Human Dignity and the Ethics of Artificial Intelligence: A Framework for Responsible Design and Use from the Perspective of Catholic Social Teaching

Joseph R. Laracy
Systematic Theology and Mathematics & Computer Science
Seton Hall University
South Orange, NJ, USA
joseph.laracy@shu,edu

Vassilka D. Kirova
Informatics
New Jersey Institute of Technology
Newark, NJ, USA
kirova@njit.edu

Cyril S. Ku
Computer Science
William Paterson University
Wayne, NJ, USA
kuc@wpunj.edu

Thomas J. Marlowe
Mathematics & Computer Science
Seton Hall University
South Orange, NJ, USA
thomas.marlowe@shu.edu

***Abstract*—The increasing use of artificial intelligence (AI) has brought ethical issues into sharp focus, affecting generative AI, predictive AI, robotics and autonomous devices, and more. In this paper, we examine AI ethical questions from the perspective of Catholic Social Teaching (CST), a framework with philosophical and theological foundations, governing interpersonal interactions and those with the natural world. CST offers a coherent, integrated, and non-utilitarian vision of social ethics, grounded in human dignity and the objective demands of justice and the common good. Other ethical approaches may fragment or relativize these principles, or lack an ultimate grounding for them. Thus, CST has the potential to make a contribution to AI ethics.**

Keywords— ethics of artificial intelligence, philosophical ethics, artificial intelligence, Catholic Social Teaching, human dignity

I. Introduction

The interdisciplinary approach of cybernetics pioneered by Norbert Wiener contributed to early artificial intelligence (AI) research by introducing ideas of learning, adaptation, and automation [1]. Researchers like Alan Turing built on these principles to develop symbolic AI, aiming to create machines that could reason and solve problems like humans [2]. Over time, AI evolved beyond cybernetics, incorporating deep neural networks, machine learning, and data-driven algorithms, but cybernetics' foundational concepts of self-regulation, decision-making, and automation remain central to AI today.

From its formal definition in the mid-1950's, AI has continually gained influence, significance, and impact. With advances in computer architecture and remote storage, the rapid development of predictive AI using data analytics, the increasing use of robotics and other autonomous devices, and the coming of age of generative AI, AI now has a major presence in industry, education, healthcare, and beyond. While there are clear benefits from the use of AI, there are also costs, risks, and numerous ethical challenges. There are many works examining the ethics of AI, as discussed in Section 3, including a number based on philosophical or theological principles. These works have examined the topic from various perspectives, including technical, social, enterprise, and personal dimensions.

In this paper, we use Catholic Social Teaching (CST), also called Catholic Social Thought or Catholic Social Doctrine, to codify an ethical approach, focusing on human dignity, but touching also on other aspects [3]. Although CST is inspired by Catholic theology, it suggests a social philosophy broadly applicable in the secular realm. The rest of this paper is organized as follows. In Section II, we overview the principles of Catholic Social Teaching, also touching briefly on its history. Section III deals with related work. Section IV then considers the core ethical challenges, followed by the principles and guidelines derived therefrom these considerations, and Section V looks at case studies and challenges for implementation. Section VI then offers practical guidelines for stakeholders, comparing them to IEEE guidelines and other standards, and Section VII presents our conclusions and discusses future work.

II. Catholic Social Teaching as an Ethical Framework

The Greek word ἠθικός (*ethikos*) denotes "moral" or "expressing character." Ethics, also known as moral philosophy, studies those acts of the will affirming or rejecting the order proposed by reason. It deals with the acts of the will in their order to each other and in their order to their end [4]. Ethics is a practical discipline that refers to human action with the purpose of being morally good. Whereas the natural sciences are descriptive in character, concerned with empirical facts, ethics is prescriptive, concerned with normative values. As a normative value-based discipline, ethics is more concerned with what ought to be than what is. The classical Thomist approach to ethics focuses on three criteria:

979-8-3315-3228-4/25/$31.00 ©2025 IEEE

TABLE I. Principles of Catholic Social Teaching and Resulting Responsibilities

Principles of Catholic Social Teaching					
Human Dignity	Subsidiarity	Solidarity & Common Good	Charity	Universal Destination of Goods & Social Justice	Stewardship & Sustainability
Respect for all Sanctity of life Lack of bias Privacy & confidentiality	Local decisions Local organizations Local ownership where possible	Rights & responsibilities go together Rights of workers, users, etc. Prioritize families & communities	Care for all, especially the poor and vulnerable	Ownership and responsibility are broadly shared Fair participation for all	Care for creation Environmental conservation Climate concerns Sustainable infrastructure

I. the moral object (i.e., the substance of the action, the end inherent to the chosen act),
II. the intention of the acting subject (i.e., the reason or purpose, for which the act was chosen), and
III. the circumstances or concrete conditions surrounding and touching upon the action (i.e., the totality of the reasonably anticipated consequences of the act).

The moral object is the substance of a human act and answers the question: "What have you done?" The intention answers the question, from the agent's point of view, "Why is this act being done?" The circumstances of an action are the particular conditions of each act that are outside the act (and yet in some way touch upon the act). To be a morally good act, all three elements must be good [5]. As Anderson et. al. points out, the principles of Catholic Social Teaching (CST), and some of its analogs from other traditions, can also be fruitfully applied to ethical issues that arise in engineering and technology [6].

CST is established on a number of foundational principles that include human dignity, subsidiarity, solidarity and the common good, the universal destination of goods, charity, social justice, and stewardship and sustainability (see Table I). It has deep roots in the Bible and early Christian doctrine. The Fathers of the Church emphasized care for the poor, just economic practices, and the dignity of labor. In the nineteenth century, CST became more formally articulated with Pope Leo XIII's encyclical *Rerum Novarum* (1891), which addressed the rights of workers, the dangers of both unregulated capitalism and socialism, and the need for just wages [7]. Throughout the twentieth century, CST engaged with social and civil rights movements, with encyclicals like *Pacem in Terris* (1963) promoting human rights and *Laborem Exercens* (1981) affirming the dignity of work [8, 9]. In the modern era, CST continues to address environmental justice, e.g., *Laudato Si'* (2015), globalization, e.g., *Fratelli Tutti* (2020), and now digital ethics, advocating for policies that promote solidarity, subsidiarity, and the common good in an increasingly interconnected world [10, 11].

The philosophical tradition of natural law and natural rights, which undergirds CST, significantly influenced the founding of the United States by shaping its commitment to inalienable rights, human dignity, and the common good. The U.S. Declaration of Independence and Constitution reflect the natural law tradition, affirming that rights are endowed by the Creator and not merely granted by governments. CST more directly played a role in the development of the Universal Declaration of Human Rights (1948), particularly through the influence of the French Thomist philosopher, Jacques Maritain, who helped frame human rights as rooted in universal moral order rather than cultural or political consensus [12]. Pope Pius XI's 1931 articulation of the principle of subsidiarity helped the authors of the Maastricht Treaty of 1992 articulate the relationship between the European Union and its sovereign nation-states [13]. All these documents reflect the natural law principle that justice and rights are inherent to human nature, forming the basis for human dignity and social order [14].

Most recently, the Vatican's 2025 document, *Antiqua et Nova*, addresses the ethical and anthropological challenges posed by AI, emphasizing the need for responsible development that aligns with human dignity and the common good. It underscores that AI should complement rather than replace human intelligence, highlighting the importance of human oversight in critical areas such as healthcare, education, and warfare. The document warns against potential risks, including the erosion of truth through deepfakes and misinformation, and stresses the moral responsibility of developers and users to ensure AI serves humanity positively. Reflecting CST, it calls for collective efforts to guide AI development in ways that promote justice, protect human rights, and foster the integral development of all people [15].

CST emphasizes human dignity, respect, and community as foundational ethical principles that extend beyond religious contexts and are highly relevant to broader secular society and technology. Human dignity asserts that every person, made in the image of God, has intrinsic worth. It thus calls for shaping policies that protect individual rights and prevent exploitation in workplaces, healthcare, and cyberspace. Respect calls for fair treatment of all people, influencing areas like data privacy and bias in algorithms. Community underscores the importance of the common good, encouraging technology to serve humanity rather than further isolate individuals or exacerbate inequality. Together, these principles guide ethical decision-making in governance, business, and innovation, ensuring that technological advancements uplift rather than undermine human well-being.

CST places human dignity at the center of moral and social decision-making, affirming that every person has inherent

worth regardless of status or utility. This perspective offers a unique contribution to AI ethics by emphasizing that technology should serve humanity, not replace or diminish it. A key application of this is human-centered design, an approach that prioritizes user needs, accessibility, and ethical responsibility in engineering. This means designing systems that enhance human well-being rather than prioritizing efficiency or profit at the expense of humans. For example, in software engineering, human-centered design ensures that algorithms, data sources and collection approaches, and Large Language Model (LLM) training are transparent, fair, and aligned with ethical values, preventing bias, safeguarding privacy, and fostering trust. By aligning CST's focus on dignity with human-centered design, engineers have the opportunity to develop technology that truly benefits society while also respecting fundamental human rights.

III. Related Work

A. Philosophical and Theological Perspectives

Overviews of philosophical approaches to AI ethics can be found at [16, 17, 18]. Other works examine the issue in the light of specific philosophical approaches, including Kantian [19], dignitarian [20], classical Greek philosophy [21], value ethics [22], virtue ethics [23], honor ethics [24], and feminist political philosophy [25]. Boddington [26] offers a metaethical and metaphysical perspective. Other philosophy-based works look at specific situations or specific issues. For example, Borkovich et al. [27] consider workplace issues, with an emphasis on integrity. Karpouzis [21] and Haderlie [23] are among the many works that deal with education and pedagogy. And Diaz-Rodriguez et al. [28] treat the issue of hallucination, misrepresentation, and trustworthiness. Explicitly theological approaches are somewhat less common. Ahmed et al. [29] explore a multi-religious perspective, and Xu [30] and Oveido [31] also investigate AI using a religious or theological lens.

B. Pragmatic Ethical Perspectives

There are many works addressing the ethics of AI with a more pragmatic or "day-to-day" view. Many such references will be cast in terms of data science or data analytics, machine learning, or robotics and autonomous devices, rather than AI per se. We cannot hope to present a full overview, so we will rather look at the coverage of relevant issues. This subsection covers the problems of AI and the development process; the next deals with sector issues and social implications. Some cited works could fit into either category.

Vardi [32], a distinguished computer scientist and former editor of ACM Communications, writes that he is now far more concerned about the ethics of AI, in ways that resonate with many of the concerns of CST, than he was five years ago. Diakopoulos et al. [33] consider professional ethical responsibilities of AI in light of the ACM Code of Ethics and Professional Conduct [34], with particular application to journalism. The ethics of AI for practitioners and in the software development process is treated in [35, 36], and the concept of Responsible AI in [33, 37]. The regulatory issues are discussed in [28, 38], and the user perspective in [39].

Works discussing or reporting on the individual problems of AI applications include Banciu and Cimu [40] on data privacy, with a focus on compliance, Vosoughi et al. [41] on misinformation, specifically in the news, and Mouschoutzi [42] on three cognitive biases in ChatGPT. Legal issues are treated in Appel et al. [43], discussing generative AI with respect to intellectual property and copyright infringement, Paterson [44] on deepfakes from both a social and a legal perspective, and Sebastien [45], with a careful look at consent, and its interaction with trustworthiness and explainability. Finally, Vainio-Pekka et al. [46] discuss the importance of explainability in AI applications.

Robotics and autonomous devices have different ethical challenges, many having to do with the interaction with humans and with infrastructure. Vemoori [47] presents a survey of ethical issues for autonomous vehicles. Coskun [48] looks at the overuse of power and water needed for complex algorithms and very large databases, a serious social and environmental concern. These are discussed further in Section V.

C. Social and Sector Perspectives

There are also works with a greater resonance with the principles of CST, addressing specific social, environmental, and moral issues. There is however great overlap with the technical analysis and citations of work on bias, misinformation, property rights, and consent included above.

Gautam [49] surveys the social impacts of AI. Chi et al. [50] look at the interaction of AI with diversity and inclusion. Howard and Borenstein [51] and Sharma and Mallory [52] discuss the social effects of bias in AI. Irving [53] discusses social media manipulation and disinformation, and Christensen et al. [54] the effects of AI hallucination. Tangentially related to misinformation, He [55] looks at the ethics of marketing with generative AI. Finally, Alves et al. [56] discuss the social implications of AI in the Internet of Things (IoT), with distinct echoes of the CST principle of solidarity and the CST inspired economic theory of distributism.

Seemingly all sectors of life, learning, and the economy are affected by AI. Holmes and Tuomi [57] and Siau and Wang [58] are among many looking at AI ethics and education and academia, with Bowden [59] reporting on the misuse of AI in education, and Gray [60] presenting a rights framework for AI research. Char et al. [61] look at ethical challenges in implementing machine learning in healthcare, and Khattar [62] and Brougham and Haar [63]. The interaction with labor and workers' rights is studied by Vakkuri et al. [64]. Shakilla and Saputro [65] examine AI ethics in the context of industry and leadership.

IV. Core Challenges, Principles, and guidelines

Computer science and engineering, data science, and now deep learning models built on complex neural networks and other developments in AI touch nearly all aspects of human life—from the arts and economy to infrastructure, medicine, and beyond—interacting and reinforcing one another. Their direct impact is profound, but their influence on education and workforce training will be even greater, shaping future generations' skills, opportunities, and ethical responsibilities.

Professionals in science, medicine, and engineering are often guided to apply various ethical frameworks—such as teleological, deontological, and virtue ethics—to navigate moral challenges in their fields. Teleological ethics, derived from the Greek word *telos* meaning "end" or "goal," assesses the morality of actions based on their outcomes. This approach determines whether an action is right or wrong by evaluating the consequences it produces. For instance, one might consider how implementing AI in a specific area would ultimately impact human safety. Deontological ethics, rooted in the Greek term *deon* meaning "duty," judges actions based on their inherent morality, independent of the results they yield. This perspective emphasizes the importance of following moral rules or duties. An example question might be: What are my responsibilities to clients, patients, or the public when developing or using a particular AI system? Virtue ethics, originating from the Latin word *virtus* meaning "excellence of character," is considered a form of teleological ethics. It focuses on the character traits that a virtuous person would exhibit in a given situation. For example, when facing a challenge related to healthcare privacy and data analytics, one might ask: How would a person of the highest moral character respond in this scenario? These ethical frameworks provide structured approaches for professionals to evaluate and address moral dilemmas within their respective disciplines [66, 67, 68].

A. Core Ethical Challenges

One ethical challenge is the reduction of persons to data. AI systems often reduce human experiences, decisions, and relationships to data points and algorithms, treating people as objects of analysis rather than subjects of moral worth. AI is increasingly used in hiring, healthcare, policing, and social services, raising concerns about impersonal and biased decision-making. Many AI models reflect the biases of the data they are trained on, thus potentially exacerbating unjust discrimination. When AI makes life-altering decisions, such as who receives medical treatment or financial aid, it risks treating people as mere problems to be solved, rather than as persons with inherent dignity. CST offers a helpful corrective by teaching that each human being has inherent dignity and should never be treated merely as a means to an end [69]. Furthermore, it demands a preferential option for the poor and vulnerable, ensuring that AI does not reinforce injustice and exclude marginalized communities.

Advances in AI-driven automation or in generative AI could potentially threaten human labor, a key aspect of human dignity [70]. If AI replaces human workers without ensuring just compensation, retraining, or meaningful work opportunities, it could violate the right to dignified labor. In some countries AI is used in mass surveillance, tracking, and social scoring, raising serious privacy concerns [71]. CST on the other hand emphasizes the right to privacy and personal autonomy—a surveillance state undermines the dignity of individuals by treating them as potential threats rather than as persons deserving respect.

Another serious issue emerges around moral responsibility and accountability. When AI causes harm from autonomous weapon systems, deepfake deception, or biased sentencing algorithms, who is morally and legally responsible—the developer, the user, or the machine? CST insists that true moral responsibility belongs to persons, and AI should never be given moral autonomy in a way that displaces human accountability. AI offers great potential for caregiving and encouraging interaction with isolated people, e.g., in healthcare facilities. However, AI therapists and robotic caregivers can create illusory relationships. CST values authentic human relationships, and AI should enhance, not replace, human connection, especially in care for the elderly, children, and the vulnerable.

The transhumanist movement advocates for enhancing human abilities through technology, such as AI, genetic modification, and cybernetic augmentation, ultimately aiming to surpass natural human limitations and even achieve immortality [72]. Transhumanism is problematic from a CST perspective because it often rejects the inherent dignity and natural limits of the human person as created by God. CST upholds that human dignity is intrinsic, not dependent on intelligence, physical ability, or technological enhancement. Additionally, transhumanism's pursuit of radical self-determination and control over human nature contradicts CST's emphasis on humility, the common good, and the moral order of creation. This raises concerns about inequality, loss of authentic human relationships, and ethical misuse of technology. From the CST perspective, AI must remain at the service of human dignity and the common good. The Church urges caution against AI systems that treat people as data points, undermine the dignity of work, entrench discrimination, or replace human relationships. AI must be ethically designed, justly implemented, and always subordinate to the moral responsibility of human beings. AI must serve humanity, not the other way around.

B. Principles from CST and Secular Ethics

CST's principles align very well with secular ethical values like fairness, transparency, and accountability. For example, solidarity is the commitment to the well-being of all, particularly the most vulnerable, fostering unity and social responsibility. In the context of AI, solidarity aligns with fairness, ensuring that AI systems are designed to promote

equity, reduce biases, and prevent discrimination. Both perspectives advocate for AI that serves humanity rather than exacerbates existing inequalities. This can be seen in efforts to mitigate algorithmic bias, promote inclusive datasets, and ensure broad access to AI benefits.

Subsidiarity in CST emphasizes decision-making at the most appropriate level, empowering local and individual agencies while avoiding over-centralization. In AI governance, this principle corresponds with transparency, ensuring that AI decision-making processes are understandable, explainable, and accessible. Subsidiarity suggests that AI systems should not concentrate power in the hands of a few but rather provide users with knowledge and control over how AI influences their lives. This aligns with calls for open-source AI, explainable AI (XAI), and policies that promote participatory governance in AI development.

The common good refers to creating conditions where all people can thrive, emphasizing shared responsibility for societal well-being. In AI ethics, this principle is reflected in accountability, ensuring that AI developers, corporations, and policymakers are responsible for the impact of AI on individuals and society. Mechanisms such as ethical AI audits, regulatory oversight, and human-in-the-loop systems help ensure that AI serves the common good rather than harming communities through unchecked automation or surveillance.

By integrating CST principles with secular AI ethics, both religious and non-religious stakeholders can find common ground in advocating for AI that is just, inclusive, and oriented toward human flourishing. This shared ethical perspective can shape policy discussions, corporate AI strategies, and global governance efforts, ensuring that AI technologies reflect fundamental values that protect human dignity and the social fabric.

V. Case Studies and Implementation Issues

In this section, we analyze seven case studies from a CST perspective, illustrating how the CST principles outlined above align with fairness, transparency, and accountability in AI governance and development, and, in particular, whether human dignity was upheld, on the one hand, or compromised, on the other, or if other CST principles were violated by systems relying on AI. In the latter case, we suggest how applying CST principles could tend to lead to better outcomes.

A. Misinformation on Social Media

An important case involves Facebook's AI and election misinformation. Here we consider the secular value of accountability and the CST principle of promoting the common good. Facebook (now Meta) previously employed AI algorithms (e.g., social bots) as well as third-party organizations to curate news feeds as well as detect misinformation. However, studies quickly revealed that its AI-driven engagement model amplified false information during the 2016 U.S. election and others [41]. When AI algorithms prioritize engagement over accuracy, they facilitate the spread of misinformation, eroding public trust and hindering informed decision-making. This lack of accountability in information dissemination can lead to misinformed voting choices and a weakened democratic system. CST's notion of the common good indicates that AI should be designed to serve society's well-being rather than profit-driven manipulation that threatens democratic processes. From the perspective of accountability, AI developers must take responsibility for unintended consequences, ensuring AI systems do not spread misinformation or harm social cohesion.

In addition to utilizing AI, Meta hired third-party fact-checkers such as PolitiFact in 2016. However, a study conducted earlier in 2011 by Eric Ostermeier, a political science professor at the University of Minnesota, found that PolitiFact assigned "false" or "pants on fire" ratings to Republican statements over three times more often than to Democratic ones [73]. This created further difficulties for Meta. Third-party fact-checking has been shown not to be a panacea when it comes to increasing media trust [74]. Thus, in January 2025, the company announced a shift to a "Community Notes" model, allowing users on platforms like Facebook, Instagram, and Threads to identify and provide context for potentially misleading posts. This change empowers the broader user community to monitor and address misinformation, moving away from reliance on AI bots which were amplifying false information, and official fact-checking organizations which brought their own political bias.

From a CST perspective, the design and use of AI in social media must align with the principles of the common good, ensuring that technologies promote societal well-being rather than exacerbate polarization or undermine democracy. Developers and platforms must be held accountable for the unintended consequences of their systems, prioritizing transparency, fairness, and the protection of free speech. Ethical responsibility lies in ensuring that AI fosters informed decision-making and social cohesion, rather than manipulation or the erosion of public trust. In CST the right to the truth is rooted in the dignity of the person. People have the right to access accurate and truthful information to make informed decisions that contribute to their well-being and participation in society. This right underpins the need for transparency, accountability, and ethical communication, ensuring that misinformation and manipulation do not harm individuals or communities, especially in a democratic context. Truth serves as a foundation for justice, freedom, and the common good.

B. Hiring Algorithms

In 2018, Amazon discontinued its AI recruiting tool after discovering it discriminated against women [75]. The system, trained on resumes submitted over a decade, learned to favor male candidates, penalizing resumes that included the word "women's" or referenced women's colleges. This incident highlighted the ethical risks of using historical data that reflect societal biases in AI decision-making. It underscores the ethical risks of embedding and perpetuating biases through AI, particularly in critical areas like employment. From a CST

perspective, this incident raises concerns related to human dignity, the common good, and the preferential option for the marginalized. The dignity of every person, a foundational principle of CST, was compromised when women were systematically disadvantaged by the AI system. By reinforcing gender bias, the tool denied women equal opportunities, violating principles of justice and fairness. Applying CST to this issue calls for proactive measures to correct biases in AI decision-making. This includes creating diverse and inclusive training datasets, implementing ethical oversight mechanisms, and prioritizing the needs of marginalized communities. True technological progress must align with moral responsibility, ensuring that innovation serves humanity rather than deepening social divides. Similar observations and conclusions apply for situations reported by Hoffman et al. [76], reporting how AI generates or reinforces covertly racist perceptions in hiring and other recommendations.

C. Facial Recognition Software

Next, we consider San Francisco's Ban on Facial Recognition in 2019 [77]. San Francisco became the first U.S. city to ban the use of facial recognition technology by local agencies, citing concerns over surveillance, racial bias, and civil liberties. Critics argued that facial recognition disproportionately misidentifies people of color and infringes on privacy rights. According to the principle of subsidiarity, decision-making should occur at the most local, appropriate level. Communities should have the right to regulate technology that impacts their citizens rather than having policies imposed by higher authorities. From the perspective of transparency, citizens deserve to know how AI-based surveillance systems operate, what data they collect, and how decisions are made. San Francisco's decision inspired similar policies in other cities, demonstrating how subsidiarity can empower local communities in AI governance. The debate reinforced the need for transparent AI policies, ensuring public awareness and informed consent regarding surveillance technology.

D. Two HealthCare Examples

Here we present two examples, the first in which human dignity is upheld, and the second in which it is compromised.

In the first example, Google's DeepMind developed an AI system capable of predicting acute kidney injury (AKI) up to 48 hours before it is typically detected by prior methods [78]. By analyzing a comprehensive dataset of electronic health records, the AI accurately identified over 90% of patients whose condition would deteriorate to the point of requiring dialysis. This early detection allows for timely interventions, improving health outcomes. Here, human dignity is preserved, and fairness is exhibited. The dignity of the patient is upheld, since patient suffering and long-term consequences are reduced, and fair access to life-saving care is ensured. Further, even if the AI is replacing some human healthcare professionals, it is not (entirely) to save money, but because it substantially improves outcomes. In this case, CST is satisfied for the patient, but human dignity, as well as subsidiarity, may suggest that human diagnosticians remain in the process, allowing human judgment to interpret and complement the AI results, and to interact face-to-face with the patient.

In the second case, IBM Watson for Oncology was designed to assist doctors in diagnosing and recommending cancer treatments [79]. However, reports surfaced that the system provided incorrect and potentially unsafe recommendations due to biased training data, which largely came from a single hospital in the U.S. rather than reflecting diverse global populations. The AI application used an unfair algorithm that compromised solidarity and fairness. AI should prioritize the needs of all, especially marginalized groups. In healthcare, bias in AI can disproportionately harm underrepresented communities. Thus, developers must ensure that AI models are trained on diverse, representative datasets to avoid systemic bias and unfair treatment. Greater awareness of bias in AI training data has led to initiatives promoting ethical AI development, such as Google's PAIR (People + AI Research) [80] and Microsoft's AI Fairness Checklist [81].

E. Ethics of Autonomous Vehicles

Autonomous vehicles (AV)–mostly taxis and other service vehicles–are becoming ever more common, argue their benefits [82], while reports of problems, particularly accidents, are also common [83]. There are interesting ethical tradeoffs here. Human dignity is enhanced for the disabled and others who may be unable to drive (although some would strongly prefer to interact with a human driver). On the other hand, the common good and perhaps stewardship are compromised by accidents and other safety problems, and by violations of traffic laws and customs. Autonomous vehicles present a complex ethical landscape when viewed through the lens of CST. The increased accessibility and independence afforded to those with disabilities align with CST's emphasis on the inherent worth of every person and the importance of removing barriers that prevent full participation in society. At the same time, concerns about safety, ethical decision-making in emergencies, and accountability highlight tensions with the principles of the common good and stewardship. The common good requires that technological advancements serve society as a whole, not just a privileged few.

Reports of AV-related accidents, system failures, and violations of traffic laws suggest that, while beneficial for some, AVs may introduce new risks that endanger others. The ethical principle of stewardship also calls for responsible innovation—ensuring that new technologies do not create harm, particularly when lives are at stake. Companies developing AVs must prioritize not only efficiency and profitability but also the protection of human life and societal well-being. Moreover, CST's emphasis on subsidiarity suggests that decisions about AV regulation and deployment should be made at appropriate levels, involving local communities, policymakers, and industry leaders in ethical deliberation. Transparency in AV development, clear liability structures for accidents, and

inclusive decision-making are crucial to ensuring that these technologies serve all people justly. Ultimately, while AVs hold great promise in promoting accessibility and human dignity, their implementation must be guided by a moral framework that prioritizes the common good, safety, and responsible stewardship. Ethical oversight, rigorous safety standards, and an ongoing commitment to justice should shape the future of this technology to ensure it truly serves humanity.

F. A Resource Issue

As the use of large databases for predictive and generative AI increases dramatically, the use of power and water increases even more quickly [84]. Generative (and predictive) AI uses massive amounts of energy for computation and data storage and gallons of water to cool the equipment at data centers. Excessive use could have a negative impact on climate as well as deny other customers (and even public safety and emergency response teams) needed electric power resources.

This clearly raises issues of stewardship and sustainability, as well as solidarity and the common good. Responses should include research to lower resource demands, and legislative and regulatory oversight and accountability, including local zoning and other decisions. The first may be foreshadowed by DeepSeek, which on the other hand has its own common good problems with completeness and reliability [85], among others. Also, the first steps toward the latter are underway in the US and the EU [84]. These case studies demonstrate that CST principles and secular AI ethics converge on key issues such as fairness, transparency, and accountability. By integrating both moral perspectives and technical safeguards, AI development prioritizes human dignity, social justice, and responsible innovation.

VI. Practical Guidance for Development and Use of Artificial Intelligence

We note a number of similarities between the principles and priorities of CST and practical guidance from UNESCO [86], IEEE [87], and IBM [88]. See Table II. UNESCO's ethical recommendations and CST share a strong commitment to human dignity, social responsibility, and the common good. Both emphasize the importance of fairness, accountability, and sustainability in guiding ethical decision-making. Each advocates for the protection of human rights, the promotion of peace, and the responsible stewardship of the environment. Additionally, both frameworks stress the need for transparency, awareness, and collaboration to ensure justice and equity in society. Ultimately, UNESCO and CST align in their dedication to ethical principles prioritizing human well-being and the betterment of society as a whole.

CST and the IEEE Standards Association Global Initiative on the Ethics of Autonomous and Intelligent Systems share a commitment to human dignity, justice, and the responsible use of technology for the common good. Both emphasize human rights, well-being, and accountability, ensuring that technological advancements serve society rather than harm it.

TABLE II. Three ethical guidelines for AI (reformatted from [89] based on [86, 87, 88])

UNESCO	**IEEE**	**IBM**
UNESCO ethical recommendations are based on specific core values such as human dignity and rights, promoting peace, and care for the environment. Based on these values, UNESCO specifies ten principles: 1. Proportionality and Do No Harm 2. Safety and Security 3. Right to Privacy and Data Protection 4. Multistakeholder and Adaptive Governance & Collaboration 5. Responsibility and Accountability 6. Transparency and Explainability 7. Human Oversight and Determination 8. Sustainability 9. Awareness and Literacy 10. Fairness and Non-discrimination	The IEEE Standards Association (SA) has established a Global Initiative on the Ethics of Autonomous and Intelligent Systems. The IEEE approach is established on eight fundamental principles: 1. Human Rights 2. Well-being 3. Data Agency 4. Effectiveness 5. Transparency 6. Accountability 7. Awareness of Misuse, and 8. Competence	IBM proposes three guiding values for AI: 1. The purpose of AI is to augment human intelligence, 2. Data and insights belong to their creator, and 3. Technology must be transparent and explainable. Leveraging insights from the 1979 Belmont Report, IBM defines three overarching principles for AI: 1. Respect for persons 2. Beneficence, and 3. Justice, i.e., burdens and benefits may be distributed following one or more of: a. Equal share b. Individual need c. Individual effort d. Individual contribution e. Societal contribution f. Merit

The IEEE's focus on transparency and awareness of misuse aligns with CST's call for ethical responsibility and the protection of the vulnerable. Additionally, both frameworks stress the importance of competence and informed decision-making to prevent harm and promote fairness. Ultimately, CST and IEEE share a vision of ethical stewardship, ensuring that innovation respects human dignity and contributes to a just and sustainable society.

CST and IBM's approach to ethical AI have many similarities. Both emphasize that technology should serve and uplift humanity, aligning with CST's principle of the common good. IBM's focus on respect for people reflects CST's commitment to human dignity, while beneficence aligns with CST's advocacy for fairness in social and economic structures. Additionally, IBM's commitment to transparency and accountability mirrors CST's emphasis on ethical responsibility and the need for informed decision-making. Both frameworks seek to ensure that technology promotes human flourishing, fairness, and social well-being.

Engineers developing AI should therefore integrate the ethical principles of CST and the above guidelines into their daily work by implementing the following concrete actions. First, they should conduct regular impact assessments to identify potential risks related to fairness, transparency, and human rights. For example, before deploying an AI model, engineers should check for biases in training data and in beta-test results and ensure diverse representation. Second, they must prioritize user privacy by designing systems that provide clear data ownership and consent options, following best practices in encryption and data protection. Third, transparency can be improved by making AI decision-making processes explainable, such as providing users with understandable reasons for automated decisions in hiring, lending, or medical diagnosis. Further, engineers should establish accountability measures, such as ethical review boards or internal audits, to ensure compliance with justice and sustainability standards. Finally, they should actively engage with diverse stakeholders, including ethicists, policymakers, and affected communities, to align technological developments with the common good. By embedding these practices into their workflows, engineers can create AI that upholds human dignity, promotes social justice, and fosters long-term societal well-being. Researchers and practitioners exploring, developing, or employing AI technology, enterprises and others deploying or leveraging AI applications, regulators and legislators, and social scientists and evaluators of AI applications and tools, or of products using them, should also be aware of and follow these guidelines, as appropriate to their roles, contexts, and situations.

VII. Conclusions and Future Work

Catholic Social Teaching provides a useful perspective from which to examine the ethics of AI and formulate both conceptual and practical guidelines. In this paper, we have examined that perspective in depth, related it to other approaches to the issue, showing where it differs or enriches and refines, and argued that it remains accessible to a secular audience, while offering a rigorous and consistent framework that can be useful to all.

In future work, we would like to extend this approach to more of STEM, looking at the disciplines themselves as well as their interaction with AI. It seems that overall, much of the work strictly in the formal sciences (e.g., pure mathematics, logic, theoretical computer science) needs minimal attention. Challenges arise in the application of these fields (e.g., data science) in an engineering context (e.g., software engineering). Numerous ethical issues also arise in the life sciences and healthcare. We would also like to extend the comparison to other ethical approaches. In [6], we looked briefly at the social teaching of other religious traditions. In the future, we would like to compare CST more thoroughly with other religious and philosophical approaches, including those mentioned above, with respect to AI and the issues considered above.

Acknowledgments

The authors would like to thank Lisa Rose-Wiles and Father Pawel Tomczyk for fruitful discussions on these topics, and to Carol Marlowe for finding valuable information. Cyril S. Ku's research for this paper was supported in part by the National Science Foundation under NSF Grant # 2028011.

References

[1] N. Wiener, *Cybernetics, or Control and Communication in the Animal and the Machine*, 2nd ed. Cambridge: The MIT Press, 1965.
[2] A. M. Turing, "Computing Machinery and Intelligence," *Mind*, vol. 59, no. 236, pp. 433–460, Oct. 1950.
[3] Pontifical Council for Justice and Peace, *Compendium of the Social Doctrine of the Catholic Church*. Vatican City: Libreria Editrice Vaticana, 2004.
[4] T. Aquinas, *Commentary on Aristotle's Nicomachean Ethics*, C.I. Litzinger, Trans. Notre Dame, IN: Dumb Ox Publications, 1993.
[5] *Catechism of the Catholic Church*, Part III, Chapter 1, Article 4, "The Dignity of the Human Person." Available: https://www.vatican.va/archive/ENG0015/_INDEX.HTM.
[6] J. Anderson, J.R. Laracy, and T. Marlowe, "A Multi-Disciplinary Analysis of Catholic Social Teaching with Implications for Engineering and Technology," *J. Systemics, Cybernetics, and Informatics*, vol. 18, no. 6, pp. 41–49, 2020. Available: https://www.iiisci.org/journal/pdv/sci/pdfs/SA244NH20.pdf.
[7] Leo XIII, *Rerum Novarum*. Vatican City: Holy See, 1891.
[8] John. XXIII, *Pacem in Terris*. Vatican City: Holy See, 1963.
[9] John Paul II, *Laborem Exercens*. Vatican City: Holy See, 1981.
[10] Francis, *Laudato Si*'. Vatican City: Holy See, 2015.
[11] Francis, *Fratelli Tutti*. Vatican City: Holy See, 2020.
[12] United Nations General Assembly, "Universal Declaration of Human Rights," A/RES/217(III), Dec. 10, 1948. Available: https://www.un.org/en/about-us/universal-declaration-of-human-rights.
[13] Pius XI, *Quadragesimo Anno*, Vatican City: Holy See, 1931.
[14] M. A. Glendon, "The Influence of Catholic Social Doctrine on Human Rights," Pontifical Academy of Social Sciences, Acta 15, 2009. Available: https://www.pass.va/content/dam/casinapioiv/pass/pdf-volumi/acta/acta_15/acta15-glendon.pdf.
[15] Dicastery for the Doctrine of the Faith and Dicastery for Culture and Education, *Antiqua et nova: Note on the Relationship Between Artificial Intelligence and Human Intelligence*, Vatican City: Holy See, Jan. 28, 2025. Available: https://www.vatican.va/roman_curia/congregations/cfaith/documents/rc_ddf_doc_20250128_antiqua-et-nova_en.html.

[16] V. C. Müller, "Ethics of AI and Robotics," *Stanford Encyclopedia of Philosophy, Stanford University*, April 30, 2020. Available: https://plato.stanford.edu/entries/ethics-ai/

[17] "Ethics of Artificial Intelligence," Internet Encyclopedia of Philosophy. (no date). Available: https://iep.utm.edu/ethics-of-artificial-intelligence/

[18] E. Prem, "From ethical AI frameworks to tools: a review of approaches," *AI Ethics 3*, pp. 699–716, 2023.

[19] L. Singh, "Automated Kantian Ethics." In *Proc. 2022 AAAI/ACM Conf. on AI, Ethics, and Society (AIES '22)*. Oxford, England, May 2021.

[20] R. Hanna, and E. Kazim, "Philosophical foundations for digital ethics and AI Ethics: a dignitarian approach," *AI Ethics 1*, pp. 405–423, 2021.

[21] K. Karpouzis, "Artificial Intelligence in Education: Ethical Considerations and Insights from Ancient Greek Philosophy". In *Proceedings of the 13th Hellenic Conference on Artificial Intelligence (SETN '24)*, Article 57, pp. 1-7, Piraeus, Greece, September 2024.

[22] J. Simon, G. Rieder, and J. Branford, "The Philosophy and Ethics of AI: Conceptual, Empirical, and Technological Investigations into Values," *CEPE/IACAP 2021: Introduction to Topical Collection*, Digital Society. vol. 3, February 2024.

[23] C. Haderlie, "Generative AI Policy: A Pedagogical Communication Tool and Virtue Ethics Gateway." In *Proc. 42nd ACM Intl. Conf. on Design of Communication (SIGDOC '24)*, Arlington, VA, USA, October 2024.

[24] S. T.-I. Wu, D. Demetriou, and R. A. Husain, "Honor Ethics: The Challenge of Globalizing Value Alignment in AI." In *Proc 2023 ACM Conf on Fairness, Accountability, and Transparency (FAccT '23)*, Chicago, IL, USA, October 2023.

[25] A. Kasirzadeh, "Algorithmic Fairness and Structural Injustice: Insights from Feminist Political Philosophy." In *Proc 2022 AAAI/ACM Conference on AI, Ethics, and Society (AIES '22)*, San Jose, CA, USA, October 2024.

[26] P. Boddington, "Philosophy for AI Ethics: Metaethics, Metaphysics, and More." In: *AI Ethics. Artificial Intelligence: Foundations, Theory, and Algorithms.* Springer, Singapore, 2023.

[27] D. J. Borkovich, R. J. Skovira, and F. Kohun. "Artificial intelligence in the workplace: A philosophical approach to ethics and integrity," *Issues in Info. Syst.*, vol. 25, no. 3, pp. 200-217, 2024.

[28] N. Díaz-Rodríguez, J. Del Ser, M. Coeckelbergh, M. López de Prado, E. Herrera-Viedma, and F. Herrera, "Connecting the dots in trustworthy Artificial Intelligence: From AI principles, ethics, and key requirements to responsible AI systems and regulation," *Info. Fusion*, vol. 99, 2023.

[29] S. Ahmed, A.A. Sumi, and N. Aziz, "Exploring Multi-Religious Perspective of Artificial Intelligence," *Theology and Science*, vol. 23, no. 1, pp. 1–25, .2024.

[30] X. Xu, "Understanding AI from a Theological Perspective," *The University of Edinburgh, Edinburgh Futures Institute*, August 3, 2023. Available: https://efi.ed.ac.uk/understanding-ai-from-a-theological-perspective/

[31] L. Oviedo, "Artificial Intelligence and Theology: Looking for a Positive—But Not Uncritical—Reception,". *Zygon J. Religion and Science*, vol. 57, no. 4, pp. 938-952, December 2022.

[32] M. Vardi, "I Was Wrong about the Ethics Crisis: Is Big Tech supporting the public good, and if not, what should Big Tech workers do about it?", *Commun. ACM*, vol. 68, no. 1, p. 5, December 18, 2024.

[33] N. Diakopoulos, C. T., Dietmar J., I. Costera Meijer, and E. Motta, "Leveraging Professional Ethics for Responsible AI: Applying AI techniques to journalism," *Commun. ACM*, February 23, 2024.

[34] Association for Computing Machinery. ACM Code of Ethics and Professional Conduct, 2018. Available: https://www.acm.org/code-of-ethics

[35] A. Pant, R. Hoda, S. V. Spiegler, C. Tantithamthavorn, and B. Turhan, "Ethics in the Age of AI: An Analysis of AI Practitioners' Awareness and Challenges," *ACM Trans. Softw. Eng. Methodol*. vol, 33, no. 3, Article 80, March 2024.

[36] J. Gogoll, N. Zuber, S. Kacianka, et al., "Ethics in the Software Development Process: from Codes of Conduct to Ethical Deliberation," *Philos. Technol.*, vol. 34, pp. 1085–1108, 2021.

[37] J. Woodgate, and N. Ajmeri, "Macro Ethics Principles for Responsible AI Systems: Taxonomy and Directions," *ACM Comput. Surv.*, vol, 56, no. 11, Article 289, November 2024.

[38] L. Singh, "AI regulation: Balancing innovation and accountability," *Frontier Enterprise*, April 15, 2024. Available: https://www.frontier-enterprise.com/ai-regulation-balancing-innovation-and-accountability/

[39] M. Ali, "AI Ethics for the Everyday User – Why Should You Care? A beginner's guide to understanding the importance of ethics in artificial intelligence," *Towards Data Science*, Jan 29, 2025. Available: https://towardsdatascience.com/another-article-on-ai-ethics-but-this-time-for-the-masses-c809e09e35e7/

[40] D. Banciu, and C. E. Cîrnu, "AI Ethics and Data Privacy Compliance," *2022 14th Intl. Conf. on Electronics, Computers and Artificial Intelligence (ECAI)*, Ploiesti, Romania, pp. 1-5, 2022.

[41] S. Vosoughi, D. Roy, and S. Aral, "The spread of true and false news online," *Nat. Commun.*, vol. 9, no. 1, pp. 1-7, 2018.

[42] M. Mouschoutzi. "Exploring 3 Common Cognitive Biases with ChatGPT." The Conversation. Available: https://theconversation.com/building-fairness-into-ai-is-crucial-and-hard-to-get-right-220271

[43] G. Appel, J. Neelbauer, and D. A. Schweidel, "Generative AI Has an Intellectual Property Problem," *Harvard Bus. Review*, April 07, 2023.

[44] J. M. Paterson, "AI Deepfakes on the Web: The 'Wicked' Challenges for AI Ethics, Law and Technology," In *Proc. ACM Web Conf. 2024 (WWW '24)*, Singapore, May 2024.

[45] N. Sebastien, "The AI Consent Conundrum: Do We Truly Understand What We Agree To?" The Medium, May 22, 2024. Available: https://medium.com/@neriasebastien/the-ai-consent-conundrum-do-we-truly-understand-what-we-agree-to-295201df823e

[46] H. Vainio-Pekka, M. Ori-Otse Agbese, M. Jantunen, Ville Vakkuri, T. Mikkonen, R. Rousi, and P. Abrahamsson, "The Role of Explainable AI in the Research Field of AI Ethics," *ACM Trans. Interact. Intell. Syst*. vol. 13, no. 4, Article 26, December 2023.

[47] V. Vemoori and Forbes Technology Council. "Navigating The Ethical Dilemmas Of Self-Driving Cars: Who Decides When Safety Is At Risk?" *Forbes*, October 23, 2024. Available: https://www.forbes.com/councils/forbestechcouncil/2024/10/23/navigating-the-ethical-dilemmas-of-self-driving-cars-who-decides-when-safety-is-at-risk/

[48] A. Coskun, "AI supercharges data center energy use – straining the grid and slowing sustainability efforts," The Conversation. Available: https://theconversation.com/ai-supercharges-data-center-energy-use-straining-the-grid-and-slowing-sustainability-efforts-232697

[49] N. Gautam. "Unraveling the Social Impacts of Artificial Intelligence," UC Davis Office of Research, 2024. Available: https://research.ucdavis.edu/unraveling-the-social-impacts-of-artificial-intelligence/

[50] N. Chi, E. Lurie, and D. K. Mulligan, "Reconfiguring Diversity and Inclusion for AI Ethics." In *Proc. 2021 AAAI/ACM Conf. on AI, Ethics, and Society (AIES '21)*, pp. 447-457, 2021.

[51] A. Howard and J. Borenstein, "The Ugly Truth About Ourselves and Our Robot Creations: The Problem of Bias and Social Inequity," *Sci Eng Ethics*, vol. 24, pp. 1521–1536, 2018.

[52] S. Sharma and V. Mallory. S. Sharma and V. Mallory, "Social Bias in AI and its Implications," *NASA/TM–20210010446, NASA*, March 2021. . Available: https://ntrs.nasa.gov/api/citations/20210010446/downloads/NASA-TM-20210010446.pdf

[53] D. Irving, "Social Media Manipulation in the Era of AI. Rand Corporation, Aug 29, 2024. Available: https://www.rand.org/pubs/articles/2024/social-media-manipulation-in-the-era-of-ai.html

[54] J. Christensen, J. M. Hansen, and P. Wilson. "Understanding the role and impact of Generative Artificial Intelligence (AI) hallucination within consumers' tourism decision-making processes," Taylor and Francis Online, Jan 17, 2024. Available: https://www.tandfonline.com/doi/abs/10.1080/13683500.2023.2300032

[55] F. He, M. Du, A. Filos-Ratsikas, L. Cheng, C. Song, M. Lin, and J. Vines. "AI Driven Online Advertising: Market Design, Generative AI, and Ethics," In *Companion Proc. ACM Web Conf. 2024 (WWW '24*), pp. 1407-1409.

[56] A. M. Alves, E. L. Chuma, G. C. Vaz, G. G. De Oliveira, L. L. B. Roger and Y. Lano, "AI Ethics in IoT: Involving Society in the Discussion,"

2024 Intl. Conf. on Electrical, Computer and Energy Technologies (ICECET), Sydney, Australia, 2024, pp. 1-4, Sydney, Australia, 2024,
[57] W. Holmes, and I. Tuomi, "State of the Art and Practice in AI in Education,". *European J. Education*, vol 57, no. 4.
[58] K. Siau, and W. Wang, "Artificial Intelligence (AI) Ethics: Ethics of AI and Ethical AI." *J. Database Manage. (JDM)*, vol. 31, no. 2, pp. 74-87,
[59] J. Bowden, 7 Examples of AI Misuse in Education. Inspera, March 2024. Available: https://www.inspera.com/ai/examples-of-ai-misuse-in-education/
[60] M. Gray. "A Human Rights Framework for AI Research Worthy of Public Trust," *Issues in Science and Technology*, 2024. Available: https://api.semanticscholar.org/CorpusID:269887827
[61] D. S. Char, N.H. Shah, and D. Magnus, "Implementing Machine Learning in Health Care—Addressing Ethical Challenges," *The New England J. Medicine*, vol. 378 no. 11, pp. 981-983. 2018.
[62] R. Khattar, "Will AI Benefit or Harm Workers? Center for American Progress," Aug 24, 2023. Available: https://www.americanproEgress.org/article/will-ai-benefit-or-harm-workers/
[63] D. Brougham, and J. Haar, "Smart Technology, Artificial Intelligence, Robotics, and Algorithms (STARA): Employees' Perceptions of Our Future Workplace," *J. Management & Organization*, vol. 24, no. 2, pp. 239-257, 2017.
[64] V. Vakkuri, K.-K. Kemell, and P. Abrahamsson, "AI Ethics in Industry: A Research Framework", Preprint. Available: arXiv:1910.12695. https://doi.org/10.48550/arXiv.1910.12695
[65] U. Shakilla, and P. E. Saputro, "Revolutionizing Management: The Role of AI and Technology in Modern Leadership Practices." *Solo Int. Collab. and Publication of Soc. Sci. and Humanities*, vol, 3, no. 1, pp. 53–68, 2024.
[66] C.E. Harris, M.S. Pritchard, M.J. Rabins, R.W. James, and E.E. Englehardt, *Engineering Ethics: Concepts and Cases*, 6th ed. Boston, MA: Cengage Learning, 2018.
[67] K.D. O'Rourke and P.J. Boyle, *Medical Ethics: Sources of Catholic Teachings*, 4th ed. Washington, DC: Georgetown Univ. Press, 2011.
[68] P.C. Hebert and W. Rosen, *Doing Right: A Practical Guide to Ethics for Medical Trainees and Physicians*, 4th ed. Oxford, UK: Oxford University Press, 2019.
[69] Second Vatican Council, *Gaudium et Spes*, Vatican City: Holy See, 1965, sec. 27.
[70] John Paul II, *Laborem Exercens*, Vatican City: Holy See, 1981, sec. 9.
[71] E. Wishnick, "Digital authoritarianism and technological cooperation in Sino-Russian relations," in *China-Russia Relations in Central Asia*, Palgrave Macmillan, 2022, pp. 185-207.
[72] L. Terec-Vlad, "What about eternal life? A transhumanist perspective," *Postmodern Openings*, vol. 6, no. 2, pp. 33-41, Dec. 2015.
[73] E. Ostermeier, "Selection Bias? PolitiFact Rates Republican Statements as False at 3 Times the Rate of Democrats," Smart Politics, Feb. 10, 2011. Available: https://smartpolitics.lib.umn.edu/2011/02/10/selection-bias-politifact-rate/
[74] T. A. Heikkilä, "The role of fact-checkers in a post-truth world: Analyzing the impact of fact-checking on media trust," *Media and Communication*, vol. 10, no. 3, pp. 54-66, 2022. doi: 10.1080/21670811.2022.2031240.
[75] R. Iriondo, "Amazon Scraps Secret AI Recruiting Engine that Showed Biases Against Women," *Carnegie-Mellon School of Computer Science Bias in Machine Learning Series*. October 11, 2018. Available: https://www.ml.cmu.edu/news/news-archive/2016-2020/2018/october/amazon-scraps-secret-artificial-intelligence-recruiting-engine-that-showed-biases-against-women.html
[76] V. Hofmann, P.R. Kalluri, D. Jurafsky, et al., "AI generates covertly racist decisions about people based on their dialect," *Nature*, vol, 633, pp. 147–154, 2024.
[77] D. Hall, San Francisco is first US city to ban facial recognition. BBC, May 14, 2019. Available: https://www.bbc.com/news/technology-48276660
[78] M. Suleyman, and D. King (2019). "Using AI to give doctors a 48-hour head start on life-threatening illness," Google DeepMind. Available: https://deepmind.google/discover/blog/using-ai-to-give-doctors-a-48-hour-head-start-on-life-threatening-illness
[79] IBM Corporation, "IBM Watson for Oncology," 2016. Available: https://www.ibm.com/watson-health/solutions/oncology.
[80] "People + AI Research," Google, 2017. Available: https://pair.withgoogle.com/.
[81] M. Madaio, L. Stark, J. W. Vaughan, and H. Wallach, "Co-Designing Checklists to Understand Organizational Challenges and Opportunities around Fairness in AI," In *Proc. 2020 CHI Conf. on Human Factors in Computing Systems (CHI 2020)*, Honolulu, HI, USA, Apr. 2020. . Available: https://www.microsoft.com/en-us/research/publication/co-designing-checklists-to-understand-organizational-challenges-and-opportunities-around-fairness-in-ai/.
[82] R. Krishna, 'The Future of AI in the Automotive Industry: Revolutionizing Design, Production, and Operations," Hitachi, no date. Available: https://social-innovation.hitachi/en-us/think-ahead/digital/ai-revolution-in-automotive-design-production-operations/?utm_campaign=FY24US&utm_source=SEM&utm_medium=TLAIA_Search&gad_source=1
[83] S. Krügel, and M. Uhl. "The risk ethics of autonomous vehicles: an empirical approach," *Sci Rep* 14, 960 (2024).
[84] D. Berreby, "As Use of A.I. Soars, So Does the Energy and Water It Requires," *Yale Environment 360*, 2024. Available: https://e360.yale.edu/features/artificial-intelligence-climate-energy-emissions
[85] K. Rucker, "Deepseek chatbot has 83% 'fail rate' in accuracy: NewsGuard study," Straight Arrow News, January 29, 2025. Available: https://san.com/cc/deepseek-chatbot-has-83-fail-rate-in-accuracy-newsguard-study/
[86] "Recommendation on the Ethics of Artificial Intelligence," UNESCO, 2021. Available: https://www.unesco.org/en/artificial-intelligence/recommendation-ethics.]
[87] "IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems," IEEE Standards Association. Available: https://standards.ieee.org/industry-connections/activities/ieee-global-initiative/.
[88] "AI Ethics," IBM. Available: https://www.ibm.com/impact/ai-ethics.
[89] V. D. Kirova, C. S. Ku, J. R. Laracy, and T. J. Marlowe, "The Ethics of Artificial Intelligence in the Era of Generative AI," *J. Systemics, Cybernetics, and Informatics*, vol. 21, no. 4, pp. 42-50, 2023.

AUTHOR INFORMATION

Fr. Joseph R. Laracy is Associate Professor of Systematic Theology and an affiliated faculty member in Mathematics and Computer Science. He has broad research interests, including theology and ethics, cybernetics, and computer engineering.

Vassilka D. Kirova is an Adjunct Professor of Informatics at NJIT and an ex-Bell Labs Consulting director. Her research focus is on software engineering, automation, AI applications, operating models and autonomous systems.

Cyril S. Ku is Professor of Computer Science at William Paterson University. He does research in artificial intelligence, data science, natural language processing, databases, and software engineering.

Thomas J. Marlowe is Professor Emeritus of Mathematics and Computer Science at Seton Hall University. He has broad research interests including software engineering, computer science pedagogy, and interdisciplinary studies.