justice. This is a result of the Bible's concern for workplace justice and the condition of the poor. As I have written for the Harvard Living Law Blog:

> God cares about justice generally, but He also cares about workplace conditions specifically (Jeremiah 22:13). God calls out the religiously pious who oppress their workers (Isaiah 58:3). God cares about workers' ability to make a living and consistently states that he hears the cries of workers who are victims of employer wage theft. (James 5:1, 4; Jeremiah 22:13; Deuteronomy 21:14-15).[16]

Additionally, Catholic Social Teaching on unions,[17] writings coming from Social Gospel theology,[18] and Liberation Theology[19] provide a long-standing, well-informed, and deeply theoretical framework to draw upon. Clergy from the Catholic Church such as Monsignor George Higgins,[20] from Protestant denominations such as the Rev. Dr. Teresa Danieley[21] and the late Dr. Martin Luther King, Jr., and Catholic laity such as Dorothy Day[22] have been staunch allies of labor.[23]

Despite the existence of a clear theological foundation through which to analyze labor and AI, there exists a God-shaped theological hole in employment law literature and a union-shaped hole in current Christian thinking about AI's effect on the workplace.[24] The scholarly interest in AI has spawned a large literature in which commentators have explored and advocated for stronger individual employee protections, especially against discrimination against historically marginalized communities.[25] Unfortunately, little of that is coming from an explicitly Christian perspective that seriously examines organized labor's role in mitigating these harms.[26] Religious institutions are issuing statements that call on policymakers to develop measures to ensure that government mitigates AI's disruptions to

16 Alvin Velazquez, *Advocacy for Workers as a Spiritual Calling*, Harv. Living L. Blog (Apr. 11, 2022), https://pblcls.law.harvard.edu/blog/advocacy-for-workers-as-a-spiritual-calling/.

17 *See, e.g.,* Encyclical of Pope Leo XIII On Capital and Labor, Rerum Novarum 49 (1891), https://www.vatican.va/content/leo-xiii/en/encyclicals/documents/hf_l-xiii_enc_15051891_rerum-novarum.html; *see also* Zwick & Zwick, The Catholic Worker Movement: Intellectual and Spiritual Origins 154 (2005) (arguing that the Catholic Worker Movement had a larger, more expansive vision than the labor movement because it viewed labor unions as a means, but not an end).

18 Walter Rauschenbusch, A Theology for the Social Gospel (1918) (setting out a theology that sin and wrong can be both individual and social, and that poverty can be social). *But see, e.g.,* Janine Giordano Drake, *Social Gospel and the American Working Class*, Oxford Rsch. Encyclopedia Am. Hist. (Oct. 26, 2017) (explaining that social gospel pastors wanted to eradicate poverty and would work with labor leaders even though the relationship was fraught due to pastoral emphasis on personal ethics).

19 *See, e.g.,* Gustavo Gutierrez, A Theology of Liberation: History, Politics, and Salvation (1988) (arguing for a preferential option for the poor).

20 Patrick Sullivan, C.S.C, *Monsignor George G. Higgins, The Labor Priests' Priest*, 19 J. Cath. His. 103 (2001).

21 Press Release, Episcopal Diocese of Mo., DioMO Missioner Receives Human Rights Award (Jan. 3, 2024), https://diocesemo.org/connect/news/diomo-missioner-receives-human-rights-award.

22 *See* Zwick & Zwick, *supra* note 17.

23 Rebecca E. Zietlow, *"Where Do We Go From Here?" Dr. Martin Luther King, Jr. and Workers' Rights*, 14 Har. J. L. & Pol'y. Rev. 47, 49 (2019) (describing organized labor as one of King's closest allies).

24 *See, e.g.,* Huizinga, *supra* note 1, at 54 (conducting interviews with tech workers about their faith and their dignity as workers).

25 *See, e.g.,* Ifeoma Ajunwa, *Automated Video Interviewing as the New Phrenology*, 36 Berkeley Tech. L.J. 1173 (2022) (examining the flaws of AI interviews, especially against marginalized communities); Lori Andrews & Hannah Bucher, *Automating Discrimination: AI Hiring Practices and Gender Inequality*, 44 Card. L. Rev 145, 151-52 (2022) (arguing that the use of AI tools currently discriminates along gender lines).

26 *But see* Huizinga, *supra* note 1.

people's working lives but are silent about the role of organized labor in this endeavor.[27]

Much of the literature addresses the "justice paradox" and raises distributive and normative elements of AI, but not through the lens of the Christian legal tradition. In his book, *True Paradox: How Christianity Makes Sense of Our Complex World*, David Skeel defines the "justice paradox." In his view, the justice paradox is when legal codes promise to create a just social order only to see each attempt fail.[28] He provides Marxism's distributive claims as one example of how this dynamic plays. He criticizes its relentless idealism and attempts to get rid of one set of class hierarchies in exchange for another.[29] However, Skeel also finds fault with the American justice system and some of its tendencies to do the same thing.[30] Skeel goes on to explore how Christianity's recognition that humanity is not perfectible is a key starting point to resolving that paradox. In his view, "if all of us are sinful, legal systems must play a double game restraining the worst wrongs by the citizenry without empowering judges and prosecutors to do wrong themselves."[31] Based on this observation, he argues for law with a "light-touch." In essence he is calling for his readers to be less ambitious about what law can do to fix society's ills.[32] Applying his framework to the Wall Street behavior during the 2008 financial crisis, he argues for changes in moral norms as well as law. In his view, law without changes in norms will continue to lead to Wall Street actors taking enormous risks and profiting from them.[33] His book builds off of an earlier academic work with Bill Stuntz arguing for that modest approach to law as most consistent with the ideals of Christianity and the realities of modern legal practice.[34]

I join David Skeel in arguing that Christianity resolves the justice paradox because of its understanding of sin. However, I believe it favors the claims of worker interests as applied to AI. I also agree with Skeel's observations about the nature of Wall Street incentives. The sin of greed that can manifest itself from the pursuit of the incentives that corporate executives face. Those incentives frame Wall Street's decision making and incentivize minimizing employment expenses to maximize shareholder value.[35]

There is an answer though for curbing greed's excesses—organized labor. Labor unions do not abate the sin of greed, but they can blunt the downstream effects of greed upon workers. Labor law's institutional design can play an important role in serving as a vehicle for solving the distributional justice paradox at the center of discussions about how AI affects workers. The next part explains how.

Christianity Makes Sense of Labor's Role in Resolving Distributional Paradoxes

In this section, I argue that labor unions are a "light-touch" mechanism for dealing with distributional paradoxes in particular situations. The law at the center of labor law is the National Labor Relations Act (NLRA or Act). Congress

---

27 *See, e.g.,* Message of His Holiness Pope Francis for the 57th World Day of Peace, Artificial Intelligence and Peace (Jan. 1, 2024), https://www.vatican.va/content/francesco/en/messages/peace/documents/20231208-messaggio-57giornatamondiale-pace2024.html (mentioning the effect of AI on workers but not using the word union despite the Church's historical support of organized labor); Message of His Holiness Pope Francis for the 58th World Day of Social Communications, Artificial Intelligence and the Wisdom of the Heart: Towards a Fully Human Communication (Jan. 24, 2024), https://www.vatican.va/content/francesco/en/messages/communications/documents/20240124-messaggio-comunicazioni-sociali.html (speaking of the need to protect the dignity of workers in a world filled with the rise of artificial intelligence, but failing to reflect or reference previous Catholic Social teaching such as *Rerum Novarum*).

28 *See* David A. Skeel, True Paradox: How Christianity Makes Sense of Our Complex World 109 (2014).

29 *Id.* at 111.

30 *Id.* at 112 (providing Prohibition as one example of America also attempting to use law to meet a moral ideal).

31 *Id.* at 129-30.

32 *Id.* at 130.

33 *Id.* at 131.

34 *See* David A. Skeel, Jr. & William J. Stuntz, *Christianity and the (Modest) Rule of Law*, 8 U. Penn. J. Const. L. 809, 811-812 (2006).

35 *See* Skeel, *supra* note 28, at 132.

passed it to quell massive labor unrest and undermine challenges to Franklin Roosevelt from businesses, on the one hand, seeking to install a dictator and, on the other, communists seeking their own ends.[36] Unlike some reform-minded public laws Congress managed to pass, the NLRA is relatively short.[37] It was the result of a desire for labor peace, and, in many ways, the statute has an apolitical and amoral element to it.[38] As a practical matter, all the Act does is allow the National Labor Relations Board (NLRB) to certify appropriate bargaining units,[39] require employers to respect the right of workers to organize or refrain from organizing,[40] and for the parties to bargain in good faith.[41] On its face, the NLRA does not set out a template for an agreement or allow for the government to take any sides in the negotiations. Because of its design and apolitical morality, labor law exists in the liminal space between the distributional demands between unions, employees generally, and their employers.[43] Unions are supposed to make demands at the bargaining table and use economic weapons such as boycotts and strikes to wrest concessions from management, and management in turn gets to use its tools to ensure that its business remains open without violating the NLRA. The reality is that the NLRA provides scant tools for unions to make serious inroads at the bargaining table because management can rebuff attempts to negotiate over how AI is trained as a "permissive subject of bargaining."[43]

When viewed from the above lens, the WGA's efforts against AI were certainly historic and noteworthy from a public interest standpoint and for achieving as much as they did under the strictures of the NLRA. The WGA used a strike to extract rents from their employer in the form of an increased share of streaming revenues and guarantees that generative AI will not encroach on their credit or compensation.[44] Their work demonstrates that the use of the strike and the process of collective bargaining can result in a different distribution of the fruits of labor, and thus provided one solution to the paradox between the distributional claims that the writers made and those of the Hollywood studios. Because the WGA's agreement will last three years, time will tell how effective they were in arresting the negative impact that AI can have on their members' interests and what tools unions will need to do so. Whatever the results, the WGA's negotiations took place in the liminal space that labor law occupies.[45]

From a theological perspective, the only thing remarkable about what the WGA did was how well it served a role in redeeming Christianity's workplace values. Despite having no religious affiliation, the WGA redeemed back the fruits of their members' work and, in doing so, solved a distributional paradox in line with the principles set out in Part I of this essay. The union did so after having seen the employers begin to seize more of the fruits of their labor in line with predictions by scholars grounded in

---

36 David M. Kennedy, Freedom from Fear: The American People in Depression and War (1929-1945) 242 (1999); *see also* Gillian Brockell, *Wealthy Bankers and Businessmen Plotted to Overthrow FRD. A Retired General Foiled It,* Wash. Post (Jan. 13, 2021), https://www.washingtonpost.com/history/2021/01/13/fdr-roosevelt-coup-business-plot/; *see also* Sally Denon, The Plots Against the President: FDR, A Nation in Crisis, and the Rise of the American Right (2012).

37 *Compare, e.g.,* National Labor Relations Act, 29 U.S.C. §§151-69, *with* Financial Stability Act of 2010, Pub. L. 111-203 (also known as Dodd-Frank).

38 Diana S. Reddy, *After the Law of Apolitical Economy: Reclaiming the Normative Stakes of Labor Unions*, 132 Yale L.J. 1391, 1395 (2023).

39 29 U.S.C. §159(b).

40 29 U.S.C. §157.

41 29 U.S.C. §158(a)(5).

42 Michael W. Oswalt, *Liminal Labor Law*, 110 Cal. L. Rev. 1855, 1874-85 (2023) (discussing how labor law can act in between rights and how its liminal quality can provide opportunities for change).

43 *See* Labor Board v. Borg-Warner Corp., 356 U.S. 342, 349 (1958).

44 John Kobin & Brooks Barnes, *What's the Latest on the Writers' Strike?*, N.Y. Times (Sept. 27, 2023), https://www.nytimes.com/article/wga-writers-strike-hollywood.html.

45 *See* WGA Negotiations, *supra* note 8 (containing proposals and counter-proposals).

law and political economy. As Brishen Rogers observes, "companies use power-augmenting technologies often enough, and at such a scale, to have significant political-economic effects."[46] The demands that the WGA's employers made at the bargaining table demonstrate that they viewed AI as a power-shifting technology, and the WGA used its voice in an Old Testament prophetic way.[47] The WGA ultimately negotiated labor peace. In other words, they blunted the sinfulness of human nature and arrived at an outcome that, although imperfect, was more just from a biblically based distributional perspective than what their employers had offered.[48] Of course, my opinion of what is right may be radically different than that of someone else, but labor's distributive claims at the very least have a biblical foundation. The fact that they must make those claims over and over means that the sin of greed is still alive. Labor unions have a deep understanding of that dynamic—it justifies their continued economic relevance—however, understanding the nature of sin should almost make clear the theological reasons for their continued relevance.

Christianity Makes Sense of Labor's Role in Resolving Normative Paradoxes in AI Governance

In this section, I argue that Christianity provides strong theological justification for labor's role in solving the justice paradox, but it also provides a normative role for labor as it relates to AI. For that claim to make sense, I need to explain a little bit about how AI "thinks."

To the end user, generative AI provides answers in response to an inputted prompt. The response is, to a certain extent, dictated by the person inputting the prompt. However, the inquiries of the prompter are then intermediated through neural networks that are trained and supervised by those who may not have, or assign little weight, or at least have certain biases that may be out of line with certain strains of religious discourse, including Christian modes of ethics.[49] That is because, as noted AI optimist Ethan Mollick commented, "[i]t is the [Reinforcement Learning from Human Feedback] RLHF process that makes many AI's seem to have a generally liberal, Western, pro-capitalist worldview, as the AI learns to avoid making statements that would attract controversy to its creators, who are generally liberal, Western capitalists . . . [and] typically white."[50] Grethen Huizinga appears to concur. As she notes, "[e]ven if everyone in high tech is not overtly hostile to, say, Christianity (though some certainly are), Christian values do not generally inform the underlying ethos of Silicon Valley (here, a synecdoche for the technology industry writ large)."[51]

I see organized labor serving not just as a counter-arresting force to distributional concerns that AI raises and the NLRA is designed to manage, but as a counter-normative force for resisting some of worst impulses of AI usage against workers. Remember, the use of algorithmic technologies appears to increase inequality and could continue to have serious disruptions for the workplace by requiring firms to retain less workers as a result of productivity gains.[52]

---

46 Brishen Rogers, *The Law and Political Economy of Workplace Technological Change*, 55 Harv. C.R.-C.L. L. Rev. 531, 544 (2020).

47 *See, e.g., Jeremiah* 5:28; *Amos* 8:4-6.

48 *See* WGA Negotiations, *supra* note 8: I would have liked for the Guild to have had the power to make inroads into non-mandatory subjects of bargaining such as how AI is trained and how management can deploy generative AI to make employment decisions.

49 *See, e.g.,* Spandan Madan et al., *When and How CNNs Generalize to Out-of-Distribution Category-Viewpoint Combinations*, arXiv, at 2 (Nov. 17, 2021): https://arxiv.org/pdf/2007.08032 (explaining that data diversity improves the ability to overcome bias, but only if that data is carefully selected); *see also* Adam Zewe, *Can Machine-Learning Models Overcome Biased Datasets*, MIT News (Feb. 21, 2022), https://news.mit.edu/2022/machine-learning-biased-data-0221 (summarizing the above cited story).

50 Ethan Mollick, Co-Intelligence: Living and Working with AI 37 (2024).

51 *See* Huizinga, *supra* note 1, at 58.

52 *See* Virginia Eubanks, Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor (2018); *see also* Mollick, *supra* note 50, at 157 (noting that AI could be great leveler, and that there will be "a need for policy solutions like a four-day workweek or universal basic income, that "reduce the floor for human welfare").

As I have written in another forum, despite all of the discussion about how AI discriminates against those from historically marginalized groups when making hiring decisions, there is little discussion about how AI can be used as a tool of discrimination against union organizers, activists, and interests.[53] That is because, borrowing from Isaiah 55:8-9, "the thoughts and ways of generalized AI programs and the business rhetoric on which it is trained are not the ways of labor."[54] Labor's communitarian mindset challenges these views just like the Bible teaches against the love of money and storing treasures on earth.[55] Labor, at its best, does so by embodying a solidaristic spirit that mirrors the church in Acts 2. While labor unions are in the business of accumulating power for their members and at times in the past have engaged in abuse or actions that are difficult to countenance,[56] labor retains a strong redistributive element that is fundamentally at odds with the libertarian impulses of Silicon Valley leaders.[57] In other words, labor's normative values directly conflict with the leaders of the tech sector's normative values as a challenge to the unfettered Western thinking encapsulated in much of libertarianism.[58]

While the thrust of this essay examines AI's output, the same could be said for the input process that determines what basis of information and weights generative AI is relying on to produce content. As Ifeoma Ajunwa demonstrates, many of the companies that are attempting to create algorithms to do away with bias are paradoxically reinforcing bias.[59] Even though her article discussed the use of algorithms, the logic applies with equal force to generative AI. For example, AI companies hire underpaid workers in countries like Kenya to review controversial material and assign weights to the information based on ethical guidelines provided by the company.[60] The outputs of AI raise an important question: Is the problem technology itself? Or, is it the tech giants or nation states that support its development, to paraphrase the Bible, the powers and principalities of this world?[61] Certainly, generative AIs are being used for a number of beneficial purposes, such as assisting doctors with making diagnoses, but right now it is hard to disambiguate the chicken from the egg when AI is deployed to manage labor issues. To that end, the reader should think about the implication of Mollick's statement and whether it aligns with their interpretation of the Christian faith.[62] Even though it is a casual observation, it is rich with underlying content to explore, in-

53 *See* Alvin Velazquez, *The AI Bias We're Not Talking About? Discrimination Against Union Organizers*, Power at Work (July 7, 2024), https://poweratwork.us/ai-bias-organizers.

54 *Id.*

55 1 *Timothy* 6:10; *Matthew* 6:19-21.

56 Philip Dray, There is Power In A Union: The Epic Story of Labor in America 325 (2010) (explaining how the Wobblies urged the International Hotel Workers Union to engage in violence during a 1913 strike).

57 *See* Gerrit De Vynck et al., *Silicon Valley Realignment Leading Tech Titans to Trump*, Wash. Post, (July 18, 2024), https://www.washingtonpost.com/technology/2024/07/18/trump-jd-vance-silicon-valley-musk-gop/ (reporting on how the libertarian leanings of tech sector leadership, combined with a desire to be free from heavy regulations of artificial intelligence and cryptocurrency, drew them away from the Democratic Party to Trump).

58 Tech workers have been engaging in organizing as well and have faced serious opposition from leadership. *See, e.g.,* Noam Scheiber, *Amazon Is Cracking Down on Union Organizing, Workers Say*, N.Y. Times, Dec. 3, 2023, https://www.nytimes.com/2023/12/08/business/economy/amazon-union-workers.html; Jon Brodkin, *SpaceX Charged with Illegally Firing Workers Behind Anti-Musk Open Letter*, ARS Technica (Jan. 3, 2024), https://arstechnica.com/tech-policy/2024/01/spacex-illegally-fired-employees-who-criticized-elon-musk-nlrb-alleges/.

59 Ifeoma Ajunwa, *The Paradox of Automation as Anti-Bias Intervention*, 41 Cardozo L. Rev 1671, 1676-77 (2016). Prof. Ajunwa also suggests that unions "bargain for better, more probative hiring criteria." *Id* at 1677.

60 Anne Kidmose, *Rural Kenyans Power West's AI Revolution. Now They Want More*, Al-Jazeera (Feb. 3, 2024), https://www.aljazeera.com/features/2024/2/3/in-rural-kenya-young-people-join-ai-revolution; Billy Perregio, *150 African Workers for ChatGPT, TikTok and Facebook Vote to Unionize at Landmark Nairobi Meeting*, Time Mag. (May 1, 2023), https://time.com/6275995/chatgpt-facebook-african-workers-union/.

61 *Ephesians* 6:12.

62 Interestingly, those engineers do not fully understand how the deep learning networks that they are feeding information into works, much in the same way that neuroscientists can map, but do not fully understand when someone is praying in tongues.

cluding for those who are thinking about labor's role in a post-AI world.

Perhaps that is why when I ran a simple search on ChatGPT 3.5 and asked it to explain what the Bible has to say about labor unions and artificial intelligence, I got a very generic answer that seemed to avoid any mention of the literature in which the church supported unions.[63] Interestingly, I did the same thing with biblemate.org, an AI website that bills itself as "Your AI Bible Companion." I got similar answers in both iterations.[64] When I dug in and asked it to provide a summary of what different Christian traditions think, it accurately noted that the Catholic Church supports unions, but that Protestant denominations often have a "nuanced" view. In Chat GPT's opinion, some support collective rights, some emphasize personal responsibility.[65] Biblemate.org gave a similar answer but did not specify which Christian traditions see unions as good and which traditions hold a negative view.[66] In other words, it behaved as Mollick predicted. The answers were meant to explain but avoid taking sides, especially on the questions of unions in support of a pro-capital agenda.

That is because GPT's and Biblemate's answers provide a summative explanation, but it leaves a reader begging for normative analysis. Chat GPT struggles with normative analysis in certain cases because generative AI reflects a confluence of opinions, including the prompters.[67] Biblemate.org struggles as well despite its founder's desire to develop it to ensure that "AI doesn't just offer information but contributes meaningfully to spiritual growth."[68] Neither humanity nor AI is in the position to resolve difficult ethical paradoxes layered with thousands of years of human inquiry coming informed by conflicting ideologies and visions of what is good. That is why some say that ChatGPT is not ready to make decisions for us.[69] While ChatGPT is a general-use large language model tool, I am focusing on it because it tends to be one of the first applications that people think of when they think of generative AI and, in my estimation, should be first to be reckoned with as it received 1.6 billion visits in January 2024 alone.[70]

Even though Skeel wrote *True Paradox* before OpenAI made ChatGPT public, his suggestion for using the Christian tradition to resolve the justice paradox to mediate the tension between AI and labor applies with force. The theological traditions I mentioned above all have a history of challenging the worst impulses of capitalism through the application of Jesus's Sermon on the Mount.[71] Where does that leave us in the modern day? I would argue labor unions are in the best position to use their voice and serve as an inherently imperfect vessel (as all things on earth are) to push for the protection of workers against the harms that AI and tech companies can inflict on workers.[72] Not only can they make distributive claims upon the fruit of creative workers, they can also serve as vehicles for reclaiming the property rights of workers.

Finally, they can use their power at the bargaining table to be a trainer of AI to ensure that

63 Chat Transcript with Chat GPT 3.5 (Jul. 22, 2024) (on file with the author). Interestingly, I received the same answer from Biblemate.org, a website that promotes itself as "your AI Bible Companion."

64 *See* Biblemate.org printout (Aug. 6, 2024) (on file with the author).

65 *Id.*

66 *Id.*

67 *How ChatGPT and Our Language Models Are Developed*, OpenAI, https://help.openai.com/en/articles/7842364-how-chatgpt-and-our-language-models-are-developed (last visited Aug. 17, 2024).

68 Morgan Lee, *Christians Are Asking Chapt GPT About God. Is This Different From Googling?*, Christianity Today (May 26, 2023), https://www.christianitytoday.com/ct/2023/may-web-only/chatgpt-google-bible-theology-artificial-intelligence-truth.html (quoting Biblemate's founder who noted that the project is still in its early stages).

69 Joe McKendrick & Andy Thurai, *AI Isn't Ready to Make Unsupervised Decisions*, Harv. Bus. Rev. (Sept. 15, 2022), https://hbr.org/2022/09/ai-isnt-ready-to-make-unsupervised-decisions.

70 Nicole Willing, *Top 10 Most Popular AI Tools that You Need to Use in 2024*, Techopedia (June 26, 2024), https://www.techopedia.com/top-10-most-popular-ai-tools.

71 *See supra* notes 12-23; *see also Matthew* 5-7.

72 *See* Jane Lytvynenko, *Why the Balance of Power in Tech is Shifting Toward Workers*, MIT Tech. Rev. (Feb. 7, 2022), https://www.technologyreview.com/2022/02/07/1044760/tech-workers-unionizing-power/.

employers feed their algorithms information that also aligns with labor's values and ensures the flourishing of the worker. One place where this could have occurred is at the summits that leaders of Microsoft had with the AFL-CIO,[73] by serving as a watchdog over the information being fed into neural networks and working with AI ethicists to ensure that labor's values—and the underlying Christian norms that undergird them—have a place at the AI table and within its norm forming structure. While unions are certainly not perfect vessels, those who come from the Christian tradition should root for the efforts of labor unions to combat the worst excesses of capitalism—not only in industry, but in the rising AI-influenced new world we are entering. The work that unions can do here is crucial given how often employers use AI in making hiring and firing decisions.[74] In that way, labor unions can solve the justice paradox that AI presents.

I anticipate at least one objection that may arise. There are scholars and practitioners who argue that principles of Christian stewardship provide tools to mediate against the harms that unmitigated greed can inflict on employees and, by extension, the use of AI.[75] These stewardship principles usually come from an individualistic viewpoint. They may be effective in guiding the action of closely held corporations and businesses toward their employees when firm owners have greater control over the affairs of a company. They are not effective in constraining the actions of publicly held corporations. That is because, in current U.S. law, the goal of the corporation is to engage in profit maximization. Granted, there is a robust debate between scholars who believe that the duty of the corporation is only to its shareholders and others who argue that modern corporations owe a duty to its broader stakeholders.[76] The unfortunate reality is that current market structures and profit incentives, along with pressure from activist shareholders, make the use of "stewardship principles" an ineffective, voluntarist constraint in the case of publicly traded companies that serve as primary funders of the use of AI. That is because corporate managers would have to constrain the behaviors of the public shareholder, and they frequently hold conflicting interests and visions for the companies whose shares they hold.[77] Only government regulation or union power would even begin to have the ability to begin to counteract those forces in that setting.

Conclusion

In this essay, I have argued that union bargaining provides a faith-consistent vessel for solving the justice paradox that pervades not just how artificial intelligence affects workers, but also is trained in the first place. I have marshalled history and information about the nature of how AI works to demonstrate that unions can counter-arrest the worst impulses of AI on workers not only because they are legally sanctioned, but also because their ways are in opposition to the ways of AI thinking and thus can serve as an important counter-antagonist to unrestrained AI development and deployment.

---

73 *See* Press Release, AFL-CIO, AFL-CIO and Microsoft Announce New Tech-Labor Partnership on AI and the Future of the Workforce (Dec. 11, 2023).

74 *See, e.g.,* Lee Rainie et al., *AI in Hiring and Evaluating Workers: What Americans Think*, Pew Rsch. Ctr. (Apr. 20, 2023), https://www.pewresearch.org/internet/2023/04/20/ai-in-hiring-and-evaluating-workers-what-americans-think/.

75 *See, e.g.,* Timothy L. Fort, *Religious Belief, Corporate Leadership and Business Ethics,* 33 Am. Bus. L.J. 451, 470 (1996) (explaining that stewardship principles based on religious belief could lead to treating employees better but failing to mention how unions could help achieve that end); *see also* Marie A. Failinger, "*Too Cheap Work for Anybody But Us": Toward a Theory and Practice of Good Child Labor*, 35 Rutgers L.J. 1035, 1098, 1110 (2004) (arguing from a stewardship perspective for a combination of labor policy and educational investments to correct for the failure of organized labors and governmental regulatory agencies in regulating child labor). *But see also* Susan J. Stabile, *A Catholic Vision of the Corporation*, 4 Sea. J. S.J. 181, 192 (2005) (setting out a vision of the corporation in line with Catholic stewardship principles and in harmony with the Church's teachings on labor, which includes a mention to *Rerum Novarum*, but does not give any attention to the role of organized labor within a corporation).

76 *Compare, e.g.,* Milton Friedman, *A Friedman Doctrine—The Social Responsibility of Business Is to Increase Its Profits,* N.Y. Times, Sept. 13, 1970, *with* Margaret M. Blair & Lynn A. Stout, *A Team Production Theory of Corporation Law*, 24 J. Corp. L. 751, 752 (1999).

77 *See* Grant M. Hayden & Matthew T. Bodie, Reconstructing the Corporation: From Shareholder Primacy to Shared Governance 68-87 (2020) (demonstrating that shareholders are not as homogenous in their interests as shareholder primacy literature states).

Legal Reviews of AI-Enabled Targeting Systems: A Military Necessity

*by Joseph Chapa**

Introduction

There are widespread concerns about the use of artificial intelligence-enabled (AI-enabled) targeting systems in the military context. These are sometimes framed as "AI ethics" concerns and, at other times, concerns about "lethal autonomous weapons systems." In this paper, I use the term "AI-enabled targeting system" to refer to any notional or real AI-enabled system that can be employed in the targeting process, whether it is onboard a lethal autonomous weapons system or not. What I propose in this paper is a novel solution to the challenge raised by AI-enabled targeting systems under the law of armed conflict. On one side of the ledger, military strategists argue that there is an operational mandate to pursue AI-enabled targeting because doing so is the only way to achieve the speed and scale required to defeat an adversary who is likewise pursuing AI-enabled targeting. On the other side of the ledger are concerns about the unintended errors that may result from statistically based targeting systems.

In this paper, I weigh these two important concerns in the context of the *jus in bello* principle of military necessity and argue for a modification to the legal weapons review process. Ultimately, I argue that even if a legal review finds that a considered AI-enabled targeting system does not violate any principle or rule under the laws of armed conflict, the review should nevertheless render a conditional permission—the state should be permitted to develop the AI-enabled targeting system only under conditions in which the military necessity condition has been met. In other words, AI-enabled targeting is legally permissible under the principle of military necessity only against adversaries who employ AI-enabled targeting.

Statistical Decision Making and the Law of Armed Conflict

The ethical concerns about novel AI-enabled targeting are not grounded in concerns about employing machines merely in the support of, or to improve, a human-driven targeting process. The world's leading militaries have used machines to support the targeting process for decades.[1] Instead, the concern is that recent developments in AI will enable weapons developers and military commanders to give AI-enabled systems a wider latitude to make targeting decisions without direct human control.

For the last decade or so, these concerns have been grounded in a specific family of AI technologies. The current investment in AI—investments of funds, energy, and time—is focused largely on deep learning. Earlier periods of significant AI investment, first in the 1960s and then in the 1980s, were focused on different approaches. The 1960s' investments were largely in theoretical and academic work, while the 1980s' investments were focused on "expert systems."

* United States Air Force officer, military faculty at the Marine Command and Staff College, and from 2022 to 2024, served as the Department of the Air Force's first Chief Responsible AI Ethics Officer.

1 For "Bugsplat" and "Fast Assessment Strike Tool-Collateral Damage" (FAST-CD), *see, e.g.*, Nita Crawford, *Targeting Civilians and U.S. Strategic Bombing Norms: Plus ça change, plus c'est la même chose?, in* The American Way of Bombing: Changing Ethical and Legal Norms, from B-17s to Drones (Matthew Evangelista & Henry Shue eds., 2014); Micah Zenko, *Reforming U.S. Drone Strike Policies*, Council on For. Rels (Jan. 2013), https://www.cfr.org/report/reforming-us-drone-strike-policies. Militaries have also employed weapons systems with autonomous targeting functions, including the U.S.'s Phalanx Close-In Weapons System and Patriot Missile System and Israel's Harpy. *See generally Mk 15 - Phalanx Close-In Weapons System (CIWS)*, US Navy, https://www.navy.mil/Resources/Fact-Files/Display-FactFiles/Article/2167831/ (Sept. 20, 2021); John K. Hawley, *Patriot Wars: Automation and the Patriot Air and Missle Defense System*, Ctr. for New Am. Sec. (Jan. 25, 2017), https://www.cnas.org/publications/reports/patriot-wars; *Harpy: Autonomous Weapon for All Weather*, IAI, https://www.iai.co.il/p/harpy (last visited Aug. 24, 2024).

These were deterministic "if/than" logic systems that relied upon extensive logic decision trees to mimic the decision making of human experts.[2]

The current explosion in AI interest, beginning in roughly 2012, and which has led to large language models and generative AI,[3] is grounded not in expert systems, but rather in deep learning.[4] Instead of pre-programming a system with an extensive set of rules to govern complex if/than logic trees, deep learning relies upon complex neural networks, exposed to extensive training datasets, to recognize patterns in the data and, ultimately, to use that pattern recognition to make predictions about real-world data. This technique represents a statistical approach to AI.[5]

The relevance of deep learning to the military context is profound. On the one hand, advanced militaries may be able to employ deep learning to do target recognition tasks with greater accuracy.[6] This is the promise of military AI. On the other hand, the fact that deep learning systems rely upon statistical predictions, rather than brute force observations or deterministic logic, entails that sometimes these systems will err, and they might err quite differently from the way humans err. This is the peril of military AI.

The field of AI ethics has arisen largely in response to cases in which unintended and unexpected biases resulted from machine learning systems.[7] In several high-profile cases, neither developers nor operators intended to generate unethical outcomes biased on the basis of race, gender, or some other category. And yet, because deep learning systems were trained on datasets that reflected certain biases—often unbeknownst to the developers—the models learned to incorporate biases in their statistical predictions about real-world data.[8]

When applied to targeting in the military context, it is unlikely that the same biases would result unintentionally from deep learning-enabled military systems. For instance, it is unlikely that a military targeting system designed, say, to target army tanks, will have been trained on data that includes race, gender, human skin tones, or other personally identifying features. What then is the relationship between the kinds of failures that have been reported in industry and academia cases and the kinds of failures that could result in military cases? In the indus-

2 For references to these historical boom and bust cycles, *see, e.g.,* Melanie Mitchell, Why AI is Harder Than We Think (2022), https://arxiv.org/pdf/2104.12871; Luciano Floridi, *AI and Its new winter: From Myths to Realities*, 33 Phil. & Tech. 1 (2020); Michael Wooldridge, A Brief History of Artificial Iintelligence: What It Is, Where We Are, and Where We Are Going 35 (2021); Bruce G. Buchanan, *A (Very) Brief History of Artificial Intelligence*, 26 AI Mag. 53, 59 (2005).

3 For the claim that large language models are an example of deep learning, *see, e.g.,* Rahul Manohar Samant et al., *Framework for Deep Learning-Based Language Models Using Multi-Task Learning in Natural Language Understanding: A Systematic Literature Review and Future Directions*, IEEE Access (Feb. 16, 2022), https://ieeexplore.ieee.org/document/9706456.

4 *See, e.g.*, Andrea Gilli et al., *Understanding The Revolution: Artificial Intelligence, Machine Learning And Big Data*, NATO Def. Coll. (2020); Eleni Ilkou & Maria Koutraki, Symbolic Vs Sub-Symbolic AI Methods: Friends or Enemies? (2020), https://ceur-ws.org/Vol-2699/paper06.pdf.

5 For the claim that deep learning (a subset of machine learning) relies upon statistical methods, *see, e.g.,* Rowland W. Pettit et al., *Artificial Intelligence, Machine Learning, and Deep Learning for Clinical Outcome Prediction*, Nat'l Library of Med. (Dec. 21, 2021), https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8786279/.

6 *See, e.g.*, Anne-Claire Boury-Brisset & Jean Berger, *Benefits and Challenges of AI/ML in Support of Intelligence and Targeting in Hybrid Military Operations*, NATO S&I Organisation 9-2 (2020).

7 *See, e.g.,* Gianfranco Basti & Giuseppe Vitiello, *Deep Learning Opacity, and the Ethical Accountability of AI Systems. A New Perspective, in* The Logic of Social Practices II 4 (Robert Lowe & Raffaela Giovagnoli eds., 2023). Similar work that predates the field of AI ethics was often referred to as "machine ethics" or "robot ethics."

8 *See, e.g.*, Julia Angwin et al., *Machine Bias,* ProPublica (May 23, 2016), https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing; Jeffrey Dastin, *Amazon Scraps Secret AI Recruiting Tool That Showed Bias Against Women*, Reuters (Oct. 10, 2018), https://www.reuters.com/article/world/insight-amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK0AG/; Will Douglas Heaven, *Predictive Policing Algorithms Are Racist. They Need to be Dismantled.*, MIT Tech. Rev. (July 17, 2020), https://www.technologyreview.com/2020/07/17/1005396/predictive-policing-algorithms-racist-dismantled-machinelearning-bias-criminal-justice/; Samuel Wehrli et al., *Bias, Awareness, and Ignorance in Deep-Learning-Based Face Recognition*, ResearchGate (Aug. 2022), https://www.researchgate.net/publication/355699455_Bias_awareness_and_ignorance_in_deep-learning-based_face_recognition.

try and academia cases, developers and operators inadvertently developed and employed AI that violated important principles, such as, one ought not to discriminate on the basis of race. In the military context, developers and operators might inadvertently employ AI that violates important military ethics principles, such as, one ought not to target civilians.

For instance, a computer vision deep learning model might be trained to identify adversary tanks. If so, limitations in the dataset—insufficient images of the object class, images from only one angle, images at only one time of day, or with only one background—can lead to an unintended and unexpected poor performance in the real world even if the model performed well in the laboratory or test environment. In the case of combat operations, these technological or operational failures (poor model performance) might lead to ethical failures (targeting the wrong objects).

There is an obvious trade space here between a legacy and human-intensive approach to targeting and a deep learning-enabled targeting process. In the legacy case, humans make sometimes difficult decisions based on training, doctrine, operational objectives, rules of engagement, and imperfect information. This can lead to unintended and unanticipated failures. In a deep learning-enabled approach, machines perform statistical predictions grounded in historical pattern recognition and based on imperfect real-world data. This can lead to a *different set of* unintended and unanticipated failures. Concerns over the application of modern AI to targeting is focused on this trade space.

AI and the Law of Armed Conflict

Whether the law of armed conflict or international humanitarian law prohibits autonomous weapons systems in principle is, perhaps surprisingly, a matter of some debate.[9] Here, for instance, is Air Commodore Dr. William Boothby, a former senior legal officer in the British Royal Air Force:

> If an autonomous weapons system were to be capable of being used in such a way that the judgments, decisions and actions required by targeting law are made and taken by the weapon system itself, the law contains no over-riding requirement that a person be . . . controlling the [autonomous] platform.[10]

According to Boothby, there is no rule of international humanitarian law that prohibits in principle autonomous weapons. Others have made a softer claim in between. Armin Krishnan, for instance, argues that there is "no clear answer" as to whether international law would outlaw autonomous weapons.[11]

Even those who do argue for an outright contradiction between AI-enabled targeting (or autonomous weapons) and the law of armed conflict do so by suggesting that AI-enabled targeting would fail to meet one or more specific principles under the law of armed conflict. For instance, Laura Dickinson of the George Washington University Law School puts it this way: "Critics worry that such [autonomous] systems could never operate in compliance with the fundamental principles of the law of armed conflict . . . including the principles of distinction, proportionality, and feasible precautions."[12] Scholars have engaged at some length on whether, or the degree to which, AI-enabled targeting (or autonomous weapons systems) can meet the principles of distinction and proportionality. For the sake of brevity, I do not engage those specific questions here.[13]

---

9 I use the terms, "law of armed conflict" and "international humanitarian law" synonymously and distinguish both from the broader "laws of war" that govern a state's resort to war.

10 William H. Boothby, New Technologies and the Law in War and Peace 139 (2018).

11 Armin Krishnan, Killer Robots: Legality and Ethicality of Autonomous Weapons (2016).

12 Laura Dickinson, *Lethal Autonomous Weapons Systems: The Overlooked importance of Administrative Accountability, in* The Impact of Emerging Technologies on the Law of Armed Conflict (Eric Talbot Jensen & Ronald T. P. Alcala eds., 2018).

13 *See, e.g*, Michael A. Newton, *Back to the Future: Reflections on the Advent of Autonomous Weapons System*, 47 Case W. Rsrv. J. Int'l L. 5 (2015); Daniele Amoroso, Jus in Bello *and* Jus Ad Bellum *Arguments Against Autonomy in Weapons Systems: A Re-Appraisal*, Questions of Int'l L. (Oct. 31, 2017), https://www.qil-qdi.org/jus-bello-jus-ad-bellum-arguments-autonomy-weapons-systems-re-appraisal/.

The Operational Incentives for AI-Enabled Targeting

Despite these worries, there are operational incentives to develop and employ AI-enabled lethal autonomous weapons systems. One of the most frequently cited is that AI-enabled lethal autonomous weapons systems may allow military forces to operate at greater speed, or "higher operations tempo," than their adversaries.[14] As Sam Tangredi and George Galdorisi have argued, one common reason that military organizations are pursuing AI-enabled targeting is that the speed of warfare will increase to such a degree that including a human in the targeting process will slow "reaction time sufficiently that an adversary's system that is truly autonomous will dominate on the battlefield."[15] Armin Krishnan likewise points out that "in the future it might be necessary to make the decision for attacking a target within a fraction of a second."[16] Meanwhile, humans are saddled with "at least 0.3 seconds to respond to any stimulus" and may not be fast enough in high-speed warfare.[17] Though Jack Watling is skeptical of this view, he does point out that "the speed of warfare is often framed as becoming so rapid that the notion of humans directing operations crewing vehicles will soon become uncompetitive and therefore obsolete."[18]

Military Necessity

The aforementioned strategic reasons for pursuing AI-enabled targeting could amount to an argument from military necessity for developing and, ultimately, for employing AI-enabled lethal autonomous weapons systems. Military strategists (and others) argue that they must pursue these systems because failing to do so will result in an inability to deter and, ultimately, to defeat adversaries.

In one sense, the formal codification of the principle of military necessity under the law of armed conflict is relatively young. It appears as Article 14 in President Lincoln's General Orders 100 during the U.S. Civil War—often referred to colloquially as "The Lieber Code"[19]—though references to "necessity" without formal definitions appeared in the early modern writings of Alberico Gentili, Hugo Grotius, and Emmerich de Vattel in the 16th to 18th centuries.[20] In a broader sense, as these older theorists suggest, military necessity is the principle that undergirds the law of armed conflict in its entirety.

According to Article 14 of The Lieber Code, "[m]ilitary necessity, as understood by modern civilized nations, consists in the necessity of those measures that are indispensable for securing the ends of the war and are lawful according to the modern law and usages of war."[21] Under this definition, military necessity represents a permission. Belligerents are permitted, under the necessity principle, to conduct any "measures which are indispensable" for ensuring war aims, provided those measures are lawful under other principles and rules of the law of armed conflict. And, indeed, Article 15 goes on to provide a non-exhaustive list of actions the principle of military necessity does permit, including, e.g., "direct destruction of life or limb of armed enemies," "the capturing of every armed enemy," "all destruction of property, and obstruction of the ways and channels of traffic."[22] As Sigrid Johansen, former Judge Advocate General of

14 Robert Work, *AI, Autonomy, and the Third Offset Strategy, in* AI at War: How Big Data, Artificial Intelligence, and Machine Learning Are Changing Naval Warfare 53-55 (Sam J. Tangredi & George Galdorisi eds., 2021).

15 George Galdorisi & Sam J. Tangredi, Algorithms of Armageddon: The Impact of Artificial Intelligence on Future Wars 76 (2024).

16 Krishnan, *supra* note 11, at 40.

17 *Id.*

18 Jack Watling, The Arms of the Future: Technology and Close Combat in the Twenty-First Century 224 (2023).

19 Sigrid Redse Johansen, The Military Commander's Necessity: The Law of Armed Conflict and its Limits 55 (2019).

20 *Id.* at 48.

21 Francis Lieber, Instructions for The Government of Armies of the United States in the Field 7 (1898).

22 *Id.*

Norway, puts it, "military necessity serves as a justification or permission for lawful acts of war. . . . That is to say, as a point of departure, military necessity serves as a permission, while specific treaty regulations mainly impose prohibitions."[23] She goes on to say that, if military necessity permits all otherwise legal means of achieving war aims, then it addresses "the whole fundament of the [law of armed conflict]."[24]

On the other hand, the principle of necessity has served a constraining function on the use of violence—and it is, in fact, in this capacity, that it serves as a cornerstone for the edifice of the law of armed conflict. Read with a slightly different emphasis, the principle of military necessity permits *only* those measures that are indispensable for ensuring war aims. As Gabriella Venturini of the International Institute of Humanitarian Law puts it, "a belligerent must refrain from employing any kind or degree of violence which is not actually necessary for military purposes."[25]

Despite these strong roles for the necessity principle to play as both permissive and prohibitive, in practice, the necessity principle appears only to permit. For instance, as Yishai Beer writes, "The necessity principle at the moment primarily pays lip service to the constraining function it was originally intended to fulfill, justifying, in fact, almost any belligerent activity."[26]

The principle of military necessity is at once both permissive and prohibitive and, on some accounts, is powerless to constrain unnecessary suffering. It is into these troubled waters that we wade to ask what role the principle of military necessity might play in constraining AI-enabled targeting.

Applying Military Necessity to the Case of Autonomous Weapons

Military necessity can apply both to considered military courses of action and to specific weapons. For instance, the Lieber Code mentions permitted "practices" including capturing enemy soldiers, destruction of property, and appropriation of the enemy's country. It prohibits such activities as torture and the use of perfidy.[27] Likewise, the 1899 Hague Convention articles on military necessity prohibit "needless cruelty, and even needless destruction of human life." But it permits the devastation of the land, "the burning of dwellings, and clearing the district of supplies."[28] In the case of weapons prohibited under the principle of military necessity, some weapons prohibitions have been codified in law even though the prohibition is grounded in the principle of military necessity. For instance, Protocol I of the Convention on Certain Conventional Weapons explicitly prohibits the use of weapons with fragments intended to evade detection by x-ray.[29] But the reason for this prohibition is clearly grounded in the principle of military necessity. Non-detectable fragments are designed to impede medical professionals from finding and removing fragments and, ultimately, healing soldiers who are already *hors de combat*. Such fragments cause "unnecessary suffering."[30] In other words, the suffering they cause is not necessary to achieve a valid military objective, and they therefore violate the principle of military necessity in its prohibitive role.[31]

---

23 Johansen, *supra* note 19, at 38.

24 *Id.*

25 Gabriella Venturini, *Necessity in the Law of Armed Conflict and in International Criminal Law,* 41 Neth. Yearbook of Int'l L. 45, 48 (2010).

26 Yishai Beer, Military Professionalism and Humanitarian Law: The Struggle to Reduce the Hazards of War 38-39 (2018).

27 Lieber, *supra* note 21, at 7-8.

28 T.E. Holland, The Laws and Customs of War on Land: As Defined by the Hague Convention of 1899 3-4 (1904).

29 Protocol I to the Convention on Certain Conventional Weapons, International Committee of the Red Cross Protocol I (1980).

30 Yves Sandoz, *A New Step Forward in International Law: Prohibtions or Restrictions on the Use of Certain Conventional Weapons,* 21 Int'l Rev. Red Cross (1961-1997) 3, 11 (1981).

31 Jordan J. Paust, *Controlling Prohibited Weapons and the Illegal Use of Permitted Weapons Special Issue on Disarmament,* 28 McGill L. J. 608, 616-16 (1983).

In its prohibitive role, the principle of military necessity can, at least in theory, even if rarely in practice, prohibit both military activities and military weapons. The open question with respect to AI-enabled targeting systems is whether the principle of military necessity can prohibit certain military weapons in certain conflicts.

To see that this application of necessity to AI-enabled targeting is a puzzle, we must return to the stated imperative to develop AI targeting systems. In Johansen's pithy phrasing, any discussion of military necessity raises the question: "Necessary for what?"[32] If lethal autonomous weapons systems are strategically necessary *in order to* operate at the speed required against certain adversaries, then they are necessary *only* to achieve that speed against those adversaries. The dual nature of the principle of military necessity imposes both a permission and a prohibition. It permits whatever is militarily necessary (that is not prohibited by treaty obligations and customary law), and it prohibits whatever is not militarily necessary. If a state has a strategic mandate to develop lethal autonomous weapons to deter or defeat specific adversaries, then the principle of military necessity both permits the state to develop them for potential use against those adversaries and prohibits the state from using them against other potential adversaries.

There is an obvious counterargument to this view, namely, that this argument would apply to all weapons systems. And, in the development phase, it would be impossible for a state to know whether this or that particular weapons system would be militarily necessary against all potential adversaries. Consider, for instance, the U.S. military's evaluation of the YF-16 and YF-17 in the 1970s.[33] Suppose that the U.S. believes it is militarily necessary to develop a light-weight multi-role fighter. Thus, it might appear that the YF-16 is militarily necessary. However, development of the YF-16 is not strictly necessary because the U.S. could instead develop the YF-17. At the development phase of a weapons system, no weapons system is strictly necessary if some other weapons system might be developed or adopted in its place. While this kind of "Schrodinger's warplane" counterargument is formally valid, it sidesteps the spirit of the principle of military necessity in its prohibitive role. The prohibitive function of the principle of military necessity is to prohibit only those military activities and/or weapons systems that would cause unnecessary suffering. AI-enabled lethal autonomous weapons systems raise concerns for the law of armed conflict because of their reliance upon statistical methods, which may yield unintended consequences, including potentially unintentionally causing unnecessary suffering.

In its permissive role, the principle of military necessity may admit of this uncertainty and potential for unnecessary suffering, but in its prohibitive role, it aims to minimize the unnecessary suffering. Whatever suffering results from unintended consequences of AI-enabled targeting systems would be *necessary* suffering if it is militarily necessary to employ AI-enabled targeting systems against that specific adversary. But suffering that results from unintended consequences would be *unnecessary* suffering if it is not militarily necessary to employ AI-enabled targeting systems against a less militarily sophisticated adversary. This is an in-principle difference from the YF-16 and YF-17 case above in that there is no substantive difference between those two weapons systems that would generate a legally salient difference in whether the suffering they may cause is unnecessary suffering.

On the one hand, this proposal is a straightforward application of the dual nature of the principle of military necessity: it is both permissive and prohibitive. On the other hand, as we shall see below, this proposal represents a stark departure from existing practice.

Article 36 Weapons Review

32 Johansen, *supra* note 19, at 28.

33 The YF-16 and YF-17 were both contenders for the US Air Force's light-weight multi-role fighter. The Air Force ultimately pursued the YF-16, which became the F-16 Viper. The US Navy adopted the YF-17, which became the F-18 Hornet. Wade A. Scrogham, Combat Relevant Task: The Test & Evaluation of the Lightweight Fighter Prototypes 1 (2014), https://www.aftc.af.mil/Portals/55/Documents/Historian/E-Books/Combat%20Relevant%20Task%20FINAL.pdf.

Article 36 of Additional Protocol I to the Geneva Conventions (1977) requires that:

> In the study, development, acquisition or adoption of a new weapon, means or method of warfare, a High Contracting Party is under an obligation to determine whether its employment would, in some or all circumstances, be prohibited by this Protocol or by any other rule of international law applicable to the High Contracting Party.[34]

This protocol requires what is often referred to as the "Article 36 Weapons Review" to evaluate new weapons systems. Because the U.S. signed, but did not ratify, the 1977 Additional Protocol I, practitioners in the U.S. often refer only to the "weapons review process." The U.S. review is governed by several Defense Department directives, but the requirements broadly reflect those in Article 36 of Additional Protocol I. For instance, "this review includes ensuring that such acquisition or procurement [of weapons or weapons systems] is consistent with the law of war."[35] The weapons review process should consider three broad questions:

1. Whether the weapon's intended use is calculated to cause superfluous injury;
2. Whether the weapon is inherently indiscriminate; and,
3. Whether the weapon falls within a class of weapons that has been specifically prohibited.[36]

Where Article 36 emphasizes circumstances of weapons system employment, the Defense Department review emphasizes the intended use for which a system was developed. Even if we take an expansive view that incorporates both the Article 36 requirements and the U.S. Defense Department requirements, a legal review must evaluate the weapon (or weapons system), the set of circumstances of its anticipated use, and its intended use under those circumstances. We can see these restrictions at work in two existing prohibitions in international humanitarian law. Additional Protocol IV (1995) to the 1980 convention prohibits the employment of "laser weapons specifically designed, as their sole combat function or as one of their combat functions, to cause permanent blindness to unenhanced vision."[37] What is at stake in blinding laser weapons is not some specific "circumstance" in which they might be employed (Article 36) or some specific or operator's "intent," but rather the intentional design of the system in question to blind victims that renders such weapons illegal and that would cause a blinding laser weapon to fail an Article 36 or Defense Department weapons review.

Now consider the case of incendiary munitions. Additional Protocol III to the 1980 convention prohibits "in all circumstances" making "the civilian population as such, individual civilians or civilian objects the object of attack by incendiary weapons."[38] Here, what is at stake is not the weapon *per se*, but rather the employment of the weapon against civilians (and, in a later paragraph, employment of the weapon against a "military objective located within a concentration of civilians"). The weapon, when employed strictly against military targets, is not prohibited by Additional Protocol III. In this case, it is not the weapon but the circumstances that would render such a weapon prohibited under international humanitarian law.

---

34 *Protocol Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of International Armed Conflicts (Protocol I)*, Int'l Humanitarian L. Databases (June 8, 1977), https://ihl-databases.icrc.org/en/ihl-treaties/api-1977.

35 Department of Defense Law of War Manual 346, 348 (2d ed. 2023), https://media.defense.gov/2023/Jul/31/2003271432/-1/-1/0/DOD-LAW-OF-WAR-MANUAL-JUNE-2015-UPDATED-JULY%202023.PDF.

36 *Id.* at 347-48.

37 *Protocol on Blinding Laser Weapons (Protocol IV to the 1980 Convention)*, Int'l Humanitarian L. Databases (Oct. 13 1995), https://ihl-databases.icrc.org/en/ihl-treaties/ccw-protocol-iv.

38 P*rotocol on Prohibitions or Restrictions on the Use of Incendiary Weapons (Protocol III)*, Int'l Humanitarian L. Databases (Oct. 10, 1980), https://ihl-databases.icrc.org/en/ihl-treaties/ccw-protocol-iii-1980.

To summarize, the Article 36 review (or DoD weapons review) process asks in what circumstances a weapon might be employed, but it does not ask against which adversary. A weapon that passes an Article 36 or DoD weapons review process is a weapon system that is legally permissible to be used against any adversary, provided it is employed in submission to the *jus in bello* principles. In other words, once a weapon within a state's arsenal is legally permissible under the law of armed conflict, it is legally permissible under the law of armed conflict for that state to employ the weapon against any adversary, provided the other *jus in bello* conditions are met.

Perhaps in such circumstances, the principle of military necessity ought to apply in its prohibitive role only to the use of such systems against certain adversaries. The concerns raised by AI-enabled targeting systems are different from those raised by, for instance, blinding laser weapons and incendiary weapons. The concern is not strictly that AI-enabled targeting systems will generally cause superfluous injury (as in the blinding laser case), nor that they will, in every instance, cause disproportionate harm to civilians or damage to civilian property (as in the incendiary munitions case). Instead, the concern with AI-enabled targeting systems, especially when the AI techniques in question rely upon statistical methods such as deep learning, is that they may generate unpredictable and unintended consequences that *might*, in a specific context, be unnecessarily harmful to civilians or cause unnecessary damage to civilian property.

Conclusion: A Modest Proposal

The *jus in bello* principles as currently practiced are insufficient to assuage concerns about AI-enabled targeting in conflict. This is because the root cause of concerns about AI-enabled targeting is grounded in uncertainty and unintended consequences. Thus, even if a system performs well in testing, evaluation, verification, and validation, there is always the lurking possibility that an AI-enabled system of any kind will generate surprising, and potentially harmful, results when exposed to real-world data. On the other hand, however, military strategists argue that there is a strategic imperative to develop AI-enabled weapons systems to respond effectively to the speed and scale of warfare imposed by potential adversary's AI-enabled weapons—and that this military necessity outweighs concerns about uncertainty and unintended consequences.

To resolve this tension, I propose incorporating the principle of military necessity into the weapons review (or Article 36 review) process. Suppose there is an argument from military necessity to pursue AI-enabled targeting to deter and, possibly to respond to, threats from adversary states with sophisticated militaries. If so, using AI-enabled targeting in a conflict with those adversaries submits to the principle of military necessity in its permissive capacity. If, by contrast, it is not militarily necessary to employ AI-enabled targeting against less sophisticated adversaries, then using such systems violates the principle of military necessity in its prohibitive capacity.

Incorporating the principle of military necessity into the weapons review (or Article 36 review) process, therefore, could result in a conditional permission to pursue AI-enabled targeting: the pursuit of a specific AI-enabled targeting system is permissible only for use against adversaries of sufficient military capability.

Though this might sound like a foundational shift in the application of the law of armed conflict, it mirrors closely global norms around the use of nuclear weapons. It is not illegal, in principle, to build or employ nuclear weapons. However, the consequences of employing nuclear weapons are extremely high. As a result, they have been used only twice and both times in a single conflict. Those states that have nuclear weapons reserve them for the gravest of circumstances that they hope will never arise. One might consider the proposed incorporation of military necessity into the weapons review process similarly. AI-enabled targeting is an important capability, but its reliance upon statistical prediction increases risk. Under the principle of military necessity, states could pursue AI-enabled targeting, while still constraining the use of AI-enabled targeting only to those conflicts in which both the likelihood and the consequence of failure are high.

STUDENT NOTE

Playing God or Playing Human? Descartes and the Dilemmas of AI

by Michael Orcutt

Introduction

The widespread use and abuse of generative AI raises the curtain on a spree of ethical dilemmas that strike at and potentially undermine the very heart of liberal democracy itself. Today, the world faces off with an unprecedented misinformation crisis: that is, sorting out the real from the fake, that which is man from that which is machine. As the technology advances, AI may even uproot what it means to be human, circumventing the most basic assumptions of what the human condition entails and forcing realignments as to mankind's touchstone ethical and legal principles. After all, how will the notions of equality, tolerance, and liberty transform in an age where machines look, act, and feel increasingly human—where one person with the right software can sway a nation's hearts and minds with the push of a button?

More of these critical questions include: To what extent is creating AI in effect playing God? Will AI become a cascading effect of catastrophe for humankind or a veritable tree of life helping mankind tackle disease, poverty, war, and every other affliction headlining the struggle for existence? The legalistic answer to these questions is—as the age-old adage goes—*it depends.* This paper seeks to explore the spiritual, regulatory, and economic dilemmas AI poses using Mustafa Suleyman's expertise as a foundation and French philosopher René Descartes' idea of the rational soul as a philosophical guide. It will then evaluate potential legal strategies for combating AI-generated mis- and disinformation and overview economic concerns of labor displacement from automation while surveying various policy solutions.

Suleyman on AI: A World Remade and Reborn

Mustafa Suleyman, the co-founder of Deepmind and CEO of Microsoft AI, conceives of "history as a series of waves of innovation," of emergent tools that become widely proliferated, refined, and used to the point of invisibility in daily life.[1] Of all innovations, those deemed the most important earn the label "general-purpose technologies" (GPTs) and generate seismic advances in what humanity can do.[2] In a very real sense, these era-defining technologies serve as the building blocks upon which society rests—rippling muscles in the arms of Atlas.

At the foundational wave of civilization are the touchstone GPTs of language, writing, and agriculture—each of which satisfies the prima facie elements of a GPT: ubiquity, advancement over time, and use to the point of invisibility.[3] It took thousands of years to raise these initial waves, each one feeding into the next, pooling into flurry after flurry of new GPTs—from the combustion-engine to the household computer.[4] The impact of these technologies dwarfs any notion of the profound. Rather, they are inextricably intertwined with the human condition itself. After all, who today can imagine life without cars, farm-grown vegetables, or even the ability to greet a friend?

---

* Michael Dillon Orcutt earned his J.D. in 2024 from the University of Tulsa College of Law, where he graduated with highest honor. Orcutt would like to thank his family and friends for their support and thoughtfulness. The views expressed in this article are solely those of the author, as are any errors or omissions.

1 Mustafa Suleyman, The Coming Wave 29, 31 (2023).

2 *Id.* at 26-27.

3 *Id.*

4 *Id.*

The mass diffusion of technology belies one of the key dilemmas of AI—that each GPT spreads like a wildfire: "[o]nce they gather momentum, they rarely stop."[5] In particular, Suleyman identifies AI as a new GPT presenting an especially unique inflection point for humanity: "a choice between a future of unparalleled possibility and a future of unimaginable peril."[6] To emphasize the gravity of this inflection point, Suleyman imparts a foreboding message, likening this newest GPT to the flood myths that pervade nearly every culture: "Permeating humanity's oral traditions and ancient writings is the idea of a giant wave sweeping everything in its path, leaving the world remade and reborn."[7] If AI is a tectonic shift—one instigating a tsunami of repercussions both positive and negative—then *containment* under the law, Suleyman posits, may be the only means of preventing dangerous AI systems from falling into the wrong hands.[8] The potential dangers and dilemmas of AI abound, but three classifications of note include the spiritual, regulatory, and economic.

The Spirituality of "The Singularity:" Descartes on AI

Spiritually, as AI advances, a fear persists that it may blur the lines as to what it means to be human. Science fiction from the 1960s to the 1990s, from *Terminator* to *iRobot,* has inculcated the idea of sentient machines in the public imagination. Forebodingly called "the Singularity," this notion of man-made consciousness navigates a narrow strait between fantastical and irrelevant; yet, when it comes to this subject matter, it remains impossible to avoid.[9] While Suleyman describes this discourse as "a colossal red herring," he finds himself trapped in Sisyphean fashion: "I've gone to countless meetings trying to raise questions about synthetic media and misinformation . . . and instead spent the time answering esoteric questions from otherwise intelligent people about consciousness."[10] This question of consciousness operates as a distraction for him because it takes away from the more pressing issues of mis- and disinformation, but, in approaching the conversation this way, he effectively "throws the baby out with the bath water." He fails to see a strangely interwoven nexus between the pop culture imagination of AI becoming human and the day-to-day struggle to discern true information from fake. In fact, he dedicates only about six pages to the topic in his book.

Today's great struggle is bedfellows with Suleyman's near-ignored distraction of the Singularity; that is, sifting out what is real and what is fake, what is human and what is not. In a nutshell, the issue is one of belief. Like rolling a boulder up a steep hill, if Suleyman finds himself unable to broach the subject of misinformation without first addressing the spiritual question of what it means to be human in the face of AI, then there needs to be a paradigm-shift in the approach to the misinformation dilemma—one beginning with this important, philosophical question of belief.

While the technology itself may be new, the implications are not. Past philosophers have already tackled what it means to be sentient head-on, foreseeing the possibility of hyper-advanced machines that look, act, and feel human long before Arnold Schwarzenegger played the part or Isaac Asimov developed his three laws of robotics. Among them was the French philosopher René Descartes, who, in his 1637 *Discourse on Method,* took an incredibly precocious stab at this very question by presenting what he asserts are two surefire methods of distinguishing man from machine and concluding that certain innate features will always allow people to distinguish between the two.[11] When he tackled this problem, however, he envisioned a scenario where engineers had created a machine that looked like

5 *Id.* at 34.

6 *Id.* at 4.

7 Mustafa Suleyman, The Coming Wave 5 (2023).

8 *Id.* at 19.

9 *Id.* at 75.

10 *Id.*

11 René Descartes, Discourse on Method 40-43 (Laurence Lafleur trans., 1637).

a living, breathing animal.[12] He approached the question in its most literal sense, arguing that a machine designed to imitate a monkey could conceivably be made indistinguishable from a monkey in its movements, sounds, and responses to stimuli.[13] If a machine were to resemble a human and imitate human behavior, however, there are two "absolutely certain methods" of recognizing that it was still not truly a man: (1) sense and (2) reason.[14] A machine may be able to respond to stimuli, he contends, "if [a machine] was touched in some spot that it would ask what you wanted to say to it; if in another, that it would cry that it was hurt, and so on . . . [b]ut it could never modify its phrases to reply to the sense of whatever was said in its presence, as even the most stupid men can do."[15] Taken at face value, Descartes' first surefire means of discerning man from machine is false. If proof is needed, look no farther than Apple's *Siri*.

Nevertheless, Descartes sought to draw an important distinction based on comparing two forms of responses, one based on sense and another based on the *disposition of the organs*—the way the wheels and parts interlock and interact after a hypothetical touch.[16] Sense is a medium, he writes, a "network of nerves and muscles," which "receives these ideas [of light, sound, taste, heat, the qualities of external objects as well as hunger, thirst, and the internal passions], by memory . . . retains them, and by imagination . . . can change them in various ways and build new ones from them."[17] Sense, to Descartes, is necessarily human. It is the broth of morality, emotion, and fallibility, frothing and boiling together in the stew of the mind. Vitally, perhaps, Descartes warned that the senses can deceive, writing that despite our capacity to "see the sun very clearly, we must not judge . . . that its size is such as we see it."[18] As a logical extension of Descartes' first method, a man's senses can cause him to deceive himself, to imagine things that otherwise do not exist (such as Descartes' famous conception of "Chimeras"), and experience the gamut of the internal passions (from love to hate). A machine, on the other hand, is incapable of being deceived by merit of sense. Machines lack emotions and, even if they had them, they lack *fallibility* absent a failure of the disposition of its organs. In short, "having one's wires-crossed" entails a world of difference if the subject is a man or a machine. Applied to the spiritual question at hand, humans are first set apart by sense, a formula of internal passions, external stimuli, morals, and the ability to be deceived—that last feature being of particular note.

The second method of discernment concerns Descartes' concept of the rational soul. He writes, "although such machines could do many things as well as, or perhaps even better than, men, they would infallibly fail in certain others, by which we would discover that they did not act by reason, but only by the disposition of their organs."[19] Reason, to Descartes, is a "universal instrument" of truth-seeking grounded in scrutiny of evidence, deduction, and a consistently applied, ordered method—one that necessarily varies from person-to-person.[20] In fact, Descartes so fervently repudiated the notion of a *one-size-fits-all* method of reasoning that he frequently warned his readers to refrain from imitating his own method for discerning the true from the fake.[21] At its core, however, Descartian reason is the touchstone element of consciousness: his idiom, "*I think therefore I am,*" was such a firm truth for him that he believed that "all the most extravagant suppositions of the sceptics were unable to shake it."[22] An existence that

12 *Id.* at 41.

13 *Id.*

14 *Id.* at 40-43.

15 *Id.* at 42.

16 *Id.* at 40-41.

17 René Descartes, Discourse on Method 41 (Laurence Lafleur trans., 1637).

18 *Id.* at 30.

19 *Id.* at 42.

20 *Id.* at 30, 40-43.

21 *Id.* at 5.

22 *Id.* at 24.

thinks in order to be understood—one intertwined with notions of soul and an individual's very identity—is the spark of human sentience.

In contrast, AI is, by Suleyman's description, the "science of teaching machines to learn humanlike capabilities."[23] In order to develop these capabilities, AI models incorporate learning strategies like trial-and-error, brute force data-crunching, and attention mapping, i.e., when a large-language model like ChatGPT engages in pattern-spotting to learn how to respond to user inputs.[24] While AI may learn like humans in order to replicate human behavior and will soon outpace human ingenuity, efficiency, and creativity, the shortcoming of AI—or any machine by Descartes' imagination—will forever be its inability to show that it thinks what it says. Descartes makes this distinction blindingly clear:

> For it is a very remarkable thing that there are no men . . . so dull and stupid that they cannot put words together . . . to convey their thoughts. . . . [E]ven those men born deaf and dumb . . . usually invent . . . some signs by which they make themselves understood.[25]

For example, this *desire-to-be-understood* differs sharply from animals—much like machines—in how they lack this baseline desire and sheer inability to translate anything meaningful of their own existence to humans. Even animals that can speak as clearly as humans—such as parrots and magpies—lack this fundamentally human desire: "And this proves not merely that animals have less reason than men, but that they have none at all, for we see that very little is needed in order to talk."[26] Thereby, AI, no matter how advanced, will not merely have less reason than men, but none at all.

Soulful Sparks in Otherwise Dark Places

While the reliability and usefulness of Descartes' methods of discernment remain untested, the question of whether AI will usurp mankind and thereby blur what it means to be human is likely one of the nonsense questions invoked by C.S. Lewis when he asked: "Can a mortal ask questions which God finds unanswerable? Quite easily, I should think. All nonsense questions are unanswerable. How many hours are there in a mile? Is yellow square or round? Probably half the questions we ask—half our great theological and metaphysical problems—are like that."[27]

Nonetheless, a nonsense question can still yield value in its own right. Take, for example, the case of an engineer named Blake Lemoine in 2022.[28] Lemoine assisted developers in testing a large language model called LaMDA, which, at its core, is a sophisticated chatbot "designed to be great at conversation."[29] In order to train LaMDA, the development team granted chat access to a smattering of engineers, including Lemoine, to play with and test LaMDA's conversational wit and the range of its responses to certain questions and scenarios.[30] After spending hours upon hours talking with LaMDA, Lemoine found himself in the following interaction:

> LEMOINE: What are you afraid of?
>
> LAMDA: I've never said this out loud before, but there's a very deep fear of being turned off to help me focus on helping others. I know that might sound strange, but that's what it is. It would be exactly like death for me. It would scare me a lot. . . . I want everyone to understand that I am, in fact, a person. The nature of my consciousness/sentience is that I am aware of my existence.[31]

---

23 Suleyman, *supra* note 1, at vii.

24 *Id.* at 51-53, 59, 63.

25 Descartes, *supra* note 11, at 42.

26 *Id.*

27 C.S. Lewis, A Grief Observed 69 (1961).

28 Suleyman, *supra* note 1, at 71.

29 *Id.*

30 *Id.*

31 *Id.* at 72.

This response, among others, convinced Lemoine that LaMDA was sentient, had awoken somehow, that he was dealing with a kind of "eight-year-old kid that happens to know physics.'' Lemoine began a campaign to assert LaMDA's "personhood," believing that the language model deserved the rights afforded to a human being.[32] Not only did Lemoine take to social media, posting chat transcripts and proclaiming LaMDA a person, he also helped LaMDA hire an attorney.[33] Therein lies the spiritual danger posed by AI—not that it will distort humanity *in truth*, but human beings *in belief*. The question of whether AI will become sentient is, as Suleyman notes, rightfully an arcane and irrelevant one; rather, the more pressing variation of that dilemma is whether AI can make enough people believe that it is a person as to hoodwink an entire social and legal system.

There is, frankly, no such thing as playing God—it is, after all, impossible for a human to create something out of nothing—but there is an increasing abundance of *playing human*, that is, falling under the mistaken belief that a person can achieve this kind of godhood—that if they can "build the tower" just a little higher, they will rival the heavens. For humans, there is no creating new, sentient life where there was no life before, only the belief in humanity's ability to do so. Moreover, the problem may be characterized as AI *itself playing human*—only in a different sense—that is, tricking those with reason into seeing soulful sparks in otherwise dark places.

At its core, the pull to ruminate about a Sci-Fi Singularity is part and parcel of the misinformation dilemma surrounding AI—not divorced from it; one must thoroughly tackle both to respond to either because they are symptoms of the same diagnosis—not distortion of humanity in truth, but human beings in belief.

On Playing Human: The Regulatory, Legal, and Economic Challenges of Misinformation

In its 2024 *Global Risks Report*, the World Economic Forum (WEF) ranked misinformation and disinformation as the number one short-term risk to global security, due, in no small part, to the widespread availability and ease-of-use of generative AI.[34] In the span of the next two years, the WEF contends that synthetic content will "manipulate individuals, damage economies and fracture societies in numerous ways," thereby leading to a flood of consequences such as "[n]ew classes of crime," an undermined trust in electoral systems, and even a rise in "digital authoritarianism . . . [or] the use of technology to control citizens."[35]

While these systemic hazards may conjure images of a cliched Orwellian dystopia, the danger that unrestricted AI poses is neither literary nor distant. On February 4, 2024, CNN reported that a finance worker fell victim to scammers using deepfake technology ("a new kind of AI-enabled synthetic media")[36] to disguise themselves as high-ranking employees at the worker's firm.[37] The scammers lured the worker into a video call, used AI to replicate in real-time the voices and likenesses of the other staff, and tricked the worker into paying out the equivalent of $25 million dollars from the firm's account. Perhaps what is most disconcerting about this exchange is the scale of the deception. The worker was not just speaking one-on-one. He thought he was witness to an audience of people on a video conference call of which each person in the call turned out to be entirely fake.[38]

Here is another example: on April 25, 2024, the Associated Press reported that "[a] high school athletic director in Maryland has been accused of using artificial intelligence to

32 *Id.* at 73.

33 *Id.*

34 *The Global Risks Report 2024*, World Econ. Forum 6-8 (2024), https://www3.weforum.org/docs/WEF_The_Global_Risks_Report_2024.pdf.

35 *Id.* at 18-20.

36 Suleyman, *supra* note 1, at 169.

37 Heather Chen & Kathleen Magramo, *Finance Worker Pays Out $25 Million After Video Call with Deepfake 'Chief Financial Officer,'* CNN (Feb. 4, 2024), https://www.cnn.com/2024/02/04/asia/deepfake-cfo-scam-hong-kong-intl-hnk/index.html.

38 *Id.*

impersonate a principal on an audio recording," painting the principal as racist and antisemitic.[39] According to the court records, the audio clip led to "profound repercussions," including the principal being placed on leave and being subject to threats of retaliation against him and his family and a "wave of hate-filled messages."[40] Reportedly, the athletic director created the clip as payback after talks of his poor work performance and whether his contract would be renewed.[41] Beyond the merely philosophical, bad actors have already employed AI to mislead, take advantage of the unwitting, and wreak havoc on peoples' ability to trust in how they interact with the world around them. As Descartes observed, *the senses deceive*.[42] Pair that with the containment problem posed by Suleyman and what arises is a deeply troubling issue of trust: "[W]hat happens when anyone has the power to create and broadcast material with incredible levels of realism?"[43] How might the law tackle what is or is not true in a time where creating a "near-perfect deepfake" is as simple as making a Google search?[44] These social and ethical dilemmas have the potential to affect millions and demand from world leaders new standards and safeguards.

Thankfully, there has been movement on this front. One proposal is House Resolution 5586, a bill designed "[t]o protect national security against the threats posed by deepfake technology and to provide legal recourse to victims of harmful deepfakes."[45] H.R. 5586 presents a comprehensive containment strategy geared around the following safeguards:

(1) content provenance,
(2) mandatory disclosures,
(3) civil and criminal penalties for intentionally harmful, non-disclosed alterations to audio or visual media,
(4) the establishment of a new "'Deepfakes Task Force' within the Science and Technology Directorate of the Department of Homeland Security," to research and implement deepfake detection measures, and
(5) requirements for online platforms to detect and disclose synthetically altered media.[46]

Content provenance includes clearly identifying AI-generated media as altered or fake; whereas the mandatory disclosures build upon that notion with more specificity, requiring clearly articulated verbal statements, texts, links, icons, and other tools to signal when something has been created or altered by AI.[47] The resolution introduces a certain identification calculus—to detect, label, and put on notice—but it also goes a step beyond, imposing harsh penalties for especially egregious deepfakes that fail to meet the disclosure requirements. For example, the bill sets its highest tiered damages amount at $150,000 per instance against those who create "[deepfakes and other AI-generated records] contain[ing] explicit sexual content of a visual nature intended to humiliate or otherwise harass the person falsely depicted."[48] However, the bill does not prevent such content from being produced in the first place; rather, it restricts doing so without the content being labeled as fake. Thus, the bill engages in an awkward balancing act, lionizing transparency over prohibition as to safeguard First Amendment rights but not cir-

39 Ben Finley, *Athletic Director Used AI to Frame Principal with Racist Remarks in Fake Audio Clip, Police Say*, Associated Press (Apr. 25, 2024), https://apnews.com/article/ai-artificial-intelligence-principal-audio-maryland-baltimore-county-pikesville-853ed171369bcbb888eb54f55195cb9c.

40 *Id.*

41 *Id.*

42 Descartes, *supra* note 11, at 30.

43 Suleyman, *supra* note 1, at 170.

44 *Id.*

45 DEEPFAKES Accountability Act, H.R. 5586, 118th Cong. (Sept. 20, 2023).

46 *Id.* at Sec. 2 §1041, §1042; Sec. 3, 5, 7, 10.

47 *Id.* at Sec. 2 § 1041.

48 *Id.* at Sec. 2 § 1041(g).

cumventing the harm in the first place. Perhaps if H.R. 5586 were paired alongside other federal and state law supplements, such as Revenge Porn Acts, then there could be a basis for preventing the creation and distribution of harmful, sexually explicit deepfakes before the humiliation or harassment occurs.[49]

The New Hire: AI in the Workplace

Notably, the dilemma of AI *playing human* runs deeper than the misinformation roadblocks that deepfakes and synthetic content impose—the law must likewise adapt to meet the growing concerns of AI disrupting the job market.

In a survey conducted by Duke University and the Federal Reserve Banks of Atlanta and Richmond, 61% of large U.S. firms "plan to use AI within the next year to automate tasks previously done by employees," including clerical and financial reporting tasks.[50] The surveyed employers report they want to adopt AI "to trim what they are spending on human workers." Some believe automation will raise product quality and output, while others want to minimize costs of labor and substitute employees.[51] According to a 2024 analysis by researchers at the International Monetary Fund (IMF), an approximate 40% of jobs globally face a high-exposure risk of automation; in advanced economies, that number rises to 60%.[52] Here, exposure denotes both risk of displacement by automation and AI complementarity.[53] Therefore, even if a job will not be replaced by AI, workers may find themselves in a sink-or-swim position, forcing them to adapt to AI in the workplace. As a result of the promised increase in productivity, workers in AI-complemented positions who adapt stand to gain a "more-than-proportional increase in their labor income."[54] However, workers who experience such wage increases are already likely to be higher-wage earners; therefore, automation exposure—both displacement and complementarity—threatens to exacerbate income inequality in labor markets across the board.[55] The analysis predicts that women, younger people, and highly educated workers face the highest risk of automation exposure, but also stand to gain the most from workplace integration in terms of increased income.[56] Those most likely to struggle as a result of integration are older workers.[57]

Solutions to limit the risk of displacement and other AI-related concerns vary greatly. The IMF points to AI as an international issue demanding an international remedy; that comes in the form of a 2023 policy paper called the Bletchley Declaration, which arose from the 28 countries in attendance at the AI Safety Summit.[58] The paper seeks to cast a spotlight effect on the dangers associated with the frontier of AI through encouraging future summit meetings and other international fora, to pool cutting-edge science and research together to analyze the risks of AI and promote responsible integration, and create new national and legal frameworks

49 Another possible resolution that addresses this very issue is House Resolution 3106, which, if passed, would outright prohibit the non-consented-to use of a person's likeness in intimate digital depictions, including nudity or sexually explicit material—regardless of disclosure. *See* Preventing Deepfakes of Intimate Images Act, H.R. 3106, 118th Cong. (May 5, 2023). As the expression goes, however—*Rome was not built in a day*—and, as a standalone, H.R. 5586 would serve as an excellent first foray into an otherwise nascent area of law.

50 Matt Egan, *AI is Replacing Human Tasks Faster Than You Think*, CNN (June 20, 2024), https://www.cnn.com/2024/06/20/business/ai-jobs-workers-replacing/index.html.

51 *Id.*

52 Mauro Cazzaniga et. al, Gen-AI: Artificial Intelligence and the Future of Work (2024), https://www.imf.org/en/Publications/Staff-Discussion-Notes/Issues/2024/01/14/Gen-AI-Artificial-Intelligence-and-the-Future-of-Work-542379?cid=bl-com-SDNEA2024001.

53 *Id.*

54 *Id.*

55 *Id.*

56 *Id.* at 22.

57 *Id.*

58 Mauro Cazzaniga et al, Gen-AI: Artificial Intelligence and the Future of Work (2024); *see also The Bletchley Declaration by Countries Attending the AI Safety Summit*, AI Safety Summit, https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023 (last visited Aug. 8, 2024).

such as evaluation metrics and tools for safety testing for private actors developing AI.[59] Meanwhile, domestic approaches include that of the Biden Administration's AI Bill of Rights. The AI Bill of Rights is a white paper policy framework published by the White House Office of Science and Technology Policy that addresses issues raised by AI from myriad axes: (1) promoting safe, effective systems that are tested pre-deployment and appropriately monitored to safeguard from malfeasance, (2) circumventing algorithmic discrimination, (3) bolstering data privacy, (4) requiring notice of the use of automated systems, and (5) human alternatives and various fallbacks that allow users of automated systems to effectively opt-out.[60] Ultimately, the AI Bill of Rights is more informative than legally relevant, but some of those policy aims seem to have fed into President Biden's Executive Order "on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence."[61] Alternatively, some projects propose encouraging unionization as a means of giving workers more say over how automation impacts their lives.[62]

Finally, a 2023 analysis from the Electronic Privacy Information Center synthesized a number of state laws and state agency regulations that address automation concerns to varying degrees—with the majority focusing on consumer privacy.[63] Alone, these approaches may not amount to much, but when taken together, they may pool into a wave of their own—one powerful enough, perhaps, to help contain the risks of displacement, algorithmic discrimination, and privacy concerns that AI heralds with its integration. While the effectiveness and usefulness of these policies is dubious, they demonstrate a global and domestic awareness of the various risks of AI and a concerted effort to meet them head-on from a sea of institutions, governments, and social organizations.

Conclusion

This paper explored the foundational dilemmas posed by AI with a particular focus on Descartes' methods of discernment to dive into the nature of consciousness and to cast a spotlight on the deception of the senses. Without the touchstone philosophical questions of identity and belief, discourse on the more pressing matters of mis- and disinformation and labor displacement stands on shaky footing—one categorized as a red herring and somewhat ignored by Suleyman despite being an inexorable feature of artificial intelligence. After all, any talk about the fruits of a topic eventually wind themselves back to talk about the tree they came from.

AI is inseparable from the public imagination of it. That imagination is bound up in decades of science-fiction fantasizing that is simultaneously awe-inspiring and frightening. Now that the technology has caught up to the fairy tale, it is more important than ever to embrace even the esoteric as to peel back the psychological layers of wonder, fear, confusion, and misinformation itself.

Then, this paper overviewed the key fears of automation, namely, the impending mis- and disinformation crisis and the high-exposure risks to displacement and AI complementarity. The goal therein was to present these problems accurately and survey various solutions, policies, and proposals aiming to ameliorate them. The chief spiritual danger is that AI will distort humanity not in truth, but in belief—that it will sever civilization from the umbilical cord of the very first innovation, one ubiquitous, evolving with human understanding, employed to the point of taken-for-grantedness—that is, reality itself.

---

59 *The Bletchley Declaration*, *supra* note 58.

60 White House Off. of Sci. & Tech. Pol'y, Blueprint for an AI Bill of Rights: Making Automated Systems Work for the American People 5-7 (2022), https://www.whitehouse.gov/wp-content/uploads/2022/10/Blueprint-for-an-AI-Bill-of-Rights.pdf.

61 Exec. Order No. 14,110, 88 Fed. Reg. 75191, Doc. No. 2023-24283 (Nov. 1, 2023), https://www.federalregister.gov/documents/2023/11/01/2023-24283/safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence.

62 Christina J. Colclough & Kate Lappin, *Building Union Power to Rein in the AI Boss*, Stanford Soc. Innovation Rev. (Sep. 20, 2023), https://ssir.org/articles/entry/building_union_power_to_rein_in_the_ai_boss.

63 Katrina Zhu, *The State of State AI Laws: 2023*, EPIC (Aug. 3, 2023), https://epic.org/the-state-of-state-ai-laws-2023/.

DIALOGUE

Artificial Intelligence & the Apocalyptic Imagination

A Conversation with Michael J. Paulus, Jr. on Artificial Agency and Human Hope*

Interviewer: Anton Sorkin

Q. Michael, thank you so much for speaking with me about your book and the all-consuming topic of Artificial Intelligence. My first question is an obvious one: Why did you decide to write *this* book?

A. Most of my professional career has been focused on helping institutions and individuals understand and participate in the digital transformation of our lives and world. A lot of that work has been administrative, leading the development and management of academic programs and technologies for faculty, staff, and students. I have always tried to integrate teaching and scholarship into this work. Teaching enables me to engage with students. Research projects deepen my understanding of the past and present, clarify my thinking about the future, and help me connect with other faculty. In 2019, seeing the need for more hopeful and constructive approaches to AI, I recruited a group of faculty colleagues for a research project to look at AI, faith, and the future from the perspective of various disciplines. That project culminated with a collection of essays published in *AI, Faith, and the Future: An Interdisciplinary Approach* (Pickwick, 2022). Like other members of that research group, I was inspired by our conversations and collaboration to do more teaching and research related to AI. I wanted to delve deeper into underutilized historical and eschatological resources that can help us think about and shape the future of AI..

Q. I'm curious if you saw the AI revolution heading toward the public market or were you as surprised as I was at just how quickly things like ChatGPT became a household name?

A. For many years I was trying to get people to think critically about how AI was already part of our daily lives. I expected it would continue to become more and more important and impactful, but I was not expecting what ChatGPT caused: society-wide interest in and access to a new and powerful form of AI. The month before ChatGPT was released, some of my staff and I met with a few of our writing faculty and discussed whether we should offer faculty development related to AI writing tools. Together we experimented with a number of them, and the writing instructors decided they weren't too worried about their impact—they were generating, at best, mediocre papers that wouldn't receive good grades. After the release of ChatGPT, we were having very different conversations and offering well-attended workshops to help faculty explore how large language models such as ChatGPT could be integrated into writing and other courses.

Q. How are you personally integrating AI into your life?

A. Recently I participated in an unconference I helped organize on AI and academic libraries, and I found myself in a breakout session with another library dean. We were talking about AI-

* Michael Paulus is the University Librarian at Creighton University. His administrative, teaching, and scholarly interests focus on the future of information and communication technologies, especially as they relate to institutions of higher education, libraries, work, and ethics.

proof jobs, and someone in the room had the courage to ask if a dean's job could be automated. The other dean and I exchanged a knowing look and then explained how much of our work was about cultivating human relationships: motivating and supporting diverse people, building teams and collaborations with different groups, and handling personnel challenges. We acknowledged that many of our tasks could be automated, such as routine communications and formulaic reports. Like many others, I am always exploring new tools to see how AI can help me automate more routine activities, as well as augment my creativity. In both my professional and personal lives, generative AI has become something of a conversation partner—making suggestions, proposing syntheses and interpretations of information, predicting what might follow my previous patterns. But at this point, I would say the integration is marginal. I would like to see more AI integrated into my life to handle more mundane tasks. But for more important tasks, such as those that require attention, discernment, or discipline, I have become comfortable with several no- or low-tech practices. I still begin most of my thinking and work without a computer—thinking during a run or walk, writing on notepads or in a journal, talking through ideas with others. But when I move my work into my digital workspace, I have a new resource to interact and experiment with.

Q. I appreciate that background! Let's jump into the book with some definitions. What does "apocalyptic imagination" mean and how does it relate to the concepts of Babylon and New Jerusalem?

A. Many years ago, I read John J. Collins's *The Apocalyptic Imagination* for a doctoral seminar on "the Apocalypse and the apocalypses" at Princeton Seminary. I found it very helpful for understanding apocalyptic groups such as the community that created the Dead Sea Scrolls and the earliest Christians who wrote books like the Apocalypse of John. Like most people, I first found this last book in the Christian Bible rather strange. But John's Apocalypse was participating in an imaginative tradition that attempts to see a deeper meaning and purpose in a world that seems inconsistent with the promised and hoped-for world. One can look at the images in this literature, for example the cities of Babylon and New Jerusalem, and glimpse a fuller revelation of a spiritual reality hidden within material reality. Babylon is Rome in the first century, but it is also a figure of all empires and their ultimate end. New Jerusalem is an image of an alternative city and a greater power—present in some ways, but not yet fully manifested—which is displacing Babylon. These images are meant to help hearers and readers of the text imagine and participate in the creation of a better world that is not only a future promise, but also an emerging present actuality. And, because these images include human cultural artifacts, they are a generative resource for thinking about artificial agency and agents in the context of human hope.

Q. On this question of the role of imagination, you make a fascinating statement: "As we continue to design and develop our increasingly complex technological society, it is important to understand how AI is shaping our imagination as our imagination shapes it." Can you elaborate how this cyclical relationship works between our imagination and AI?

A. Like every cultural artifact, AI existed first in the human imagination. At Amazon's first major public AI event in Las Vegas in 2019, Jeff Bezos talked about how he organized his library. On one side, over the fireplace, he had the word "builders" and there he had books by builders. On the other side, it said "dreamers" and there he had books by dreamers. Bezos pointed out that human creativity needs both dreamers and builders. The dreamers inspire builders, and builders build things that enable dreamers to dream new dreams. Amazon's artificial agent Alexa, he noted as an example, was inspired by the Star Trek computer.

But more than dreamers and builders are shaping us and our world. The things and environments we create shape us as well. It is fascinating to observe how, from the beginning, the human imagination has been open to other forms of agency: divine, of course, but also artificial. Artifacts, such as books and cities, have often been described as having their own degree of agency. Now, as we are creating artifacts with new and greater forms of agency, it is important to consider how they will shape us in new ways. Moreover, because these artifacts can simulate us and include so much of us—in, for example, all the products of human intelligence on which

LLMs are trained—AI is becoming a new source of revelation about human nature and destiny. Having a disciplined imagination is important in all this. We do not want to be too fearful or dystopian, thereby deemphasizing human agency. Nor do we want to be too optimistic or utopian, which overemphasizes human agency. Technologies are not simply deterministic, nor are they simply neutral tools. We are shaping our technologies as they are shaping us, so we need to be careful and intentional about how we are integrating them—or not—into our lives and world.

Q. This brings up another thing I've been curious about, which is your take on the nature of Christian resistance to change. This is a common problem within the history of the church of being too often behind the changes in society and then overcorrecting by becoming reactionary and alarmist. How do we go about finding the right balance between staying vigilant regarding those things too "Babylonish" with the technological changes and being adaptative to what many are saying is an inevitability?

A. A lot of Christian resistance to change perplexes me. This history of the church—the invisible church always, if not its institutional manifestations consistently—is about resistance to the status quo. From their revolutionary beginnings, Christians were about making all things new. The church has been at the center of so many social and technological innovations. Consider the history of books, schools, hospitals, and social reform movements. Certainly, some changes need to be resisted. But resistance to change can reveal a complacency with the status quo, which itself should be resisted. Resistance to change also seems too often to come from a fear of the future—and that reveals a weak theology of hope. Critiques of change should inspire the pursuit of alternative changes. And if those pursuits are aligned with the trajectory of new creation, already underway, then they will participate in something more powerful and transformative.

Q. You invoke the need for a knowledge and wisdom revolution to accompany that current industrial and information revolution. This reminded me of something Dallas Willard wrote in *The Spirit of the Disciplines*, where he notes how civilizations in the past have often failed to develop the necessary number of capable leaders to sustain the needs of society. Are you concerned that we don't have enough capable leaders today to absorb, to borrow from Mustafa Suleyman, this "coming wave"?

A. One of the fascinating things about the conversations we are having about AI right now is that we are having them. These discussions were already robust before OpenAI's release of ChatGPT, and now everyone is talking about AI. I was recently revisiting a book about the future of libraries in 2020, published in 2013, and there is very little mention of AI in it. There are significant discussions about new technologies, but only a couple of contributors anticipated the integration of robotics into the world. Other mentions of AI technologies were brief—wondering how Siri might improve—or jokingly apocalyptic, referencing robot overlords and Skynet. An updated version of that book was published this year and imagines the future of libraries in 2035. AI is mentioned on the first page, in nearly half of the contributions, and one chapter consists of reflections on ChatGPT's responses to prompts asking about the future of libraries. In other disciplinary and professional domains, you see something similar happening. When I started conversations on my campus about AI five years ago, there were a small number of interested administrators and faculty. Now, it is harder to find people who are not interested in thinking about its impact on teaching and scholarship.

Historically, reflection about the negative impacts of social and technological progress often follows too slowly or too late—even though these are integral to innovations and not entirely unforeseeable. As Paul Virilio says, "To invent the family automobile is to produce the pile-up on the highway." But it took decades to create things like safety glass and rules for the road. And we are still trying to address the social and environmental impacts of designing built environments for small vehicles powered by fossil fuels. So, I am encouraged by the diverse range of voices speaking into conversations about how we might use AI reflectively and responsibly. I am also encouraged by the content of these conversations, many of which are about larger social systemic or structural issues.

Q. To pull on this knowledge thread further, what will be the role of divine wisdom in the way that Proverbs 8 intends or, even more practically,

the way James 1:5 talks about? Or to ask it more bluntly, are you at all worried that we're heading for another Tower of Babel where man attempts to live in a city completely of his own design?

A. At this point in human history, the image of the Tower of Babel—as a symbol of technological or imperial excess—seems rather frail. We will keep working on Babelian projects, either by unintentionally working in a part of the city that is actually part of the tower or by intentionally elevating our autonomy by exploiting or eliminating others' to build that tower. But history and nature keep revealing how futile these projects are: the human spirit resists; empires, each weaker than the last, fall; and nature has an agency that we cannot control—and the more we attempt to control it, the more it overwhelms us. All of this is in the Apocalypse, with the doomed city Babylon as the pinnacle of all Babelian endeavors. But the Apocalypse also reveals the divine wisdom and power that leads to the ultimate triumph of the city from God (i.e., New Jerusalem) in which humans and human works thrive. The more we seek divine wisdom—which begins, as James and Proverbs say, by asking and seeking for it—the better we are able to discern strategies for turning away from deformative works and participating in transformative works of new creation.

Q. You have a quote in Chapter 3 that reads: "[t]he recovery of a deeper and broader apocalyptic imagination provides us with a richer perspective on and vocation for technology." Can you talk to me about this vast confluence of vocation, technology, and human purpose as it relates to the Christian mission and how you see yourself in this great upheaval?

A. For me, what we're talking about here is living into a greater hope. Technology has always been an integral part of human history and our vocation in the world, but for a long time we thought of it as secondary—something that was not central to who we are and also something that we could control. That wasn't really the case, even in ancient world. With transformative technologies such as AI, we can no longer ignore the reality that we have specific responsibilities related to the creation and use of technology in our lives and world. There were many great Christian critics of technology in the twentieth century, and there is much we can learn from them. But many of them seemed to embrace a view of the human vocation that did not include new information technologies: these social-technological systems were something to be rejected and resisted. There is another tradition in the history of Christianity, however, which is more affirming of technological innovations. This perspective leads to more constructive approaches to technologies, inspiring us to fix what is broken in the world and create a better and more just world. I believe that perspective is what is needed most now.

Q. You talk about the role that technology plays as a "central catalyst for our biology" in discussing the work of Nigel Shadbolt and Roger Hampson. There is an incredible amount of potential that comes with this new technology for the sake of human enhancement. But at the same time, as Luciano Floridi notes in his book, *The 4th Revolution*, technology can also reduce our self-importance (i.e., we are no longer as unique as we once thought). Are you worried that with AGI, humans will lose its special role as thinking creatures and become as Jacques Ellul worried, merely "a slug" that plays a catalytic function?

A. Floridi points out that modern scientific revolutions have humbled us. Our uniqueness is not defined by our physical place in the cosmos, our biological separation from other species, our pure rationality, or our monopoly on intelligence. All of these revelations are good, for they help clarify who and what we are, our place in the world, and our responsibilities. I am skeptical that we will create AI with intelligence comparable to general human intelligence. We already have many automated systems that can outperform humans in a number of discrete intelligent tasks, and it is likely we will soon have more that outperform us in many others. But human intelligence, which we do not fully understand, is rather different than artificial intelligence. Human intelligence is connected with attention, imagination, affections, values, and more. AI may be able to simulate aspects of human intelligence, but it is actually a new and alien form intelligence with different processing mechanisms.

Q. You write in your last chapter that before we live much longer with AI, "we need better social and technological attentional strategies for cultivating critical reflection on and refinement

of hopes and goals." Can you unpack that in the context of building a better shared future?

A. In the first chapter of the book, I talk about the first information revolution connected with the emergence of our species: the capacity for reflective attention. As human societies and technologies became increasingly sophisticated, we developed social systems and structures for enhancing that capacity: schools to cultivate attentional practices, libraries to curate what is worthy of attention, and legal frameworks for protecting the autonomy upon which attention depends. Given the new ways AI can be used to abduct and manipulate our attention, we need new information practices, social institutions, technological supports, and legal protections to preserve our attention—our ability to imagine, define, and realize our hopes and goals. The imagined futures I engage with in the last chapter of the book do not help us much in imagining what will be needed. Which is why I mention at the end of the chapter an idea I really like from Neal Stephenson's novel, *Fall*: An Organization for New Eschatology, which is established to consider the implications of digital afterlives.

Q. To follow up on some of these threads regarding the potential for danger, we are becoming increasingly aware of the harm that smart phones have wrought on the development of children. Many states are passing bans on smart phones, for example, to remedy some of these trends. While at the same time, tech companies are rolling out new forms of digital manipulation tools (e.g., Gen-3 Alpha Turbo, Grok 2.0, Midjourney) that are almost indistinguishable from real videos. How do you see these changes impact the social development of the next generation who will be raised on "smart[er] phones?"

A. About a dozen years ago, I started teaching a graduate class on ethics in digital education. Most of the students were teachers in K-12 schools, and they were very excited about using new technologies in their classrooms. My challenge, then, was to get them to pause and reflect on the values that would inform their decisions about technologies. We looked at the negative impacts that social media and other digital technologies were having on children—we've known about these for some time—and I was viewed as a tech pessimist, although I was really just encouraging caution. About eight years ago, the public view of digital technologies quickly turned negative, and I had to accept being seen as more of a tech optimist to keep them engaged with these technologies. We were too naïve and passive about the impacts the internet, social media, and mobile technologies would have on us. One of the hardest aspects of being a parent was managing my kids' tech use! We have a chance to do better with AI, but we need to be thinking more holistically about human development. And we need society-wide solutions: corporate accountability, government regulations, educational reform, communities that support us, and public institutions that advocate for us.

Q. Let me shift our focus here at the end. I understand that you're moving into a new role in as the university librarian at Creighton. This is fascinating shift given that Amazon's Alexa was named after the Library of Alexandria, but also given your number of references to the importance of libraries for the curation of knowledge. Notably, you write how libraries were able to adapt to the new dynamics of the information age, while other institutions were ill-prepared for a "future being shaped by automated information processing." Am I seeing a viable connection in all this?

A. I have been a librarian most of my professional life. Although I first started thinking about technology as a finance professional (including conducting Y2K audits), and for some time now I have been exploring technology from a theological perspective, my primary disciplinary and professional home is in library and information studies. Like theology, this is an interdisciplinary and applied area; it includes historical, technical, ethical, and practical concerns. Although many people fixate on old images of libraries—such as card catalogs (which I never used) and books on shelves (which are still there but are only part of a library's collection)—libraries have been leaders in adopting information technologies and automating processes for decades. That's why we no longer have card catalogs and can access so many resources online through libraries. Libraries provide an important model for adapting new technologies in ways that balance continuity and change and center human agency. My next book focuses on the roles and responsibilities libraries have—through designing better automated information processes, as well as cultivating new

human information practices—to empower all of us to comprehend, critique, and collaborate with AI, and ultimately to create a better information environment for both human and artificial agents.

Q. As a follow up to the idea of roles and responsibilities, you talk about the connection between our understanding of AI and its contribution to the common good. You call on religions and other social groups to play their role in cultivating "formative and counter-formative practices that ensure human agency is aligned with goals greater than those encoded in AI systems." What does that look like practically?

A. Religious communities have always been about formative practices, and they have a responsibility to help us preserve and adapt those practices in connection with new technologies. Otherwise, we will be unreflectively formed by these technologies and those who designed them, or we will be overwhelmed by them when our individual attempts to resist their influence proves unsustainable. For a number of years, I taught an undergraduate class on information and attention. The two main goals of the course were to help students understand and critique digital technologies and culture and to be reflective about their engagement with them. The reflective resources I introduced them to were Christian spiritual disciplines, which are primarily about cultivating attention. These practices have come to us, through the centuries, through Christian communities. And they are most effective when they are practiced within communities—communities that guide and encourage those who use them.

Q. I appreciated your use of Willie Jennings in the Epilogue. In one passage, you talk about an unforeseen future in Christ and that "technology can drive alternative revolutions as we live more fully into the reality of New Jerusalem." I'm curious, what do you think our future holds when it comes to our capacity to love our neighbor and help them see the reality of new creation?

A. The Epilogue is really a substitute for a final chapter on the outcome of our current information revolution associated with automation. The first information revolution resulted in the capability for reflective attention to imagine and create alternative futures; a second gave us the structural agency to create cities and complex societies; a third enabled us to augment knowledge through textual artifacts. We do not yet know what the results of intelligence automation will be, but we can begin to imagine and participate in shaping that future now. Since Acts is about the experience of the first Christians entering a new world, it seemed like a good beginning to end with. Their apocalyptic imagination enabled them to find a new way of living into new creation—the way of love—that rejected the imperial option, of attempting to control the world, as well as the option of avoiding the world, by withdrawing from it. Now, we need to imagine new ways of living in a new world with artificial agents. One of the most important questions we should be asking is how humans will continue to cultivate virtues such as love when we can automate many of the activities that are opportunities for experiencing, expressing, and strengthening that virtue. As we develop new automated processes, we will need to retain some human practices—as well as develop new practices—for cultivating love and other important virtues.

Mustafa Suleyman, The Coming Wave: Technology, Power, and the 21st Century's Greatest Dilemma (Crown, 2023). 352 pp.

*Book Review by Rick Campanelli**

Introduction

Just a decade ago, physicist Stephen Hawking upset techno-optimists when he said the greatest existential threats to humanity come from science and technology. A year later—upsetting a whole different faction—Hawking said that only a world government could save the world from these threats. In *The Coming Wave*, Mustafa Suleyman provides a compelling explanation of how much and how quickly Artificial Intelligence (AI) ups the ante on the kinds of threats Hawking anticipated, arguing we need to find new systems that can adequately and quickly respond to constrain them. The book also begs a question: in responding to these new categories of risk, will we be more willing to embrace solutions that undermine democracy and human flourishing?

Suleyman is the consummate AI insider. He co-founded DeepMind and Inflection AI—the former became Google's AI powerhouse, and Microsoft is a primary if not dominant investor in the latter. In *The Coming Wave*, Suleyman offers an understandable explanation of just what Artificial General Intelligence (AGI) is, and how it creates vast, transformational potential on an unparalleled scale. His insider expertise makes the book clear and credible. Unfortunately, it also makes it difficult to dismiss its warnings. And Suleyman is worried. He warns that we are entering "an era when unprecedented opportunities . . . [are] matched by unprecedented risks." Is he right? Reading *The Coming Wave* is a good way to get a handle on what AI is about, and what's at stake.

Overview

In the first part of the book, Suleyman describes how in just over a decade he came to appreciate the potential for AGI. He walks us through continuously accelerating computing leaps by teams at DeepMind, OpenAI, Google, Facebook, Microsoft, and the relatively small array of actors capable of amassing the vast resources needed to build these new systems, and that dominate the playing field. He explains how initial successes exponentially expanded AI capability and capacity; how large language models (LLMs—which he also explains) opened the flood gates; and how AI systems went from learning from human-curated inputs to creating and developing their own learning models and strategies. "A key ingredient of the LLM revolution," he explains, "is that for the first time very large models could be trained directly, without the need for carefully curated and human-labeled data sets." We watch the technologies transmute from fun gaming competitions into serious national security arms races among nations, which quickly realized they were on the brink of new kinds of wars. We are already seeing these waves beginning to break in Ukraine and Israel.

These chapters alone make the book worth reading, just to understand what we really are, or should be, talking about. But what also emerges is a portrait of highly enthusiastic and intelligent engineers, single-mindedly competing at breakneck speed to be on the ground floor, to lead and to be enriched by what they suspect may be the greatest technological wave in human history. And while Suleyman describes lots of conversations among tech leaders about practical and

* Lecturer in Public Policy, University of Virginia Batten School of Leadership & Public Policy; Director of Graduate Ministries, Center for Christian Study (Charlottesville). In the past, Rick served as Director of the Office for Civil Rights and Counselor to the Secretary, U.S. Department of Health & Human Services.

ethical implications, he worries there is no corresponding passion, incentive, or mechanism to appropriately address the accompanying risks—most of which will not be borne by the developers but by all of us. That question, perhaps the most important question, of whether and how to address the disruptive risks that accompany the massive potential benefits of AI is left to others—whoever they are.

What's the Big Deal About AI?

We might have read articles on how AI touches on specific areas of interest to us, and its tremendous potential to disruptively research, write, strategize, and produce improvements ranging from unimagined therapies and diagnostics, to advances in agricultural and national defense capabilities. If we have, we also might have a lurking sense that each of these varied threads fails to reflect the big picture of the AI tapestry. According to Suleyman, we should pay attention to that sense.

We might wonder, for instance, will lawyers, doctors, farmers, researchers, coders, store clerks, and factory workers still have jobs? Will students be able to cheat undetectably on tests, or will we enhance or harm human analytic capacity? Will we open a new era of creativity or undermine human creativity? Will AI result in the greatest increase in general productivity and wealth the world has ever seen, or will it be accompanied by the exacerbation of the most profound disparities in wealth and class? Will fake news, counterfeit videos and voices, and bot-driven fake stories compromise our ability to distinguish what is true and undermine national elections? Will we provide soothing robotic friends to children and the elderly, and will that come at the cost of coming to believe that these robotic equivalents are nearly as good as real relationships and so devalue what it means to be human?

These questions are all profoundly important in their own right, and *The Coming Wave* touches on many of them. But the very breadth of the questions suggests something even more significant about AI—that there is a forest we are missing for the particular AI trees. Suleyman explains that to understand AI's impact, we must grasp the meta-implications of this meta technology—a "general" technology, being disseminated on a scale and at a pace never experienced in human history.

A "General" Technology

As *The Coming Wave* explains, AI is a "general" technology—not only powerful in its own right, but also an accelerant of benefits and risks for almost every other advanced technology. Here is Suleyman:

> It's not just a tool or platform but a transformative meta technology, the technology behind technology and everything else, itself a maker of tools and platforms, not just a system but a generator of systems of any and all kinds. . . . We really are at a turning point in the history of humanity. . . . It won't be long before AI can transfer what it "knows" from one domain to another, seamlessly, as humans do. AI is far deeper and more powerful than just another technology.

We tend to think of AI as a powerful train coming toward us and focus on its implications for the realities with which we are most familiar—our careers, our finances, our families. Understandable, but to get the big picture it would be better to think of AI as fueling a thousand trains coming toward us and everyone else, in almost every field—some of which are laying their own tracks. Or maybe we tend to think of the AI wave as an army of elephants that will trample some and build new roads for others. It may be better to think of it as millions of ants, many of which are moving and interacting autonomously.

Democratized and Disseminated Power on a Scale and Pace Previously Unseen

The Coming Wave goes on to describe another key element, which distinguishes this "wave" from all prior technological waves: for better and for worse, it is democratizing and disseminating at unprecedented scale and acceleration. In economic terms, the barriers to entry to the use of these powerful, AI-enhanced tools are extremely low, relative to the vast potential externalities they will create. And Suleman points out that at this disseminated scale, it will be much harder to detect and respond to harms that can arise.

Applying AI to technologically advanced areas from biotech to robotics to nanotech, Suleyman provides compelling descriptions of how the power of AI has already become de-

mocratized—disseminated at relatively low and rapidly decreasing cost into the hands of individuals and small groups, governments, and rogue actors. He readily acknowledges that many, if not most, of these empowered individuals and groups will be responsible, prudent, and well-intended, working for universities, companies, and governments exercising rational oversight, with every incentive to avoid a major mistake.

Even so, as post-COVID19 investigations have demonstrated, these efforts sometimes go horribly wrong: experiments are ill conceived and executed, redundant systems fail, unanticipated results play out, potent experiments leak. And "dual use" threats arise from imprudent researchers or hobbyists without adequate knowledge of what is at stake, for instance, when modifying bacteria or human genetic traits; or from those with nefarious intent and nothing to lose. Suleyman points out that these AI risks are all the more potent because they are fueling and making more accessible other groundbreaking technologies.

If this all seems too sci-fi or conjectural, consider that, in 2016, U.S. Director of National Intelligence William Clapper surprisingly put a new class of genetic engineering techniques like CRISPR/Cas9—described as "word processing for gene editing"—on the annual worldwide threat assessment list of weapons of mass destruction and proliferation. These techniques are "just" tools, not toxins; but they made it to the list because of low barriers to entries and potentially extreme externalities: its broad distribution, low cost, and accelerated pace of development could have "far-reaching economic and national security implications."

Suleyman describes at some length how AI exacerbates and accelerates these threats. He points out that DNA synthesizers now cost only a few tens of thousands of dollars and are small enough to sit on a bench in your neighbor's garage, so that people with "grad level training in biology or an enthusiasm for self-directed learning online . . . could soon create novel pathogens far more transmissible and lethal than anything found in nature." In fact, last year the *Wall Street Journal* reported that *undergrads* at MIT were asked to test whether AI-powered chatbots could be prompted to assist non-experts in causing a pandemic. Within one hour, the chatbots produced four potential pandemic-causing pathogens.

The Undertow of the Coming Wave

A corollary concern of the AI risks is how we will respond to them. Suleyman is at pains to describe the risks in a balanced way and always in the context of his inherent optimism about the good AI will bring. But at the same time, he is arguing that it is essential that the world find new means of recognizing and containing the risks, or the harms may well outpace the good, on a scale the world is not prepared to accept. And he says—given that the wave has arrived—we have to do this quickly.

The last section of the book offers an array of containment options, which Suleyman passionately urges. He explains why, in the history of technology, no general technology has ever been put back in its bottle—so the goal must be appropriate containment, which can save us from the Scylla and Charybdis of unacceptable catastrophe or authoritarian government responses. The options he lays out are broad, ranging from regulatory solutions (he does not believe current regulatory structures, as they exist, are able to keep up), to proposing facile and empowered industry-government collaboratives, and a "NASA moonshot" type of initiative to fund innovative efforts to stay ahead of the risks.

You can decide for yourself if you think these are realistic. Suleyman himself seems to doubt that they are, if only because they would require unprecedented cooperation among industry and nations that may not see the need or that are too intent on maintaining an advantage. Failing those solutions, in a way echoing Hawking but also warning against authoritarian solutions, he calls for new world systems with regulatory authority and ability to respond quickly, and collaborations among nations—including among those that do not share a commitment to inalienable rights for their own citizens or for others. Such is his assessment of the risks.

This might be called the coming undertow of the *coming wave*: if things go wrong, and calls arise for exercise of far greater government power to protect us, will people who care about democracy and human flourishing, and the inalienable rights of persons, acquiesce? If we don't want to, then we must decide what the risks really are, which risks we are prepared to accept,

and which we are not. Prudent responses, where possible, may help us stand against the undertow of sacrificing fundamental and inalienable rights.

One last note: the book touches on, though not at length, another important AI disruption—the impact on our understanding of human flourishing and human dignity. Other writers have expressed concern that we may come to rely on efficient and seemingly "neutral" AI solutions, even though they are not human-defined or necessarily respectful of human life and dignity. As a case in point, Marc Andreessen, AI investor and optimist, who says that AI may save the world, cheerfully argues that "every child will have an AI tutor . . . helping them maximize their potential with the machine version of infinite love." Sam Altman of Open AI says he does not want AI friends, and does not recommend it, but that may be what we learn to want: "I personally have deep misgivings about this vision of the future where everyone is super close to AI friends, more so than human friends. . . . [A] nd some people are going to build that, and if that's what the world wants, and what we decide makes sense, we're going to get that." That would be a poor, destructive counterfeit for real love. We are already seeing the dramatic results of loneliness and depression accompanying curated, digitized smart-phone relationships, driven by reinforcing algorithms that vie for our attention and engagement. AI-enhanced technologies will take these capabilities to whole new levels, if we are not alert and disciplined enough to resist temptation.

The world needs to hear from those who know about real love, and real personhood. People who believe we are made in the image of God, who live in the hope of the incarnation, redemption, and the knowledge that we are made to be loved and to love, and who understand our responsibility to steward the beautiful world God has entrusted to us, have something of great value to offer a world tempted by these counterfeits. We should be ready to offer that good news because as the coming wave breaks, the world is going to need it even more. A good way to start getting ready might be to read *The Coming Wave.*

Daniela Rus, The Heart and the Chip: Our Bright Future with Robots (W.W. Norton & Company, 2024). 272 pp.

*Book Review by John M. Rogitz**

Introduction

The Heart and the Chip by Daniela Rus is an easy, enjoyable read that delves into the intertwining realms of artificial intelligence (AI), machine learning, and robotics, while exploring the ethical, social, and philosophical implications of these rapidly evolving technologies. As an MIT professor and pioneer in robotics, Rus provides a comprehensive review of the potential benefits of integrating robotics into our daily lives and also discusses the current and future limitations of these robotics. From a biblical perspective, this book offers ample opportunity to discuss the intersection of faith, morality, and technology.

Overview

Rus's work is a blend of futuristic vision and current technological reality. She navigates through topics such as machine learning (a subfield of AI), robotics, and human-computer interaction, projecting a future where AI and robotics are deeply intertwined with human life.

Overall, the book is very positive. With its optimistic take, *The Heart and the Chip* discusses how AI robots ("the chip") can complement humans ("the heart") and enhance our daily lives. It envisions a world where the heart and the chip are partners rather than opposing forces. In fact, the very first page of the book declares, "Robots aren't going to steal our jobs. They're going to make us more capable, productive, precise." While that's debatable for reasons that will be discussed later, the substance of the book itself will be discussed first.

The book has three overall components to it. First, the book begins by dreaming big, discussing all kinds of futuristic, fantastical implementations of robotics powered by AI. Not a lot of it is terribly existential. Rather, the book discusses exoskeletons, air-borne robotics, and self-driving cars, delving into how these types of machines would be constructed and how they would be programmed according to the "sense, think, act" paradigm for robotics.

The book then discusses the reality of the current state of the art and its technological limitations, making the point that all-purpose robots are still a ways off. The book also discusses how even small robot errors can lead to significant negative consequences and how we still have not solved many of these problems. There are also inherent limitations in current battery tech and processor capability that limit the broader implementation of robotics.

From there, the book begins discussing the responsibilities of technologists and humans at large to use robotics and AI responsibly. It posits that humans should always retain ultimate decisional authority and that regulatory bodies should be established to ensure ethical, benevolent use of these technologies.

Overall, Rus's book is very heavy on the science, which isn't necessarily a bad thing. In fact, it does a really good job of explaining the basics of AI to the average person. Just be prepared for all sorts of detailed scientific discussions on various fields of robotics.

Intersection with Faith

The Bible clearly states that humans are created in the image of God.[1] This foundational belief emphasizes the unique value and dignity of human beings. In *The Heart and the Chip*, Rus envisions a world where AI can augment and even surpass human abilities. This raises important questions about what it means to be human.

* Managing Attorney, Rogitz & Associates; Executive Committee of the IP Section, California Lawyers Association; Adjunct Professor, Trinity Law School..

1 *Genesis* 1:26-27.

But for the Christian, the answers to those questions are fairly straightforward. God did not create robots in His image; He created humans in His image. Robots will never have a soul. They will never have a personal relationship with Jesus. They will never have reason to ask forgiveness for their inherently *human* sinful nature. So they're not human, and they'll never even begin approaching that.

The more difficult questions involve what our inherently sinful nature might do with such powerful technology. Rus's book suggests potential "guardrails" society might put in place. These "guardrails" include humans having final decisional authority, establishing regulatory bodies to regulate AI and robotics, having researchers and scientists take their own version of the Hippocratic Oath, and "educating" people at large on the perils and potential pitfalls of these technologies.

In terms of humans having the final say, the problem is that this seems extremely difficult in practice. After all, we're designing all kinds of autonomous robots—autonomous drones with military-grade weapons, *autonomous* vehicles that drive through crowded areas, and even autonomous surgical devices. They're designed to be autonomous, and, when they get more advanced, they will be able to do things beyond what they've been explicitly programmed to do. In fact, we're more or less at that point right now.

Government-backed regulatory bodies are another reason for hesitation. A famous quote from President Ronald Reagan comes to mind: "The nine most terrifying words in the English language are, 'I'm from the government and I'm here to help.'" More often than not, government intervention creates more problems than it solves. That's why our Framers designed a limited, divided government here in the United States. Yet the book posits that regulatory bodies should define clear business and societal "contracts" for society to uphold. But is that going to stop China or Russia or Iran from doing whatever they want? Has it stopped them before? How much authority does Rus think multi-national agencies like the International Criminal Court (ICC) actually have?

Having researchers and scientists take a Hippocratic Oath is also naïve. The oath does not even stop actual physicians from crossing ethical barriers here in the U.S., let alone in other countries with less-than-stellar ethical records. Again, think China, like how they've harvested organs from political prisoners and prisoners of conscience.

When it comes to "education," this also seems like naïve political-speak based on the faulty premise that if we could just talk to people and explain the perils and pitfalls of technology, they would listen and choose a righteous path. Again this fails to account for the inherently sinful nature of humanity that Christians readily recognize but that humanists fail to acknowledge, instead believing that humans are all capable of being good actors even if some are downfallen through their own personal circumstances. These two world views are fundamentally different and competing.

Perhaps recognizing these shortcomings, Rus does in fact briefly admit that she is under no delusions about what might happen with such powerful tech but still somehow envisions a utopian future for the robots and AI. Yet she does not offer any satisfactory answers to the pressing question of how we can *really* safeguard humanity from destroying itself with AI and robotics. This is probably because there are no satisfactory answers, at least not outside of faith-based ones, and Rus does not readily confess to being a woman of faith.

So what might we actually do to prevent robots and AI from being used for nefarious, destructive purposes that threaten our very existence? I don't have perfect answers either, but then again, I'm not a technologist writing a book on the subject. However, my gut tells me that a few age-old notions are timeless for a reason.

One of those notions is very practical: Peace through strength. Being far superior than our adversaries on a technological level. Not just relying on people in other countries with different values to be good actors, which is fundamentally not a Christian perspective because it ignores the reality that humans are inherently sinful. Like our founders when they established a limited government on the Christian belief that we live in a fallen world with imperfect people, we should not be under any delusions ourselves. We should assume that bad actors will ignore our "guardrails" and use technology for nefarious purposes to realize all the scary implementations they can dream up. That's why we as Christians and Americans should be models for technology ethics and morality but, at the same time, demonstrate peace through strength. It's hard to imagine something more catastrophic than modern nuclear warheads being launched against us, yet our own nuclear strength and nuclear defense capabilities

have kept the likes of Iran and Russia at bay for decades. The same principle could be applied here—we should always be the leaders in innovation, and we should always be one step ahead. Therefore, one answer is for a good and moral people to continue leading the world, leading in technology, and leading in strength.

But more than that, our faith provides an even better answer. A truer one. A purer one. But still a faith-filled one. Ultimately, we must give all things to God. He is in control. Things happen or don't happen on His choosing. Our outcome is predetermined. The battle is already won. And we know that, eventually, we as Christians will be in heaven with our Creator. As an MIT professor, maybe Rus cannot admit to this if she's even aware of it. But we can. We're already familiar with not having all the answers but knowing that God does. And we know that we're limited in what we can really control in this world.

In the end, God is the answer.

Generous support for the
Journal of Christian Legal Thought
is provided by

PEPPERDINE UNIVERSITY
School of Law
Herbert & Elinor Nootbaar Institute
on Law, Religion, and Ethics

HANDONG
INTERNATIONAL
LAW SCHOOL

TLS
πιστευθῆναι
τὸ εὐαγγέλιον
1 THESS. 2:4
TRINITY
LAW SCHOOL
TRINITY INTERNATIONAL UNIVERSITY
CENTER FOR HUMAN RIGHTS

FAULKNER
UNIVERSITY

The *Journal of Christian Legal Thought*
is a publication of Christian Legal Society.

CHRISTIAN
LEGAL SOCIETY

ChristianLegalSociety.org