



Frankfurt Study Day
of Foundation Centesimus Annus Pro Pontifice (CAPP) - German Section
February 17th, 2024

BUILDING DIGITAL TRUST - HOW CAN ARTIFICIAL INTELLIGENCE SERVE THE HUMAN COMMUNITY?

from Dr. Ulrich Schürenkrämer and OPf. Dr. Christian Stenz

A. Preamble

A new hype has recently arisen around artificial intelligence (AI). The discussion has become particularly intense. Countless articles have warned of the dangers and risks of this new technology. The negative effects on the economy and society - e.g. in the form of dramatic job losses, discrimination, disinformation, manipulation, surveillance, use in war, etc. - are often described in apocalyptic images. The increasing commercialisation of Chat GPT, which was originally a non-profit organisation, has also inflamed emotions.¹ In this respect, the call for control of AI is becoming ever louder.

However, from a scientific and political point of view, it is not enough to steer the use of AI into orderly legal channels. Rather, the associated ethical and educational issues need to be taken into account in order to utilise technological progress holistically for human development.

Archbishop Vincenzo Paglia, President of the Pontifical Academy for Life, has called for an ethical framework for AI². The World Day of Peace 2024 will also focus on AI: the corresponding papal message is entitled "Artificial Intelligence and Peace" because - according to the Dicastery for the Service of Integral Human Development - advances in the field

¹ A.I. belongs to the Capitalists Now <https://www.nytimes.com/2023/11/22/technology/openai-board-capitalists.html>

² <https://www.vaticannews.va/de/vatikan/news/2023-09/kuenstliche-intelligenz-vatican-ethik.html>



of AI are having an increasingly profound impact on human behaviour, personal and social life, politics and economy.³

AI can bring a great deal of progress - but its risks must also be closely scrutinised, Pope Francis demands: "We must be aware of the rapid changes that are now taking place and manage them in such a way that fundamental human rights are safeguarded and the institutions and laws that promote integral human development are respected."⁴ Technological development is unstoppable. AI will become increasingly important. It should serve the best human potential and our highest aspirations, not compete with them.⁵

With a positive approach, the human perspective and the desire for a better future, AI will be examined below. How can AI contribute to a better order of human society, to growing freedom and fraternal community? How can we ensure that people act responsibly when using AI and that fundamental human values such as inclusion, transparency, safety, justice, confidentiality and reliability are respected? But where and how do we also need to draw red lines for the use of AI?

B. The technical perspective of artificial intelligence - data analysis and data literacy

Informed citizens, informed politicians and an informed administration seem to be a good basis for a functioning democracy.⁶ In fact, there has been exponential growth in data and AI for some time now. But "more data" does not automatically mean "more knowledge". Are AI applications a threat to democracy? How can privacy, transparency and informed politics be ensured in a democratic constitutional state, even in times of AI? The problem is that there are more and more new data sources - we need to find out where the "good" data is. Data is the "oil of the 21st century", it is initially - even if available in abundance - only raw material, it is about the value that can be gained from it to satisfy existing needs, for example the need for better decisions.

The beginning of AI can be dated back to the 1950s (the so-called Turing Test). According to the OECD definition, AI is "*a machine-based system that, for explicit or implicit objectives, infers from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments.*"⁷

³ see <https://www.vaticannews.va/de/papst/news/2023-08/papst-franziskus-botschaft-welt-frieden-kuenstliche-intelligenz.html>

⁴ <https://press.vatican.va/content/salastampa/it/bollettino/pubblico/2023/12/14/0884/01927.html>

⁵ *ibid.*

⁶ see as well as follows Schüller, Katharina, Presentation at Frankfurt Study Day on 17.02.2024 incl. question and answers

⁷ <https://www.oecd-ilibrary.org>



There are many opportunities in the application of AI, such as the application for fraud detection, the application in the detection of diseases (e.g. breast cancer), and AI can also be used to fight against hostile AI (see Russian Bots on Twitter). The context of AI use is always important. The distinction between intended and unintended use, compliance with fairness (cf. AI in applicant selection, for example, when only certain age groups are selected) must be taken into account.

The aim of "*ethics by design*" is to observe ethical principles as early as the development stage and to take into account how human empathy, for example, develops through AI (how does humanity change, for example?). So is AI about, for example

- more targeted therapies or countless misdiagnoses?
- helpful assistants or loss of autonomy?
- Fighting criminals or suspecting the innocent?

Not least from an ethical point of view, AI users need to be "*literate*" with regard to the responsible use of AI. Standards for digital intelligence must be developed: from the informed prosumer, to the skilled user, to the expert creator.

One standard for digital intelligence is IEEE 3527.1-2020 - Digital Intelligence (DQ): a bundle of technical, cognitive, meta-cognitive, socio-emotional competences, that are based on universal moral values and enable individuals to deal with the challenges and opportunities of digital life.⁸

The goal is data and AI literacy:

- the ability to generate, process, analyse, present meaningful information from data,
- develop, utilise and apply AI and related algorithmic tools and strategies,
- to guide informed, optimised, context-related decision-making processes.

However, it is important that the right mindset is developed alongside the toolset and skillset. Behind this are social norms, binding regulations and values:

Mindset 1: Scrutinise the quality of data and information sources (ultimately also in government, in order to pursue evidence-based policies).

⁸ <https://standards.ieee.org>



Mindset 2: Recognise manipulation (in order to counteract distorted data and framing of false data and visualisation).

Mindset 3: Separate facts from opinions (what should be measured and what should not be measured?)

Mindset 4: Accepting the limits of the informative value of data.

The aim of data and AI "*literacy*" is therefore to develop a "*compass*" of knowledge, skills and (value) attitudes. An important building block is skills development, starting at school. However, data protection is interpreted too rigorously in Germany. As a result, we are robbing ourselves of opportunities to utilise AI. "*Information hygiene*" is important, i.e. critical selection of data and criteria when searching for data.

C. The ethical perspective on artificial intelligence - data ethics

AI systems are trained with extremely large data sets (e.g. Chat GPT 4 with an estimated 1-2 trillion parameters). In order to obtain sufficient data, use is made of the data that is available (i.e. not necessarily released) in the network, including private information, location information, protected data and company data.⁹

One example from practice is the LAION-5B dataset. It has been available for Chat GPT since 2022 and contains 5.85 billion image data, including emails. It is currently the only comparable, publicly accessible data set for training AI models. The problem is that it also contains faces and names, geocoordinates and account numbers, and that links to images of child abuse were also used for AI training.¹⁰

When using AI, a self-test is now possible at <https://haveIbeentrained.com>. This can be used to check which data the AI has used for learning. AI has become an important topic in data ethics because there is little transparency regarding the training data of large companies such as Open.ai, Google, etc. It can be assumed that in some cases images of entire parts of the internet are created to train AI. Google, for example, changed its privacy policy in July 2023: all information posted online can be used to train large language models; this

⁹ see as well as following Schröder, André, Presentation at Frankfurt Study Day on 17.02.2024 incl. questions and answers

¹⁰ see <https://www.heise.de>: Bünthe, Oliver, "LAION-5B: Forscher entdecken Links zu Kindesmissbrauchsbildern"



also applies to Germany.

The EU AI Act is intended to contribute to regulation. An initial proposal from the European Commission was presented in April 2021. In December 2023, the negotiators of the European Parliament and the EU member states agreed on a preliminary version. The text still needs to be formally adopted by the EU Parliament and EU Council. The aim is human-centered and ethical development of artificial intelligence. There is a transparency obligation, a right to traceability and a right of appeal.

The EU AI Act distinguishes between different risk classes:

- Minimal risk (AI-supported video games, spam filters, etc.); no intervention is required here.
- Limited risk (deepfakes, chatbots, etc.); intervention here is primarily for the purpose of transparency (users should know when they are interacting with an AI).
- High-risk areas (transport, educational institutions, healthcare, security, legal, police, etc.); the aim here is to constantly assess risks, create activity reports in connection with high-quality data sets, create comprehensible and complete documentation and ensure human oversight measures.
- Unacceptable risks; these include social scoring, deliberate behavioural manipulation with a significant damaging effect, discrimination with a significant damaging effect, individual categorisation based on biometric data, facial and emotion recognition, which must then also be prohibited.

During the transitional period until the EU AI Act comes into full force, the European Commission's AI Pact applies, which provides for a voluntary commitment from companies that use AI. The subject of the voluntary commitment is compliant with the forthcoming AI Act and specific measures. The commitments will be collected and published by the Commission.

Google is in talks to participate in such a pact. There are also proposals from the tech industry. Open.ai founder Sam Altman, for example, has called for an AI regulatory body modelled on the International Atomic Energy Agency (IAEA) and warns against overregulation. And Google CEO Sundar Pichai could imagine a regulation similar to the Nuclear Non-Proliferation Treaty.



The Vatican has also developed proposals on the ethical framework for the use of AI.¹¹ The ITEC handbook focusses on the practical implementation of ethical standards in companies, while state regulation is only dealt with in a subordinate manner. Data ethics is not a central aspect of the handbook, but is addressed in various places. The ITEC handbook contains the following suggestions:

- "Anything to do with the collection and use of data sets raises privacy and security concerns. Algorithms, which are influenced by the organisation's business objectives, culture and the ethics of the people developing them, use data sets to actively make decisions that impact lives. Guardrails must be put in place to ensure that these powerful tools protect the well-being of all stakeholders." ¹²

- "Some digital surveillance technologies (e.g. facial and voice recognition) are being rapidly developed and disseminated without external oversight, posing significant societal risks. Around the globe, legislators and regulators are rushing to close this perceived gap. If industry leaders do not quickly create an actionable and reliable legal and ethical framework with proven effectiveness, they will quickly find themselves in a swamp of inconsistent and even contradictory laws and regulations." ¹³

- "At the same time, the private sector has no shortage of interest in ensuring that technologies are "*ethical by design*". Behind this is not only an intrinsic desire to act ethically, but also a desire to avoid reputational damage, create a corporate culture that attracts top talent and maintain the commitment of ethically minded employees. As part of this process many of the leading companies working on these novel technologies are playing a leading role in trying to shape the regulatory framework in a way that that it does not restrict innovation and does not favour companies operating in unregulated countries." ¹⁴

Factors for a sustainable use of AI could include

- use machines that heed our most important values: avoidance of harm in physical, psychological/emotional, economic, social, ecological and freedom terms; taking into

¹¹ see Ethics in the age of disruptive technologies, an operational roadmap, ITEC-Handbuch, Juni 2023

¹² see ITEC-Handbuch, 11-12

¹³ see ITEC-Handbuch, 15

¹⁴ see *ibid.*



account the safety needs of humans; decisions made by AI must not deviate radically from the moral decisions of human society.

- Controllability and protection against abusive behaviour: security standards and accountability (especially for sensitive data, but also for metadata); documentation and oversight.
- AIs should not use data against our will or interests.
- Maintain trust and respect privacy: transparency in the use of AI (cf. chatbots: am I talking to an AI or a human?); transparency in decisions made by AI (e.g. rating or scoring systems for lending, insurances, Schufa); only use and collect the data that is required for the functionality of the AI; loss of trust can have a lasting impact on the market maturity of products (e.g. autonomous driving).
- Awareness of disruptive effects: what impact does AI have on our work processes and areas of work, for example? Will the use of AI lead to a loss of trust in the workforce? (safety of workplaces and processes)

To summarise, it can be said that the interests of companies and customers are less divergent than is often assumed. Good results alone are not sufficient for the ethical use of AI. Companies should not undervalue factors such as security and damage prevention, transparency, supervision and documentation, data minimisation, customer interests and privacy protection, also in their own interests, but should implement them in the best possible way. Companies depend on customers having confidence in their product. Employees can react sensitively to disruptive change processes. AI can serve to support the compliance and dissemination of our moral values, but these should not be undermined by AI at any time.

With regard to the demands of the encyclical *Laudato si'*¹⁵, we must not lose sight of the environmental costs. The environmental costs of data storage and training alone are immense. The so-called clouds require many tons of metal, plastic and rare earths, which consume vast amounts of water and electricity.

¹⁵ see Pope Francis, encyclical *Laudato si'*, (LS, 6 ff.), 2015



D. Perspective of the state and the economy from the point of view of Catholic Social Teaching

AI creates a dilemma structure, as it offers both advantages and disadvantages. The question is which ethics should be applied here. Some see AI as a "*black box*" - how should we take a stand on this? Data is money - the question arises as to who earns the money and whether the producers should be authorised to charge fees (cf. GEMA fees for the public use of musical works in Germany). Another problem is that fact-checking is too slow compared to the speed of spreading fake news. Corrections are usually no longer recognized.

Against this backdrop, it must be ensured at a collective level (state, economy) that AI must not reduce the personal responsibility of the individual. Legal regulations must also regulate what should happen in the event of violations. Binding legal regulations with minimum requirements for the permissibility of data sources should also be created (legislation). As AI has a strong international focus, an international charter for data use ("Data Literacy Charter") would be highly beneficial. Data misuse or the improper dissemination of data, in particular the falsification of data, should be counteracted by the legislator with sanctions. The awarding of "*labels*" combined with recommendations for reliable data sources should also be considered.

E. Individual perspectives from the point of view of Catholic Social Teaching

A sound general humanistic education is essential for a conscious approach to AI. From an individual perspective, it is essential to strengthen both the ethical and technical education of the individual. This should start as early as possible (at home, daycare centre, school), developing the *digital naive* into the *digital native*, so to speak. But be careful: self-learning systems make the user passive: he thinks he is an actor however he is only the data milk cow that is milked.

Ethical education is based on the foundations of the constitution and the laws of the constitutional state, as well as on universal values, in particular the dignity of human beings. This arises from their creatureliness and is inviolable, also and in particular by the state or economic actors. It is expressed, among other things, in the UN Charter of Human Rights. This means that the principle of personality must be strictly upheld when using AI.

In the case of technical education, which should also begin as early as possible, at school, continuing professional development is also of great importance. The principle of "*lifelong*



learning" also applies here. For example, a "*driving licence*" for data literacy, combined with regular updates, should be considered.

All these efforts in the field of education are intended to enable a critical approach to data. The greatest danger is the lack of competences, in the sense of *savoir* and *pouvoir*, i.e. knowledge, skills and abilities. To this end, we need to create space for reflection in our own mental system. Appropriate time is needed when selecting data, as well as special attention, particularly with regard to data sources and their reliability. It is helpful to exchange ideas with others. Social media want to use AI to tie up the individual's time and attention, and the individual must be aware of this and act against it. At the same time, the further development of individual talents, which are unique and cannot be replaced by AI, is important and urgent. AI has no sensitivity for transcendence, no soul and no love.

The document reflects the discussion contributions of the participants of the Frankfurt Study Day of the Centesimus Annus pro Pontifice Foundation - German Section on 17 February 2024, which resulted under the moderation of Fr. Prof. Dr. Elmar Nass, Prorector at khkt, the University of Catholic Theology in Cologne, and Professor for Christian Social Sciences and Societal Dialogue, from the keynote speeches by Katharina Schüller, CEO, Stat-up GmbH, "The technical perspective of artificial intelligence - data analysis and data literacy" and André Schröder, Catholic Social Institute, "The ethical perspective on artificial intelligence - data ethics".

Dr. Ulrich Schürenkrämer is Managing Director of Machlaan GmbH, Munich and Coordinator for Germany for Centesimus Annus Pro Pontifice Foundation, Vatican City.

Senior Priest Dr. Christian Stenz is Pastor to the Federal Police, Berlin.

Other participants at the study day were Dr. Cornelia Agel, Dr. Eleonora Bonacossa, Rainer Dormagen, Andreas Hecker, Dr. Christopher Klein, Dr. Martina Köppen, Prof. Dr. Alexander Kracklauer, Volker Lauven, Ingo Ley, Barbara Pung, Sabine Rodeck-Kemper, Dr. Oliver Roethig, Sebastian Sasse, Dr. Alexander A. Stummvoll, Sylvia Trimborn-Ley and Father Johannes H. Zabel OP.