Consensus Statement of the Workshop on Artificial Intelligence, Justice, and Democracy

“When used correctly, AI assists the human person in fulfilling his or her vocation, in freedom and responsibility” Pope Francis

“There is, however, the risk that AI will be used to advance the “technocratic paradigm”, which perceives all the world’s problems as solvable through technological means alone. Within this paradigm, human dignity and fraternity are frequently subordinated in the pursuit of efficiency, as though reality, goodness, and truth inherently emanate from technological and economic power. Yet human dignity must never be violated for the sake of efficiency. Technological developments that do not improve life for everyone, but instead create or worsen inequalities and conflicts, cannot be called true progress. For this reason, AI should be placed at the service of a healthier, more human, more social and more integral development.” Pope Francis

The Pontifical Academy of Social Sciences held a workshop on Artificial Intelligence, Justice, and Democracy at the Casina Pio IV in Vatican City on March 4-5, 2025. This gathering responded to Pope Francis’s call for “algor-ethics” and a human-centered approach to technology, particularly AI. The workshop brought together leading experts from various fields to examine the profound impact of Artificial Intelligence (AI) on society, focusing on the administration of justice, the practice of democratic citizenship, and the dignity of work. The ethos and eidos animating the workshop echoed Pope Francis’s plea for an AI that respects human dignity, is human-centered, and is integrated into our care for our common home, “to live in harmony with nature, who respect it as a source of nourishment, a mutual home, and [an] altar of human sharing.” Pope Francis, Jorge Mario Bergoglio, Carlo Musso & Richard Dixon. “Hope.” Random House Publishing Group, 2025-01-14. Apple Books.

With the emergence of Artificial Intelligence (AI), we are at the cusp of grasping the promise of humankind's latest technological tools. The advancement of new large language models—such as OpenAI's o1 and others—continues to propel AI into known and unknown technological frontiers. Like all previous transformative technologies, AI creates opportunities while unleashing genuine and perceived threats. As with past global technological shifts, promise and peril coexist, necessitating thorough examination and careful guidance. AI holds significant potential, but its progress must not compromise citizen protections, data privacy, misinformation, or anti-discrimination efforts.

In the wise words of the Holy Father, "We cannot doubt that the advent of artificial intelligence represents a true cognitive-industrial revolution, which will contribute to the creation of a new social system characterized by complex epoch-making transformations. It will be for us to determine what direction the use of this fascinating—and at the same time terrible—instrument will take, an instrument even more complex than others, because it can adapt itself autonomously to the task that is assigned to it and, if designed in this way, can make choices independently of a human being in order to reach the pre-established goal. With Guardini, we must say that every problem of a technical, social, or political nature can be tackled and resolved 'only by man'. Pope Francis, Jorge Mario Bergoglio, Carlo Musso & Richard Dixon. "Hope." Random House Publishing Group, 2025-01-14. Apple Books.

As with other powerful human tools, AI is a product of human creativity and has the potential to improve work conditions, democratize access to knowledge, open new avenues in education, advance scientific research, aid medical advances, fight climate change, automate routine and complex tasks, and more. Yet Pope Francis highlights AI's duality, showing promise and concern.

We noted AI is evolving rapidly, with large language models showing exponential growth and significant improvement in accuracy and speed. Industries are quickly adopting AI to improve efficiency, productivity, and customer relations, potentially resulting in trillions of dollars in projected economic growth. However, AI raises concerns about fraud, misinformation, and bias. We noted how algorithms can contribute to confirmation bias and the possibility of AI selecting what information is delivered to whom, including when that information is unverified and/or news that lacks journalistic standards (such as deepfakes). This has negative implications for democratic citizenship, furthering a sense of unwillingness to engage in dialogue with those we disagree with and an overall lack of civility in the public square.

We also noted how AI education is essential for empowering individuals to understand, engage with, and influence the development of AI in ways that uphold fairness, equality, and dignity. We can equip people—regardless of socioeconomic background—with the skills to recognize algorithmic biases, advocate for transparency, and demand ethical AI governance. Furthermore, AI education fosters critical thinking and informed participation in policymaking, helping societies push for regulations that protect privacy, prevent discrimination, and ensure AI serves the common

good rather than the needs of the few. A well-informed global population can harness AI for social progress to combat misinformation, expand access to justice, and advance human rights protections in ways that uphold human dignity and ethical responsibility.

We recognize the need for intentional choices to protect the dignity of work, as AI poses a threat to hollow out and devalue entire categories of employment, both routine tasks and highly skilled tasks in professions such as journalism, medicine, law, and scientific research: the first fully AI-written papers have already passed peer review. Conversely, AI services will become more widely available and significantly cheaper. Intense competition among designers of large language models makes it unlikely that any single commercial provider will dominate these emerging markets. However, at the same time, human capital may be undervalued, even for individuals who have invested substantial amounts in their education. We noted three challenges: ensuring workers shape AI's design, ensuring all workers benefit from AI's opportunities, and protecting workers from AI's harms. We noted that the future of work will depend partly on whether AI is designed with and for workers, enhancing human dignity rather than diminishing it.

Qua AI incipient global regulatory mechanisms much remains *terra incognita.* Indeed, we noted the global regulatory challenges AI poses partly due to its rapid evolution everywhere and across nearly all domains of human activity. While emerging tools exist to build on, there is currently no international consensus. We noted, among other things, how the UK tends to favor a "light touch" and adaptable approach. While the EU has established a codified regulatory framework, the USA's risk-based federal approach is currently in flux with the new U. S. Administration. Ethical considerations are essential for guiding regulatory systems, protecting human dignity, and ensuring justice. Similarly, we observed the somewhat varied approaches among the South American countries represented at the Summit.

We recognize that once technology is developed, it does not retreat. Attempting to prevent the use of AI will only result in a competitive disadvantage by denying opportunities to learn and ceding progress to those who do use the technology. Instead of blocking the use of AI, it is imperative to establish guardrails that help guide its development and integrate ethical considerations. We also noted how applied AI can help deliver critical government services that improve the quality of people's lives.

We noted a series of endeavors to offer guardrails and regulatory regimes to maximize AI's promise and minimize its risks. We delved into the European Union's Artificial Intelligence (AI) Act, Regulation (EU) 2024/1689, which aims to create a consistent legal framework for developing, marketing, and using AI systems within the Union, aligning with EU values and fundamental rights. It seeks to promote trustworthy, human-centered AI, protect health, safety, and fundamental rights, and foster innovation while preventing market fragmentation. The Act introduces harmonized rules for high-risk AI systems, addressing issues such as data quality, transparency, and human oversight, and prohibits certain unacceptable AI practices, including manipulative or exploitative

techniques. It applies to AI system providers and deployers within and outside the EU if the system's outputs are utilized within the Union. However, it exempts AI employed for military, defense, or national security purposes. The Act establishes mandatory requirements for high-risk AI systems and encourages AI literacy to empower informed decision-making regarding AI's impact.

Ultimately, we noted that current regulatory systems are somewhat inchoate and incoherent and stressed the importance of ethical oversight and human accountability in AI development and implementation.

Core Principles

We affirmed the following core principles to guide the development and deployment of AI:

1. **Human Dignity:** AI systems must respect and promote the inherent dignity of every human being.
2. **Fairness and Equity:** Efforts must address and mitigate biases in AI algorithms and datasets.
3. **Transparency and Accountability:** Transparency in AI decision-making is crucial for public trust and accountability.
4. **Human Oversight:** Retaining human control over AI decisions is paramount, especially in life-altering situations.
5. **Inclusivity and Access:** Addressing the digital divide and ensuring equitable access to AI technologies is essential.

Key Concerns and Recommendations

The workshop participants acknowledged both the transformative potential and inherent risks of AI, emphasizing the need for proactive measures to address the following concerns:

1. **Misinformation and Manipulation:** Combatting AI-generated disinformation and manipulating public opinion is crucial.
2. **Erosion of Democratic Institutions:** Protecting democratic institutions from AI-driven propaganda requires careful consideration.
3. **Bias and Discrimination:** Addressing biases in AI systems, particularly in criminal justice and employment, is essential.
4. **Job Displacement:** Preparing citizens for an AI-driven world includes addressing ethical considerations and protecting displaced workers.
5. **Security & Safety:** An AI for peace must consider the security and safety risks posed by deploying LLMs and AI agents who can scale, escape alignment with human desires, and trigger existential risks for humanity, as the most qualified scientists creating such models have been warned.

Call to Action

The participants call for:

1. **Interdisciplinary Collaboration:** Fostering ongoing dialogue among experts from diverse fields.
2. **Ethical Frameworks and Regulations:** Developing and implementing ethical frameworks prioritizing human dignity.
3. **Education and Awareness:** Promoting education about AI's opportunities and risks.
4. **Global Cooperation:** Encouraging international collaboration in AI regulation.
5. **Continued Research and Evaluation:** Supporting ongoing research on AI's societal impact.
6. **Encouraging Responsible Use of AI:** Harnessing the power and possibility of AI to assist with addressing society's most vexing challenges and opportunities in an ethically responsible manner.

Conclusion

The Workshop reaffirmed the importance of placing human dignity and ethical considerations at the forefront of AI development and deployment. By embracing "algor-ethics" and promoting interdisciplinary collaboration, we can harness AI's potential while safeguarding justice, democracy, and the pursuit of truth for all. The spirit animating the workshop responded to Pope Francis's call for an AI that respects human dignity, is human-centered, and is integrated into our care for our common home, "to live in harmony with nature, who respect it as a source of nourishment, a mutual home, and altar of human sharing." Pope Francis, Jorge Mario Bergoglio, Carlo Musso & Richard Dixon. "Hope." Random House Publishing Group, 2025-01-14. Apple Books.

Expanded Insights

The Role of AI in Justice Systems

The workshop extensively discussed the integration of AI in justice systems, highlighting both potential benefits and risks:

1. **Efficiency and Access:** AI has the potential to streamline legal processes, making justice more accessible and efficient. However, this must not come at the cost of due process or individual rights.
2. **Predictive Policing:** While AI-driven predictive policing tools can enhance law enforcement efficiency, they risk perpetuating existing biases and over-policing certain communities.
3. **Judicial Decision Support:** AI systems can assist judges by providing relevant case law and precedents. However, the final decision-making authority must remain with human judges to ensure contextual understanding and ethical considerations.
4. **Bias in Sentencing:** Participants expressed concern over the use of AI in sentencing decisions, noting the potential for embedded biases to lead to unfair outcomes, particularly

for marginalized communities.

AI and Democratic Processes

The impact of AI on democratic processes was a key focus:

1. **Election Integrity:** AI can both enhance and threaten election integrity. While it can improve voter registration and detect fraud, it can also be used to create deepfakes and spread misinformation.
2. **Political Campaigning:** The use of AI in micro-targeting voters raises ethical questions about privacy and the potential manipulation of public opinion.
3. **Public Opinion Formation:** AI-driven recommendation algorithms on social media platforms can create echo chambers and polarization, affecting the quality of public discourse.
4. **Citizen Participation:** AI tools can enhance citizen participation in democratic processes but must be designed to be inclusive and accessible to all.

Ethical AI Development

The workshop emphasized the need for ethical AI development:

1. **Diverse Development Teams:** Ensuring diversity in AI development teams is crucial to mitigate biases and consider diverse perspectives.
2. **Ethical Training Data:** The importance of using accurate, ethically sourced and representative training data was highlighted to prevent perpetuating historical biases.
3. **Explainable AI:** Developing AI systems that can explain their decision-making processes is crucial, especially in high-stakes areas like justice and governance.
4. **Continuous Monitoring and Auditing:** Regular audits of AI systems for bias and unintended consequences are necessary to maintain their ethical operation.

Global Governance of AI

Participants stressed the need for global cooperation in AI governance:

1. **International Standards:** Developing international standards for AI development and deployment, particularly in justice and democratic processes.
2. **Cross-Border Data Sharing:** Establishing ethical frameworks for cross-border data sharing to enhance AI capabilities while protecting privacy and sovereignty.
3. **AI Diplomacy:** Promoting dialogue between nations to address potential AI-driven conflicts and ensure equitable access to AI technologies.
4. **Global AI Ethics:** Proposing a global framework for AI ethics to offer guidance on the ethical development and use of AI.

Concluding Remarks

Responding to Pope Francis's call, the Workshop on Artificial Intelligence, Justice, and Democracy concluded with a renewed commitment to ensuring that AI serves humanity's best interests. By addressing the challenges and harnessing the opportunities presented by AI, we can work towards a future where technology enhances justice, strengthens democracy, and upholds human dignity. The participants pledged to continue this vital dialogue and to translate these principles into actionable policies and practices in their respective fields and jurisdictions.

Signed

Cardinal Peter K.A. Turkson, PAS & PASS Chancellor	Sr Helen Alford, PASS President
Marcelo M. Suárez-Orozco, PASS Council	Gustavo Béliz, PASS
Christoph Engel, PASS	John McEldowney, PASS
The Hon. Roberto Andrés Gallardo, COPAJU	The Hon. Tamila Ipema, COPAJU
The Hon. Daniel Urritia Laubreaux, COPAJU	The Hon. Ana Algorta Latorre, COPAJU
The Hon. Gustavo Daniel Moreno, COPAJU	The Hon. María Julia Figueredo Vivas, COPAJU
The Hon. Ivonne Hao	The Hon. Jason Snyder
The Hon. Joe Kennedy, III	The Hon. David Lowy
Prof. Michael J. Ahn	Vipin Mayar
Molly Kinder	Rómulo Rubén Abregú
Elinay Almeida Ferreira	Angélica Aquino Suárez
Raúl Arroyo	Lisa L. Atkinson
Alberto Bastos Baleazeiro	Lenia Batres Guadarrama
Maximiliano Francisco Benítez	Edward Sidney Blanco Reyes
Emilia Bustamante Oyague	Ricardo Canales Herrera
Camilo Javier Cantero Cabrera	Teresa Cardenas Puente
Marta Cartabia	Hugo Cavalcanti Melo Filho
Jose Geraldo De Sousa Junior	Valeria De Los Angeles Diaz
Luis Duacastella Arbizu	Paul English
Almudena Fernandez	Genoveva María Ferrero
Marian Gaston	Gerardo Gutiérrez Gayosso
Pedro Hartung	Fabrizio Intonti

César Raúl Jiménez

Gabriela Lenz De Lacerda

Livia Cristina Marques Peres

Paul Nemitz

Roberto Carlos Pompa

Julissa Reynoso

Adriana Saavedra Lozada

Maria Dolores Sanchez Galera

Pamela Scott Washinton

Octavio Augusto Tejeiro Duque

Delio Antonio Vera Navarro

Ivana Wolansky

Ulises Augusto Yaya Zumaeta

James Julian

Joy Cossich Lobrano

Linda Strite Murname

Adriana Orocu' Chavarria

Ana Elizabeth Quinteros Castellanos

Claudia Lucía Rincón Arango

Lilian Graciela Samaniego González

Marcelo Sanchez Sorondo

Humberto Sierra Porto

Ananda Tostes Isoni

María Alejandra Villasur García

Rebeca Xicohténcatl Corona

Jesús Zuñiga González

© Fri Jun 13 15:44:49 CEST 2025 - The Pontifical Academy of Social Sciences