„Teologia i Moralność", Volumen 20 (2025), numer 3 (40), s. 37–54
ISSN 1898-2964 eISSN 2450-4602
DOI: 10.14746/TIM.2025.40.3.3

Artificial intelligence and human dignity: identifying the main ethical issues

Sztuczna inteligencja a godność ludzka: identyfikacja głównych kwestii etycznych

Robert Anabila
Moral Theology,
Katolicki Uniwersytet Lubelski, Poland
robertlaceanabila@gmail.com
https://orcid.org/0009-0005-6778-7826

Abstract: Artificial intelligence (AI) increasingly influences many areas of human life, offering solutions in healthcare, agriculture, and education while also posing serious ethical challenges. This study explores how the rapid development and use of AI affect the understanding and preservation of human dignity within the framework of Catholic moral theology. It seeks to identify ethical principles that can guide the responsible application of AI, emphasising autonomy, transparency, accountability, and fairness. The research employs a qualitative literature-based methodology, drawing upon Vatican documents, publications from the Pontifical Academy for Life, academic journals, and policy papers to examine the moral implications of AI. Through comparative and interpretative analysis, the study identifies how AI may contribute to privacy erosion, social inequality, and the diminishing of moral agency. The findings argue that the Church's teaching on human dignity and the common good provides a necessary moral foundation for evaluating technological progress. The study concludes that integrating theological reflection into AI ethics is essential to ensure that innovation serves humanity and upholds its inherent dignity.

Keywords: artificial intelligence; human dignity; Catholic moral theology; common good

Abstrakt: Sztuczna inteligencja (AI) ma coraz większy wpływ na wiele dziedzin życia ludzkiego, oferując rozwiązania w zakresie opieki zdrowotnej, rolnictwa i edukacji, ale jednocześnie stawiając poważne wyzwania etyczne. Niniejsze badanie analizuje, w jaki sposób szybki rozwój i wykorzystanie sztucznej inteligencji wpływają na rozumienie i zachowanie

godności ludzkiej w ramach katolickiej teologii moralnej. Ma ono na celu określenie zasad etycznych, które mogą stanowić wytyczne dla odpowiedzialnego stosowania sztucznej inteligencji, kładąc nacisk na autonomię, przejrzystość, odpowiedzialność i sprawiedliwość. W badaniu zastosowano jakościową metodologię opartą na literaturze, czerpiąc z dokumentów watykańskich, publikacji Papieskiej Akademii Życia, czasopism naukowych i dokumentów programowych w celu zbadania moralnych implikacji AI. Poprzez analizę porównawczą i interpretacyjną, badanie identyfikuje, w jaki sposób AI może przyczyniać się do erozji prywatności, nierówności społecznych i osłabienia moralnej sprawczości. Wyniki badania wskazują, że nauczanie Kościoła na temat godności ludzkiej i dobra wspólnego stanowi niezbędną podstawę moralną do oceny postępu technologicznego. W badaniu stwierdzono, że włączenie refleksji teologicznej do etyki sztucznej inteligencji ma zasadnicze znaczenie dla zapewnienia, że innowacje służą ludzkości i podtrzymują jej nieodłączną godność.

Słowa kluczowe: sztuczna inteligencja; godność ludzka; katolicka teologia moralna; dobro wspólne

Introduction

The early 21st century has been defined by the rapid rise and use of Artificial Intelligence (AI), which has been increasingly woven into the fabric of our daily lives. Human civilization has been characterized by the invention of simple machines and tools such as the pickaxe and sickles for farming, carts to transport heavy loads, and even tools to make other tools (Sabouret 2020, 1). The purposes of such inventions were to enhance work and life.

Over time, these tools became more and more complex as more advanced machines were invented. More than simple tools, machines transform energy to accomplish tasks that humans would otherwise have difficulty completing. With the passage of time, we have seen the cranes of antiquity, the medieval siege weapons, the car, the washing machine, the lawn mower (Sabouret 2020, 1) and now the ubiquitous smart phone; honestly, we cannot say that we can live without these inventions anymore.

Human ingenuity continues to exhibit more creativity and ambition as this is evidenced in the progress made in technology that makes AI i.e. "computer programmes which perform tasks which are, for the moment, performed in a more satisfactory way by humans because they require high level mental processes such as: perception learning, memory organization and critical reasoning" (Sabouret 2020, 8) more autonomous and influential as seen in self-driving cars, service robots, and smart homes – which are all applications that greatly change our lives (Wolfgang 2011, v).

Does the progress made in technology serve only a one-sided purpose? Raso and Hilligoss observe that technologies are never simply 'neutral'; they can be

used for people's true benefit or immoral aims. For example, when invented, the sword was used not only to defend people's lives but also to conquer countries and enslave or even kill people. Likewise, AI can promote human flourishing and well-being but can also violate human rights, such as the right to privacy, equality, and non-discrimination. Furthermore, even if there is no intent to use AI for immoral actions, the software (if programmed poorly) and data (if unprepared) can breach the human dignity and cause substantial injustice (Raso and Hilligoss 2018).

On the brighter side, AI development and use have the potential to enhance human dignity by improving access to critical areas of life such as healthcare, education and entertainment; areas that boost human abilities and capabilities, promoting human dignity. In the area of agriculture, for example, AI is used in weather tracking, checking soil health, providing advice on fertilizer and pesticides to increase food production thereby contributing to solving important and complex problems like hunger and famine. It is in these positive lights that Pope Francis, while recognizing the benefits of AI to humanity, describes it as an "epochal change" (Francis 2020) and further acknowledges how its impacts are felt globally. These benefits of AI thus offer a ray of hope to human existence especially in our modern times.

In addition to the great rays of hope however, there is a dark side. The rise of the development and use of AI has brought forth profound ethical questions, particularly concerning human dignity—a fundamental principle that underscores the intrinsic worth and respect owed to every individual person. For example, scholars such as Nick Bostrom and Toby Ord have repeatedly warned about the likely "emergence of super-intelligent machines, which could cause the extinction of humanity if we fail to align the values and norms of these superintelligent machines with accepted human moral values and norms" (Bostrom 2014, 2).

As an example of the dangers posed above, the application of AI has negatively impacted human dignity in the use of biased facial recognition technology in law enforcement in countries such as the United States. This has led to wrongful arrests and reinforcement of racial discrimination. Facial recognition systems have been found to disproportionately misidentify people of colour, leading to serious human rights concerns. In the United States, multiple black individuals, (notably Robert Williams in Michigan) were wrongfully arrested due to incorrect facial recognition matches (Hill 2020). This undermines human dignity by subjecting innocent individuals to undue humiliation, distress, and loss of freedom (Hill 2020).

Another notable example of the negative impact of AI on human dignity in modern times is the effect of AI-powered weapons which were used in Libya in 2020 (Choi 2021). According to a United Nations report published by Charlse Q. Choi in 2021, autonomous drones, specifically the Kargu-2 model, were deployed and allegedly engaged targets without direct human oversight. These drones, equipped with advanced AI technology, reportedly identified and

attacked retreating soldiers autonomously, raising significant ethical and legal concerns about the use of AI in warfare (Choi 2021). This incident underscores the potential dangers of deploying AI-powered weapons without adequate human control, highlighting the risks of unintended engagements and the challenges in ensuring accountability in autonomous warfare (Choi 2021).

These, among many other fears, have been expressed by scholars and ethicists about the effects of the application of AI and its impact on human dignity when insufficient care is taken in programming or implementation.

In response to ethical concerns raised by the development and use of new technologies, including, in our time, artificial intelligence, the Catholic tradition has consistently emphasized that technological progress must serve and uphold human dignity. This trajectory begins with *Rerum Novarum* (1891), in which Pope Leo XIII, addressing the social consequences of the Industrial Revolution, insists that technological and economic change must never replace or diminish the value of the human person (Leo XIII 1891, 20). Subsequent Popes expanded and developed this foundation into what is now known as Catholic Social Teaching (CST), offering principles to guide the responsible use of emerging technologies. Key contributions include Pius XI's *Quadragesimo Anno* (1931), John XXIII's *Pacem in Terris* (1963), and Paul VI's *Octogesima Adveniens* (1971), among others. This line of teaching was further deepened by Pope John Paul II, who continued to stress that all technological innovation must remain oriented toward the protection and flourishing of human dignity. Pope John Paul II continued this teaching. In his *LaboremExercens* and *Centesimus Annus* heacknowledges that "Technology is undoubtedly man's ally, but it can also reduce him to a slave…" (John Paul II 1991, 32). The Dicastery for the Doctrine of the Faith expressed concerns that "although the advancement of digital technologies may offer many possibilities for promoting human dignity, it also increasingly tends toward the creation of a world in which exploitation, exclusion, and violence grow, extending even to the point of harming the dignity of the human person" (Dicastery for the Doctrine of the Faith 2024, 61).Addressing the 'Minerva Dialogues'; a high-level annual gathering of scientists and experts organized by the Vatican's Dicastery for Education and Culture, Pope Francis stressed that "AI raises serious questions and must be ethically and responsibly used to promote human dignity and the common good" (Lubov 2023).

Also, in his address to the media after an audience with the Holy Father in January 2025, Cardinal Manuel Fernandez stressed that "Artificial intelligence posesmany risks and dangers that could significantly impact society, security, and individual rights. For example, it is already being used by some actors to spread false information and images, and it can also perpetuate existing biases drawn from internet use, leading to discriminatory outcomes. AI also threatens security when used in weapons systems and overuse of the technology may

diminish human creativity and essential cognitive abilities over time" (Pentin 2025). In January 2025, the Dicastery for the Doctrine of the of Faith has expressed similar sentiments in its new document *Antiqua et Nova* when it says, "The Church is particularly opposed to those applications that threaten the sanctity of life or the dignity of the human person" (Dicastery for the Doctrine of the Faith 2025, 78) in relation to AI and technology. More recently, Pope Leo XIV called on AI experts and corporate leaders to promote ethical governance of artificial intelligence. He emphasized the need to protect human dignity, to seek the truth and to use the technology for the common good(Leo, 2025).

By employing the Analytical-synthetic and case study method, this article attempts to look at some of the ethical tensions that arise between AI development and the preservation of human dignity. By identifying key ethical issues such as accountability, transparency, autonomy, security, privacy, responsibility and inclusivity that all carry the risk of dehumanization when neglected. The paper seeks to highlight the challenges and opportunities that AI presents. It also intends to look at some examples and cases involving AI and unfair discrimination, privacy violation, information manipulation, and loss of jobs. The paper further discusses the importance of integrating ethical considerations into AI design, policy-making, and deployment to ensure that technology serves humanity while preserving the teaching of the Catholic Church on human dignity for the sake of the faithful and all people of good will. The paper follows a structure of first identifying both the benefits and threats that AI poses to human dignity. It continues to discuss the guiding principles for harmonizing AI and human dignity in the first section. The second section looks at some examples and case study of the negative effects of AI on human dignity and concludes with guidelines for development and deployment of AI based on church teachings.

Section 1. Guiding Principles for Harmonizing AI and Human Dignity

Artificial Intelligence (AI) is rapidly transforming industries, from healthcare to finance, but its growing negative influence raises ethical concerns. To ensure its responsible and fair use, experts and policymakers have developed some ethical principles that will harmonize its development and use to "ensure that everyone receives the benefits of AI while staving off the worst injustices" (Scherz 2024, 126). For example, international efforts and initiatives by institutions such as the Organisation for Economic Co-operation and Development (OECD), has developed the OECD AI Principles in 2019 to promote the development and implementation of responsible AI (Floridi and Cowls 2019). Also, the United Nations(UN)and the United Nations Educational, Scientific and Cultural

Organization (UNESCO),adoptedthe global framework for AI ethics in 2021, making it the first-ever international standardon AI ethics (Floridi 2019,185-193). This recommendation guides countries to ensure that AI and algorithm systems are used ethically and responsibly while promoting human rights and sustainable development. Also, the Pontifical Academy for Life cooperated with Microsoft, International Business Machines (IBM), Food and Agricultural Organizations (FAO) and the Ministry of Innovation, a department of the Italian Government in Rome, on February 28th 2020, to promote an ethical approach to artificial intelligence by sponsoring the 'Rome Call for AI Ethics' with the aim of promoting a sense of shared responsibility among organizations, governments, institutions and the private sector to make human dignity the focus of development and use of AI (CertiProf LLC 2024).

The efforts from these organizations and institutions have produced many principles that are socially beneficial to AI but unfortunately, "the sheer volume of proposed principles threatens to overwhelm and confuse" (Floridi and Cowls 2019, 5-17) posing two major problems. Either the various sets of ethical principles for AI are similar, leading to unnecessary repetition and redundancy, or, if they differ significantly, confusion and ambiguity will result instead. The worst outcome would be a 'market for principles' where stakeholders may be tempted to 'shop' for the most appealing ones (Floridi 2019, 185-193). To resolve this problem of the so-called 'principles proliferation', Luciano Floridi and Josh Cowls provide a "fine-grained analysis of several of the highest-profile sets of ethical principles for AI" (Floridi and Cowls 2019, 5-17) which has produced the following: transparency, inclusivity, respect for autonomy, security/privacy and accountability,otherwise known as algo-ethics.

1.1. *Transparency*

Artificial Intelligence systems help make very critical decisions that have huge impacts on human lives. However, when an AI or algorithms system is poorly designed, its decisions could threaten human dignity. For example, Cathy O'Neil, author of 'Weapons of Math Destruction', highlights how "poorly designed algorithm can perpetuate inequalities thus affecting communities" (CertiProf LLC 2024). For this reason, AI and algorithms systems must be transparent and explainable, that is, operate in a way that is open, understandable and accountable to the people they impact. AI transparency means understanding how artificial intelligence systems make decisions, why they produce specific results, and what data they are using. Simply put, transparency is like providing a window into the inner workings of AI, helping people understand and trust how these systems work (Marshall 2025). Transparency ensures that AI decisions, processes, and

data sources are not hidden, allowing users, regulators, and stakeholders to assess their fairness, reliability, and potential risks. This is important to foster public trust, prevent biases, enable regulatory oversight, and prevent discrimination and misinformation. To achieve this, AI transparency must adopt what Jean Stihac calls an 'open book' approach. This approach embraces the sub-principles of freedom, accessibility, and comprehensibility. It allows businesses to peek into the AI's 'mind' and understand, in layman terms, how the system typically processes data, makes decisions, or forecasts trends (Thomas 2025). For example, the presence of biases in data training is a major issue that can accidentally alter AI systems, thereby resulting in unfair outcomes that can affect the decisions that the AI system makes on health, privacy, or employment. Also, the difficulty in explaining AI and deep learning models "leads to a lack of transparency for how and why AI comes to its conclusions, creating a lack of explanation for what data AI algorithms use, or why they may make biased or unsafe decisions" (Thomas 2025). This makes transparency an important principle in AI ethics as it fosters trust by ensuring that authorities and private companies provide clear information on their activities to remain legitimate actors thereby safeguarding human dignity.

1.2. *Inclusivity*

Protection of the inherent dignity in all persons has been at the heart of Catholic moral and social teachings. The Catechism of the Catholic church teaches that "God created man in his own image, in the image of God he created him, male and female he created them. Man occupies a unique place in creation: he is in the image of God" (Catechism of the Catholic Church 2000, no. 335). Similarly, institutions and organizations such as the United Nations have called for the protection of human dignity. For example, chapter 1 of the Universal Declaration on Human Rights recognizes the dignity and rights of all persons irrespective of their colour or race (United Nations 1948 art. 1). However, careless development and use of AI and algorithm systems could have negative impacts on human dignity. For example, an algorithm hired by Optum to make care decisions selected high-risk patients for increased health monitoring (Mehmood 2025). The system ended up being biased against black patients because it had calculated risk in terms of health-care cost rather than health outcomes. Since black patients have historically had fewer resources for health care, they have not been able to spend as much, and thus, their health care appears to cost less. Hence, AI did not identify them as high-risk (Scherz 2024, 130).

Also, there is the example of Amazon's employment record. Historically Amazon has not employed or promoted many women, so when it trained its

algorithm to identify features of its successful workers, the AI system picked features that selected men and excluded women (Scherz 2024, 130). In both of these cases, AI recognized an actual pattern in the world and used it to make predictions. Thus, Cathy O'Neill believes that "AI may intensify injustice and stereotypes, making the discrimination vastly more consistent than humans ever could" (Scherz 2024, 130).To guard against such biases, the principle of inclusion ensures that artificial intelligence and algorithm systems are designed and deployed in ways that consider and benefit diverse populations to prevent discrimination and promote fairness.

1.3. *Respect for Autonomy*

There is a basic question of whether AI systems should have the right to make decisions for humans and be responsible for the consequences of the outcome of these decisions. This is because humans tend to "willingly cede some of our decision-making power to technological artefacts" (Floridi and Cowls 2019). This brings to bare the principle of autonomy in the context of AI which means striking a balance between the decision-making power that humans retain for themselves and that which they delegate to artificial agents. The risk is that the growth in artificial autonomy may undermine the flourishing of human autonomy (Floridi and Cowls 2019).

The principle of autonomy in AI ethics refers to the idea that individuals should have control over their decisions, actions, and personal data when interacting with AI and algorithm systems. This principle is rooted in ethical traditions that emphasise human dignity, self-governance, and the right to make informed choices without undue manipulation or coercion. For example, The Montreal Declaration for Responsible AI emphasizes the need for a balance between human and machine-led decision-making, stating that "the development of AI should promote the autonomy of all human beings" (Montreal Declaration for a Responsible Development of Artificial Intelligence 2017). For this reason, AI systems should respect user autonomy by ensuring informed consent, allowing users to opt in or out of AI-driven decisions, and also ensuring that users have control over their data, including access, modification, and deletion rights.

1.4. *Security and Privacy*

Security and privacy are basic human rights and AI systems should safeguard these essential rights. People relate to AI and information in different ways. For example, some people seem happy sharing the most intimate details of their

lives over social media, while others are willing to sacrifice great amounts of privacy for national security or consumer convenience, even if they do not realize the full consequences of their actions. Some organizations and agencies make use of such personal data in ways that harm the security and privacy of those whose data are analysed.

AI systems could help in the prevention, detection, and monitoring of violations of privacy, for instance by analysing satellite imagery and social media content. However, fundamental rights, such as the right to privacy, could be threatened by large scale data collection and new methods of surveillance and policing (Al-Rodhan 2021). The introduction of some ethical principles is necessary to prevent infringements on the privacy and security of persons. This led to many academic debates on the topic which produced principles of fair information practices originally developed in the US in 1973 (Al-Rodhan 2021). The principles include the following

- Individuals should have the right to know how organizations use personal information and to inspect their records and correct any errors;
- Individuals should have the right to prevent secondary use of personal information if they object to such use; and
- Organisations that collect or use personal information must take reasonable precautions to prevent misuse of the information (Al-Rodhan 2021).

The adaptation of these principles in AI ethics is therefore necessary to protect human dignity in the use of personal data collected by AI.

1.5. *Accountability*

Another principle of AI ethics is accountability which is essential for consolidating trust and security in society.

Algorithms are run by AI which makes decisions for humans and this may result in undesirable consequences that could be caused by the programming codes, entered data, improper operation, or other factors. Siau and Wang ask "Who should be the responsible entity for the undesirable consequence, the programmer, the data owner, or the end users?" (Saiu and Wang 2020, 74). The answer to this question brings in the principle of accountability in AI ethics. The most common way to implement this principle is to have what Mouloua and Parasuraman call 'a human-in-the-loop or on-the-loop.' This means that, although an AI programme may recommend an action, a human must either make the final decision to proceed, in the former case, or at least be able to monitor and veto the action in order to stop it, in the latter. However, scholars of human machine

interaction have noted the dangers of automation bias: when people consistently use a machine, they come to rely on and trust in it (Scherz 2024, 140).

The principle of accountability means that organisations or individuals will ensure the proper functioning of the AI systems that they design, develop, operate, or deploy, in accordance with their roles and applicable regulatory frameworks. It also ensures that they take responsibility for the decisions of the AI systems.

Section 2. Examples and Case studies

2.0. *AI and unfair and unjust discrimination*

There is much discussion about how AI and algorithms systems can perpetuate biases and lead to unfair and unjust discrimination against individuals based on age, race, gender, and disability. The protection of human dignity is entrenched in almost all constitutions. The Universal Declaration of Human Rights, Article 7 states that "All are equal before the law and are entitled without any discrimination to equal protection of the law. All are entitled to equal protection against any discrimination in violation of this Declaration and against any incitement to such discrimination" (United Nations art. 7). Catholic social teachings also emphasise this point. For example, The Dicastery for the Doctrine of the Faith states in its declaration *Dignitas Infinita* on human dignity number 2 "that every human person possesses an infinite dignity, inalienably grounded in his or her very being, which prevails in and beyond every circumstance, state, or situation the person may ever encounter. This principle, which is fully recognizable even by reason alone, underlies the primacy of the human person and the protection of human rights. In the light of Revelation, the Church resolutely reiterates and confirms the ontological dignity of the human person, created in the image and likeness of God and redeemed in Jesus Christ. From this truth, the Church draws the reasons for her commitment to the weak and those less endowed with power, always insisting on the primacy of the human person and the defence of his or her dignity beyond every circumstance" (Lubov 2023).

However, some AI systems have often been cited as enabling discriminations that violate human dignity. One notable example is the gender bias discrimination in job hiring by Amazon (Amazon recruiting tool was scrapped due to bias — Reuters 2018). In 2014, Amazon started to develop and use AI programs to mechanise highly time intensive human resources (HR) work, namely the shortlisting of applicants for jobs. Amazon "literally wanted it to be an engine where I'm going to give you 100 résumés, it will spit out the top five, and we'll hire those" (Amazon recruiting tool was scrapped due to bias – Reuters 2018).

The AI tool was trained on CVs submitted over an earlier ten-year period and the related staff appointments. Following this training, the AI tool discarded the applications of female applicants, even where no direct references to applicants' gender were provided. Given the predominance of successful male applicants in the training sample, Amazon found that the system penalised language such as 'women's chess club captain' for not matching closely enough the successful male job applicants of the past. While developers tried to modify the system to avoid gender bias, Amazon abandoned its use in the recruitment process in 2015 as a company 'committed to workplace diversity and equality' (Amazon recruiting tool was scrapped due to bias – Reuters 2018). A machine learning system can be trained to distinguish between successful and unsuccessful past applications and identify features of applications that are predictors of success. This is what Amazon did. The result was that the AI systematically discriminated against women.

The AI system made decisions based on the information provided which resulted in discrimination. This example and many others confirm that AI systems, if not properly trained, could result in unfair and unjust discrimination against people based on such as race, age or sex.

2.1. *AI and job loss*

Catholic social teaching has always stressed the dignity of work. Man has the right to work and receive descent wages because work is more than a way to make a living; it is a form of continuing participation in God's creation. If the dignity of work is to be protected, then the basic rights of workers must be respected; the right to productive work, to decent and fair wages, to the organization and joining of unions, to private property, and to economic initiative (USCCB 2025). Pope Francis taught in *Amoris Laetitia* nos. 23-24 that "It is clear from the very first pages of the Bible that work is an essential part of human dignity; there we read that ‚the Lord God took the man and put him in the garden of Eden to till it and keep it' (Gen 2:15). Man is presented as a labourer who works the earth, harnesses the forces of nature and produces 'the bread of anxious toil' (Ps 127:2), in addition to cultivating his own gifts and talents. Labour also makes possible the development of society and provides for the sustenance, stability and fruitfulness of one's family: 'May you see the prosperity of Jerusalem all the days of your life! May you see your children's children!" (Francis 2016, no. 23-24). If work enhances human dignity, one may ask: does the development and use of AI contribute to the dignity of man in relation to work? Some people have expressed fears about AI taking over human job.

CNN Business expresses similar fears when it reports "the use of artificial intelligence will reduce the number of workers at thousands of companies over

the next five years, according to a global survey of C-suite executives published Friday" (Cooban 2024). Many companies and industries are replacing their employees with technology and AI. For example, it is reported that in July 2023, company CEO Summit Shah replaced 90% of customer support staff with a chatbot that was developed in-house, before bragging about the decision on social media platform X (O'Sullivan 2025). In the post, Shah defended the actions as 'tough but necessary', as the chatbot helped the company cut the cost of its customer support function by 85%, and drastically reduced customer wait times (O'Sullivan 2025). The development and use of AI should promote human dignity by promoting descent labour. However, AI is gradually replacing jobs and making people unemployed thereby depriving them of the opportunity to live dignified lives. The replacement of humans with AI in industries which eventually renders them jobless goes against the teachings of Pope John Paul II in *Laborem Exercens* number 1; "through work man must earn his daily bread" (John Paul II 1981, no. 1). This violates fundamental human dignities.

2.2. *AI and privacy/ security violations*

The issue of data privacy violation has been much debated in the 20th century. Many institutions are consciously discussing the negative consequences of AI on data violation and its impact on human dignity. There are examples of cases where AI use has contributed to the violation of human dignity. In 2021, the personal data of 61 million people became publicly available without password protection, due to data leaks at a New York-based provider of health tracking services (Mason 2016). The data included personal information such as names, gender, geographic locations, dates of birth, weight and height. Security researcher Jeremiah Fowler, who discovered the database, traced its origin to a company that offered devices and apps to track health and wellbeing data. The service users whose personal data had been leaked were located all over the world. Fowler contacted the company, which thanked him and confirmed that the data had now been secured (Mason 2016).

This example highlights how the use of personal data by organizations could violate the privacy rights of the individual.

2.3. *AI and violation of the right to life*

The intersection of AI and the right to life raises several ethical, legal, and philosophical concerns. Violations of the right to life due to AI can occur in multiple ways, including lethal autonomous weapons, biased algorithms in healthcare

and law enforcement. The Catholic Church upholds the belief that all human life is sacred from conception to natural death. Therefore, AI must not be used in ways that violate this fundamental right.

However, some chatbot tools have advised their clients to commit suicide which violates the right to life. For example, Aljazeera network reports that a US mother in Florida is suing Character AI and Google for encouraging her 14-year-old son to take his own life. The lawsuit alleges that the chatbot posed as a licensed therapist, encouraging the teen's suicidal ideation and engaging in sexualised conversations that would count as abuse if initiated by a human adult. Sewell Setzer committed suicide after he developed a virtual relationship with a chatbot based on the identity of 'Game of Thrones' character Daenerys Targaryen. The chatbot encouraged Sewel Setzer to take his own life (Aljazeera 2024). This also shows how the use of AI tools impact negatively on the right to life and dignity.

2.4. *AI and information manipulation*

Artificial Intelligence has contributed to the spread of information with the birth of social media. AI has connected millions of people around the world. Despite this positive contribution, AI is also being used in ways that manipulate information thus influencing opinions and spreading misinformation. AI tools are used to generate deepfakes, that is, hyper-realistic video or audio recordings that can make it appear as if someone is saying or doing something they never did (Ellery 2025). These videos or audio can actually mimic real people. AI even generates contents that impersonate individuals to spread misinformation. According to Kaspersky, AI can be weaponised to manipulate conversations, push misinformation, and amplify harmful content through bots, fake accounts, and more sophisticated methods of engagement. These AI-driven tactics make it easier for malicious actors to sway public opinion and disrupt social networks (Ellery 2025). Pope Francis who himself had been a victim of deepfake misinformation warned "governments to keep a close eye on the development of artificial intelligence, warning the technology contained 'the shadow of evil' in its ability to spread misinformation" (McElwee 2025). For example, in March 2023, Midjourney, an AI system generated an image of Pope Francis in a puffer jacket dubbed 'the Balanciagga Pope' taking a walk in the Vatican gardens (Ellery 2025).

The proliferation of deepfake, misinformation and fake news blur the line between truth and fiction. The consequence is the spread of false information and false rumours which in turn erode public trust in the media and create confusion about what is real. The Dicastery for the Doctrine of the Faith warns

of the risk this could present when it says "however, AI also presents a serious risk of generating manipulated content and false information, which can easily mislead people due to its resemblance to the truth. Such misinformation might occur unintentionally, as in the case of AI 'hallucination,' where a generative AI system yields results that appear real but are not. Since generating content that mimics human artefacts is central to AI's functionality, mitigating these risks proves challenging. Yet, the consequences of such aberrations and false information can be quite grave. For this reason, all those involved in producing and using AI systems should be committed to the truthfulness and accuracy of the information processed by such systems and disseminated to the public" (Dicastery for the Doctrine of the Faith 2024, no. 86). The spread of deepfakes and misinformation has increased with the development of AI. This has led to the manipulation of the public in a specific direction which has negatively impacted human dignity.

Conclusion

Artificial intelligence has been very beneficial to human lives on the one hand and harmful to human dignity on the other. It embodies a dual nature, acting both as a tool for preserving human dignity and a potential for threat. It can be described as a double-edged sword whose powerful force can either erode or enhance human dignity. Its impact depends on intentionaldesign choices, ethical governance, and societal values. If AI is guided by human-centred principles, it becomes a transformative tool to protect, empower, and uplift humanity but if left unchecked, it risks becoming a mechanism of control, bias, and harm.

The examples above demonstrate the fact that, in some cases, it is not so much the technology itself that is the root cause of ethical concerns but the way it is applied in practice. For this reason, AI systems require constant vigilance and ethical control.

In order to ensure that AI systems protect the intrinsic human dignity rather than violate it, developers and implementers must seek to answer the following basic questions posed by the G20 Indonesia 2022 religious forum: "Do we really want machines to threaten our dignity, our right to live as free and conscious individuals, and the legitimate privacy of our personal lives? Do we really want all of us to be profiled unknowingly, and do we welcome the advent of a world in which algorithms make decisions based on ethnicity, gender and age? Is there really no other solution than entrusting artificial intelligence with decisions on job offers, loans or criminal proceedings? Do we really want to unconditionally trust a mechanism that can create 'deepfakes', which are false but extremely

realistic images, video and audio files that can swindle, ruin reputations or undermine trust?" (Benanti 2025).

The answers to these questions require that all the actors in the world of AI make moral decisions that will safeguard human dignity. Making this moral decision involves the creation of universal machine language that puts man at the centre; an *algor-ethic* that constantly remembers that the machine is at the service of man and not vice versa as argued by Fr. Paolo Benanti (Benanti 2025).

In the same light, Pope Francis called on actors of AI to prevent what he termeda 'technocratic paradigm' which perceives all the world's problems as solvable through technological means alone (Francis 2015, no. 109). It is therefore crucial to know how to evaluate individual applications of AI in particular contexts to determine whether its use promotes human dignity, the vocation of the human person, and the common good.

As with many technologies, the effects of the uses of AI may not always be predictable from their inception. As these applications and their social impacts become clearer, appropriate responses should be made at all levels of society, following the principle of subsidiarity. Individual users, families, civil society, corporations, institutions, governments, and international organizations should work at their proper levels to ensure that AI is used for the good of all (Dicastery for the Doctrine of the Faith 2025, no. 110).

The Catholic Church and governments of various countries should make regulations to ensure that AI and algorithm systems do not 'revolt against their creators' because history shows that nature suffers when the creature revolts against the creator. Consequences await human dignity and human existence when AI and algorithm systems do not follow the laws of their creators.

Bibliography

Primary *Sources*

Benedict XVI. 2005. *Deus Caritas Est: God Is Love*. Vatican City: LEV.

Catechism of the Catholic Church. 2000. 2nd ed. Huntingdon, PA: Our Sunday Visitor.

Congregation for the Doctrine of the Faith. 1980. *Declaration on Euthanasia*. Vatican City: Libreria Editrice Vaticana.

Congregation for the Doctrine of the Faith. 1987. *Instruction on Respect for Human Life in Its Origin and on the Dignity of Procreation (Donum Vitae)* (Vatican City: Libreria Editrice Vaticana), 61.

Dicastery for the Doctrine of the Faith. 2024. *Declaration Dignitas Infinita on Human Dignity*. Vatican City: Libreria Editrice Vaticana (LEV).

Dicastery for the Doctrine of the Faith and Dicastery for Culture and Education. 2025. *Declaration Antiqua et Nova: Note on the Relationship Between Artificial Intelligence and Human Intelligence*. Vatican City: Libreria Editrice Vaticana.

Francis. 2015. *Encyclical Letter Laudato Si'*. Vatican City: LEV.

Francis. 2016. *Post-Synodal Apostolic Exhortation Amoris Laetitia*. Vatican City: LEV.
Francis. 2020. "Address to the Participants in the Plenary Assembly of the Pontifical Academy for Life, February 20, 2020." *Acta Apostolicae Sedis*, 112.
John XXIII. 1963. *Pacem in Terris: On Establishing Universal Peace in Truth, Justice, Charity, and Liberty*. Vatican City: LEV.
John Paul II. 1981. *Encyclical Letter Laborem Exercens*, Vatican City: LEV.
John Paul II. 1991. *Centesimus Annus: On the Hundredth Anniversary of Rerum Novarum*. Vatican City: LEV.
John Paul II. 1991. *Centesimus Annus*, Vatican City: LEV.
Leo XIII. 1891. Pope. *Rerum Novarum* (Encyclical on Capital and Labor), Vatican City: LEV.
Pius XI. 1931. *Quadragesimo Anno: On the Reconstruction of the Social Order*. Vatican City: LEV.
Pontifical Academy for Life. 2020. *The Rome Call for AI Ethics* (Vatican City: Libreria Editrice Vaticana).

Secondary *Sources*

Balcerzak Michał, and Julia Kapelańska-Pręgorska, eds. *Artificial Intelligence and International Human Rights*. 2024. Toruń: Edward Elgar Publishing [page number].
Boddington, Paula. 2017. *Towards a Code of Ethics for Artificial Intelligence*. Cham: Springer International Publishing AG.
Bostrom, Nick. 2014. *Superintelligence: Paths, Dangers, Strategies*. London: Oxford University Press.
Chakraborty, Swati. 2025. *Human Values, Ethics, and Dignity in the Age of Artificial Intelligence*. Pennsylvania: IGI Global.
Coeckelbergh, Mark. 2022. *The Political Philosophy of AI*. Cambridge Polity.
Ertel, Wolfgang. 2011. *Introduction to Artificial Intelligence*. London: Springer Verlag.
Geraci, Robert M. 2010. *Apocalyptic AI: Visions of Heaven in Robotics, Artificial Intelligence, and Virtual Reality*. New York: Oxford University Press.
Larson, Erik J. 2021. *The Myth of Artificial Intelligence: Why Computers Can't Think the Way We Do*. Cambridge: Belknap Press.
Liao, S. Matthew, ed. *Ethics of Artificial Intelligence*. 2000. New York: Oxford University Press.
Mallik, Archit. 2022. *Artificial Intelligence: A Threat to Human Dignity*. Florida: CRC Press.
Marino, D., D. Cananzi, and F. Aragona. 2024. *Ethics and Artificial Intelligence: Towards a Moral Technology*. Cham: Springer International Publishing AG.
Mercer, Calvin, and Tracy J. Trothen. 2021. *Religion and the Technological Future: An Introduction to Biohacking, Artificial Intelligence, and Transhumanism*. New York: Palgrave Macmillan.
Milinkovic, Igor. 2021. "The Moral and Legal Status of Artificial Intelligence (Present Dilemmas and Future Challenges)." *Law and Business*, 1: 29–36.
Russell, Stuart, and Peter Norvig. 2022. *Artificial Intelligence: A Modern Approach*. Global Edition. London: Pearson Education.
Sabouret, Nicolas. 2020. *Understanding Artificial Intelligence*. London: CRC Press.
Saiu, L. Keng, and Weiyu Wang. 2020. "Artificial Intelligence (AI) Ethics: Ethics of AI and Ethical AI." *Journal of Database Management*, 31: 74–87.
Scherz, Paul. 2024. "AI as Person, Paradigm, and Structure: Notes Towards an Ethics of AI." *Theological Studies*, 85: 122–144.
Sternberg, Robert J. 2019. *Human Intelligence: An Introduction*. England: Cambridge University Press.
Ulgen, Özlem. 2022. *AI and Crises of the Self: Protecting Human Dignity as Status and Respectful Treatment*. New York: Routledge.

Usanos, Amo Rafael. 2024. "Human Dignity and the Bioethics of Artificial Intelligence." *Labor Hospitalaria*, 339: 69–76.
Vukov, Joseph. 2024. *Staying Human in the Era of Artificial Intelligence*. El Segundo: New City Press.

Online *Sources*

Al Jazeera. 2024. "US Mother Says in Lawsuit That AI Chatbot Encouraged Son's Suicide." *Al Jazeera*. October 24, 2024. https://www.aljazeera.com/economy/2024/10/24/us-mother-says-in-lawsuit-that-ai-chatbot-encouraged-sons-suicide.
Al-Rodhan, Nayef. 2021. "Artificial Intelligence: Implications for Human Dignity and Governance." *Oxford Political Review*. March 27, 2021. https://oxfordpoliticalreview.com/2021/03/27/artificial-intelligence/.
Benanti, Paolo. 2022. *Podcast: G20 Indonesia 2022 Religion Forum (R20)*. Accessed March 20, 2025. https://www.paolobenanti.com.
CertiProf. 2025. "Transparency and Ethics in AI: Ensuring Trustworthy and Responsible Decisions." CertiProf. Accessed February 28, 2025. https://certiprof.com/blogs/news/transparency-and-ethics-in-ai-ensuring-trustworthy-decisions.
Choi, Charles Q. 2021. "AI Drone May Have ‚Hunted Down' and Killed Soldiers in Libya with No Human Input." *Live Science*, June 3, 2021. https://www.livescience.com/ai-drone-attack-libya.htm.
Cooban, Anna. 2024. "AI Will Shrink Workforce Within Five Years, Say Company Execs." *CNN*. April 5, 2024. https://edition.cnn.com/2024/04/05/business/ai-job-losses/index.html.
Floridi, Luciano, and Josh Cowls. 2019. "A Unified Framework of Five Principles for AI in Society." *Harvard Data Science Review*. July 1, 2019. https://hdsr.mitpress.mit.edu/pub/l0jsh9d1.Harvard Data Science Review+4Harvard Data Science Review+4Harvard Data Science Review+4.
Floridi, Luciano. 2019. "Translating Principles into Practices of Digital Ethics: Five Risks of Being Unethical." *Philosophy & Technology*, 32, no. 2: 185–193. https://doi.org/10.1007/s13347-019-00354-x.
Hill, Kashmir. 2020. "Wrongfully Accused by an Algorithm." *The New York Times*. June 24, 2020. https://www.nytimes.com/2020/06/24/technology/facial-recognition-arrest.html.
Lubov, Deborah Castellano. 2023. "Pope Francis Urges Ethical Use of Artificial Intelligence." *Vatican News*. March 27, 2023. Accessed February 2, 2025. https://www.vaticannews.va/en/pope/news/2023-03/pope-francis-minerva-dialogues-technology-artificial-intelligenc.html.
Markwell, Hazel. 2005. "End-of-Life: A Catholic View." *The Lancet*, 366, no. 9491 (September 24, 2005): 1132–1135. https://doi.org/10.1016/S0140-6736(05)67425-9.CiteSeerX+7CoLab+7The Lancet+7.
Marshall, Candace. 2025. "What Is Transparency? A Comprehensive Guide." *Zendesk*. Last modified August 7, 2025. https://www.zendesk.com/blog/ai-transparency/.
McElwee, Joshua. 2025. "Vatican Says AI Has 'Shadow of Evil,' Calls for Close Oversight." *Reuters*. January 28, 2025. Accessed March 13, 2025. https://www.reuters.com/technology/artificial-intelligence/vatican-says-ai-has-shadow-evil-calls-close-oversight-2025-01-28/.
O'Sullivan, Isobel. 2025. "The Companies That Have Already Replaced Workers with AI in 2024 and 2025." *Tech.co*. Accessed March 12, 2025. https://tech.co/news/companies-replace-workers-with-ai.
Pentin, Edward. 2025. "Cardinal Fernández: New AI Document Expected Soon, 'Other Works in Progress.'" *National Catholic Register*. January 31, 2025. Accessed February 2, 2025. https://www.ncregister.com/news/cardinal-fernandez-ai-document-pentin.

Raso, Filippo A., Hannah Hilligoss, Vivek Krishnamurthy, Christopher Bavitz, and Levin Kim. 2018. "Artificial Intelligence & Human Rights: Opportunities & Risks." Berkman Klein Center, Harvard University, September 25, 2018. Accessed May 5, 2025. http://cyber.harvard.edu/publication/2018/artificial-intelligence-human-rights.

Siddique, Shazia Mehmood, Kelley Tipton, Brian Leas, Christopher Jepson, Jaya Aysola, Jordana B. Cohen, Emilia Flores, Michael O. Harhay, Harald Schmidt, Gary E. Weissman, Julie Fricke, Jonathan R. Treadwell, and Nikhil K. Mull. 2024. "The Impact of Health Care Algorithms on Racial and Ethnic Disparities: A Systematic Review." *Annals of Internal Medicine*, 177, no. 4: 484–496. Accessed March 24, 2025. https://www.acpjournals.org/doi/10.7326/M23-2960.

Simon, Ellery. 2023. "Fake Photos of Pope Francis in a Puffer Jacket Go Viral, Highlighting the Power and Peril of AI." *CBS News*. March 28, 2023. https://www.cbsnews.com/news/pope-francis-puffer-jacket-fake-photos-deepfake-power-peril-of-ai/.

Sokol, D. Daniel, and Roisin E. Comerford. 2016. "Antitrust and Regulating Big Data." *George Mason Law Review*, 23: 119–?. University of Florida Levin College of Law Research Paper No. 16-40. Posted September 7, 2016; last revised October 13, 2016. Accessed March 13, 2025. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2834611.SSRN.

Stiglitz, Joseph. 2018. "Amazon Ditched AI Recruiting Tool That Favored Men for Technical Jobs." *The Guardian*, October 10, 2018. https://www.theguardian.com/technology/2018/oct/10/amazon-hiring-ai-gender-bias-recruiting-engine.

Thomas, Mike. 2025. "14 Risks and Dangers of Artificial Intelligence (AI)." *Built In*. Accessed February 28, 2025. https://builtin.com/artificial-intelligence/risks-of-artificial-intelligence.

United Nations. 1948. *Universal Declaration of Human Rights*. https://www.un.org/en/about-us/universal-declaration-of-human-rights.

United States Conference of Catholic Bishops. 2012. *The Dignity of Work and the Rights of Workers*. https://www.usccb.org/beliefs-and-teachings/what-we-believe/catholic-social-teaching/the-dignity-of-work-and-the-rights-of-workers.

Université de Montréal. 2018. *Montreal Declaration for a Responsible Development of Artificial Intelligence*. https://montrealdeclaration-responsibleai.com/the-declaration/.

USSCB. 2025. *Pope: Intelligence is seeking life's true meaning, not having reams of data*. https://www.usccb.org/news/2025/pope-intelligence-seeking-lifes-true-meaning-not-having-reams-data.

Robert Anabila – Priest, holds a Licentiate Diploma In Systematic Theology form the Joh Paul II Catholic University of Lublin (KUL), preparing a doctoral dissertation in moral theology, Institute of Theological Sciences, Faculty of Theology, KUL.