In Dengg (ed.): ICARUS' WINGS – Navigating Human Enhancement.
Schriftenreihe der Landesverteidigungsakademie No. 3/2025.

European AI ethics – between categorical imperative and placebo rhetoric

Gernot Stimmer

1. The unlimited scope of AI

For many observers, the "uncanny" expansion of the use of AI in all areas of life is an undisputed fact that has been viewed positively from the outset, both from a scientific-technical and economic-political perspective, but has been the subject of increasingly critical debate in recent years. The mainstream of heated debate is currently moving between the extreme positions of unconditional apology (Larry Page: AI as a "digital God") and hysterical apocalyptic sentiment (Dan Hendryck's "Against the End of the World"), with only one thing in common: the call for a global or at least regional AI order. This article is primarily dedicated to recording and comparing the various regulations that have been agreed on a global-international, transnational and nation-state level, without the multitude of additional regulations that have been and are being established by civil society actors. The exclusivity with which the media publicly uses the keyword "artificial intelligence", usually abbreviated to the magic cipher "AI", should not obscure the fact that this is only the most effective digital form of a general technical-scientific boom to improve the human way of living and working. The new magic formula "human enhancement" (HE) can be used to summarise the wealth of instruments, methods and inventions in the field of Human Enhancement.[1]

In the following, we will therefore limit ourselves to global, transnational or regional regulations (conventions, resolutions, standards, recommendations) of international actors with varying degrees of binding force that relate exclusively or at least in part to the application or restriction of new AI technologies and their respective ethical legitimation.

[1] The definition of HE: Wikipedia: Human Enhancement, last modified 24 November 2024 at 11.32 PM (UTC). https://en.wikipedia.org/wiki/Human_enhancement.

1. *The OECD AI Principles*

The declaration of principles adopted by the 36 member states at the Ministerial Council Meeting on 22 May 2019 was also adopted by six other countries, especially Latin American countries, and includes five principles and five recommendations to governments. The Davos World Economic Forum's efforts to regulate AI at global level are also closely related in terms of programme and ideology. The AI Governance Alliance (40 (predominantly) OECD countries, academia and global companies) was founded in June 2023 in an initial step. However, the hoped-for agreement at the 54th Annual Meeting of the World Economic Forum in Davos did not materialise, which is not surprising given the current 700 national or private AI regulations.[2] The international AI agreements adopted by the G7 (6 OECD countries & Japan) in February 2024, which build on the G7 Hiroshima Summit of May 2023 and include the "Hiroshima Process International Guiding Principles for All AI Actors" and the "Hiroshima Process International Code of Conduct for Organizations Developing Advanced AI Systems", are also strongly oriented towards economic policy.[3]

2. *The UNESCO Recommendation on the Ethics of Artificial Intelligence of 23 November 2021*

The Recommendation, adopted by the 193 members of UNESCO in 2021, is based on the results of multi-stakeholder workshops in 25 countries and provides for evaluation based on national reports on the implementation of the Recommendation, to be submitted every four years. A Global AI Ethics and Governance Observatory has been established to improve the sharing of experiences, to assess the capacity of member states to implement the Recommendation and to develop an ethical impact assessment tool for AI systems. In order

---

[2] OECD: Recommendation of the Council on Artificial Intelligence. OECD/LEGAL/0449; OECD: Artificial Intelligence in Society. OECD Publishing, Paris 2019. https://doi.org/10.1787/eedfee77-en, accessed 20 April 2024.

[3] KIZUNA: The Hiroshima AI Process: Leading the Global Challenge to Shape Inclusive Governance for Generative AI. 9 February 2024. https://www.japan.go.jp/kizuna/2024/02/hiroshima_ai_process.html.

to improve the interoperability of the different AI regulations in Europe, the UNESCO National Commissions of Austria, Germany and Switzerland hosted an online event in the form of a UNESCO Talk: "Ethical AI in Europe", presenting the translation of the UNESCO Recommendation on the Ethics of Artificial Intelligence agreed between the countries.[4]

3. *The "AI Ethics" (Rome Call) of January 2023*

This was agreed at the conference in Rome between the Vatican, representatives of the Muslim and Jewish religious communities and global companies (Microsoft, IBM, etc.).[5]

4. *The NATO AI Strategy adopted on 21 October 2021*

In contrast to the other civilian organisations compared and their AI policies, NATO is a purely military-security alliance of states whose understanding of AI goes far beyond the purely military sphere and includes all issues and aspects of "dual-use technologies" such as

[4] Austrian Commission for UNESCO: Ethik in Wissenschaft und Technologie. Gesellschaftliche Folgen antizipieren. https://www.unesco.at/wissenschaft/wissenschafts-und-bioethik/ethik-kuenstliche-intelligenz, accessed 20 April 2024; UNESCO: Regierungen müssen generative KI in Schulen schnell regeln. 7 September 2023. https://www.unesco.de/wissen/ethik/kuenstliche-intelligenz/regierungen-generative-ki-schulen-altersgrenze-13-jahre, accessed 20 April 2024; Güell Paule, Laia: UNESCO launches Global AI Ethics and Governance Observatory at the 2024 Global Forum on the Ethics of Artificial Intelligence. 6 February 2024. https://digital-skills-jobs.europa.eu/en/latest/news/unesco-launches-global-ai-ethics-and-governance-observatory-2024-global-forum-ethics.

[5] Vatican News: Pope Francis to participate in G7 session on AI. 26 April 2024. https://www.vaticannews.va/en/pope/news/2024-04/pope-francis-g7-summit-italy-artificial-intelligence.html; Ring-Eifel, Ludwig: Digitale Ethik Vatikan will beim Thema Künstliche Intelligenz mitreden. 23 January 2023. https://www.herder.de/hk/aktuell/vatikan-will-beim-thema-kuenstliche-intelligenz-mitreden.

quantum and biotechnology, data protection, cyberspace, etc.[6] In February 2023, the NATO Data and Artificial Intelligence Review Board created a certification standard for the user-friendly and responsible use of AI for both the military and industrial sectors in order to implement this AI regulation.

5. *The UN Resolution on the promotion of "safe, secure and trustworthy" artificial intelligence (AI) systems of 21 March 2024*

Based on a resolution from November 2021 on the creation of global standards for AI ethics, the proposal submitted by the US was adopted as a non-binding resolution on 21 March 2024 by 193 states following negotiations with 120 member states. The main objective was formulated as the creation of international standards for the global use of AI. An AI committee convened by the UN Secretary-General drew up seven proposals for the "Global Digital Compact" to be adopted at the UN Summit of the Future at the end of September 2024, including the establishment of a scientific advisory board, a global fund financed by private and public money to remedy the asymmetric development of AI technology, a global database for AI training data, permanent dialogue between states and private stakeholders and a dedicated UN AI office. The Summit of the Future, held during the UN General Assembly on 22-23 September 2024, adopted a Pact for the Future and a Declaration on Future Generations, as well as the "Global Digital Compact", which only pro-

[6] Frank, Dorothea: NATO-Zertifizierungsstandard für Künstliche Intelligenz in Behörden. In: Spiegel. 15 February 2023. https://www.behoerden-spiegel.de/2023/02/15/nato-zertifizierungsstandard-fuer-kuenstliche-intelligenz/; Cheung, Sunny; Wettrennen um KI: Chinas Volksbefreiungsarmee und die NATO. In: Friedrich Naumann Stiftung Analyse. November 2023. https://www.freiheit.org/de/taiwan/wettrennen-um-ki-chinas-volksbefreiungsarmee-und-die-nato; North Atlantic Treaty Organization: Summary of the NATO Artificial Intelligence Strategy. 22 October 2021. https://www.nato.int/cps/en/natohq/official_texts_187617.htm; Hauser, Gunter: Die NATO. Die Grundlage des europäisch-atlantischen Sicherheitsverbundes. In: Hochleitner, Erich P. (ed.): Das Europäische Sicherheitssystem zu Beginn des 21. Jahrhunderts. Böhlau Verlag 2000, pp. 267-336; Kleinwächter, Wolfgang: Entwicklungen im Internet Governance-Umfeld April bis Juli 2024. In: DENIC Internet Governance. 14 August 2024. https://blog.denic.de/entwicklungen-im-internet-governance-umfeld-januar-bis-marz-2024/.

claimed very general principles for an "open, free and secure digital future" for all people. After three failed drafts, the power struggle between governments (multilateralism) and large digital companies (stakeholderism) was reflected in the anticipated "soft" proposals. For institutionalisation, a special scientific advisory board and permanent dialogue between governments and relevant civil stakeholders should be established.[7]

6. *The Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law of March 2024*

On 15 March 2024, the Council of Europe's Committee on Artificial Intelligence (CAI) presented a framework convention on fundamental principles and standards for the compatibility of the development of AI-based systems with the principles of human rights, democracy and the rule of law. This document, which is binding upon its 46 members, was negotiated with the participation of the US and the EU. Following adoption of the draft by the Council of Europe's

[7] Research Institute AG Wien: Vereinte Nationen und Europarat beschließen Regelwerke zu künstlicher Intelligenz. 28 March 2024. https://researchinstitute.at/vereinte-nationen-und-europarat-beschliessen-regelwerke-zu-kuenstlicher-intelligenz/; Fokuhl, Josefine: UN-Gremium legt sieben Vorschläge zur Steuerung von KI vor. In: Handelsblatt. 19 September 2024; UNRIC-Regionales Informationszentrum der Vereinten Nationen: UN-Zukunftsgipfel. https://unric.org/de/un-system/un-zukunftsgipfel-summitofthefuture/, accessed 30 October 2024; Europäische Kommission: Gestaltung der digitalen Zukunft Europas. NEWS ARTICLE, 23 September 2024. https://digital-strategy.ec.europa.eu/de/node/13020; Oswald, Fiene: Erste UN-Resolution zu Künstlicher Intelligenz. 2 April 2024, pp. 1-6. https://www.basecamp.digital/ki-verstehen-erste-un-resolution-zu-kuenstlicher-intelligenz/; Österreichische UNESCO-Kommission: UNESCO Talk: Ethische KI in Europa. 23 June 2023. https://www.unesco.at/wissenschaft/artikel/article/unesco-talk-ethische-ki-in-europa; United Nations, Office for Digital and Emerging Technologies: Global Digital Compact. 15 October 2024. https://www.un.org/global-digital-compact/sites/default/files/2024-09/Global%20Digital%20Compact%20-%20English_0.pdf.

Committee of Ministers in the summer of 2024 and ratification by member states, the treaty will be open for accession by other states.[8]

7. *The EU Artificial Intelligence Act of 13 March 2024 (AI Act)*

The proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence, submitted by the European Commission in April 2021, was finally adopted by the European Parliament on 13 March 2024 and by the Council of the European Union in May 2024 after three days of trilogue negotiations, following the opinions of the Council of the European Union and the European Parliament, and thus entered into force on 21 May 2024. Implementation will be staggered until May 2027 (6 months for banned AI systems, 24 months for generative AI, 36 months for high-risk AI systems).[9] The European Commission-

---

[8] Council of Europe, Secretary General Marija Pejčinović Burić: Artificial Intelligence, Human Rights, Democracy and the Rule of Law Framework Convention. 15 March 2024. https://www.coe.int/en/web/portal/-/artificial-intelligence-human-rights-democracy-and-the-rule-of-law-framework-convention; Council of Europe: European Convention on Human Rights. As amended by Protocols Nos. 11 and 14 Supplemented by Protocols Nos. 1, 4, 6, 7, 12, 13 and 16. September 2019. https://edoc.coe.int/en/european-convention-on-human-rights/8363-european-convention-on-human-rights-as-amended-by-protocols-nos-11-and-14-supplemented-by-protocols-nos-1-4-6-7-12-13-and-16.html; Köver, Chris: KI-Konvention des Europarats. Viel Abkommen um Nichts. In: Netzpolitik Org. 20 March 2024. https://netzpolitik.org/2024/ki-konvention-des-europarats-viel-abkommen-um-nichts/16.

[9] European Commission: Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules On Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts. 21 April 2021. https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52021PC0206; European Parliament: EU AI Act: first regulation on artificial intelligence. The use of artificial intelligence in the EU will be regulated by the AI Act, the world's first comprehensive AI law. Find out how it will protect you. 8 June 2023 (Last updated: 18 June 2024). https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence; European Commission: Artificial Intelligence – Questions and Answers. In: Questions and Answers, 1 August 2024. https://ec.europa.eu/commission/presscorner/detail/en/qanda_21_1683; KPMG International: Das bedeutet das EU-Gesetz zur KI. 10 May 2024. https://kpmg.com/de/de/home/themen/2024/03/das-bedeutet-das-eu-gesetz-zur-ki.html; Council of the EU: Artificial Intelligence Act: Council calls for promoting safe AI that respects fundamental rights. 6 December 2022. https://www.co

funded SIENNA project, which collects AI regulations and codes of conduct, lists 18 AI regulations and recommendations for behaviour based on ethical values for 2024, ranging from the SHERPA project 2019 to the ACM Code of Ethics and Professional Conduct of the Association for Computing Machinery, and from the Barcelona Declaration for the Proper Development and Use of Artificial Intelligence in Europe to the Montréal Declaration for a Responsible Development of Artificial Intelligence 2023. This civil society engagement is not discussed separately below.[10]

nsilium.europa.eu/en/press/press-releases/2022/12/06/artificial-intelligence-act-council-calls-for-promoting-safe-ai-that-respects-fundamental-rights/; Hilbricht, Benjamin: EU Trilog zur KI-Verordnung endet mit Kompromiss. In: *Behörden Spiegel*. 22 December 2023. https://www.behoerden-spiegel.de/2023/12/22/eu-trilog-zu-ki-verordnung-endet-mit-kompromiss/; Milicic, Milka/Zimmermann-Gassner, Carla: WEKA Business Solutions GmbH. Die neue KI-Verordnung in der EU. 7 March 2024. (Updated 01.08.2024). https://www.weka.at/news/Datenschutz-IT/Die-neue-KI-Verordnung-ist-seit-01.08.2024-in-Kraft?searchtext=1164787; Meltzer, Josh/Tielemans, Aaron: *The European Union AI Act. Next steps and issues for building international cooperation, in Global Economy and Development at Brookings*. Policy Brief, May 2022. https://www.brookings.edu/wp-content/uploads/2022/05/FCAI-Policy-Brief_Final_060122.pdf. Meltzer, Josh/Tielemans, Aaron: Comment: The EU AI Act: its international implications for AI policy development and cooperation, in Encompass. September 2022. https://encompass-europe.com/comment/the-eu-ai-act-its-international-implications-for-ai-policy-development-and-cooperation; Noerr: KI-Verordnung (AI-Act) final verabschiedet. In: *News*. 21 May 2024. https://www.noerr.com/de/insights/ki-verordnung-ai-act-final-verabschiedet; European Commission: European AI Office. Accessed 30 October 2024. https://digital-strategy.ec.europa.eu/de/policies/ai-office.

[10] SIENNA: AI & Robotics: Codes and guidelines. Last updated 19 April 2024. https://www.sienna-project.eu/w/si/robotics/; The SIENNA Project: AI & Robotics: Codes and guidelines. Accessed 15 July 2024. https://www.sienna-project.eu/w/si/robotics/codes-and-guidelines/; ACM Code of Ethics and Professional Conduct. Accessed 15 July 2024. https://www.acm.org/code-of-ethics; Université de Montréal: The Montréal Declaration for a Responsible Development of Artificial Intelligence. Accessed 15 July 2024. https://montrealdeclaration-responsibleai.com/the-declaration/.

1.2 Comparison of AI regulations

(For a comprehensive overview of Artificial Intelligence policies depicted in this contribution, refer to the ***paper insert*** *titled* ***"AI Regulations in Comparison"*** *which is enclosed with this publication.)*

Comparing global (UN, UNESCO, Rome Call), transnational (OECD, NATO) and regional (EU, Council of Europe) AI regulations according to criteria

1. Definition and impact assessment (positive vs negative) of AI
2. Scope of social and political areas
3. Ethical support
4. Restrictions, prohibitions
5. Control, instruments

These show a number of similarities, but also significant differences:[11]

1.3 Definition and assessment of the impact

An internationally applicable definition of AI is yet to be developed, despite the UN calling for global standards for AI ethics in November 2021 and UNESCO producing a guide to AI ethics. In the course of the EU's AI legislation, this call has also been echoed by the European Parliament (a single "technology-neutral" definition of AI) and the Council of the European Union. In its opinion of November 2021, the latter still assumed a technocratic understanding of AI, limited to "machine learning systems and concepts and logical-scientific concepts". Article 3 of the final EU AI Act, on the other hand, introduced a very broad definition in order to be compatible with other AI regulations:

> "software that is developed with one or more techniques and approaches listed in Annex I and can for a given set of human defined objectives, generate outputs such as content, prediction, recommendations of decisions influencing the environments they interact with"[12]

---

[11] All data and facts cited (except for sources specifically cited in the text) are from the literature cited above.

[12] Meltzer, Josh/Tielemans, Aaron: Comment: The European Union AI Act. Next steps and issues for building international cooperation. In: B/Global Economy and Development at Brookings. Policy Brief May 2022, https://www.brookings.edu/wp-content/uploads/2022/05/FCAI-Policy-Brief_Final_060122.pdf.

Behind this battle for sovereignty of definition lies the EU's clear interest in dominating the global AI market through joint AI regulations with other countries, especially those outside Europe. For example, the EU is working with the UN on the creation of common AI rules, as well as with the Council of Europe via its members and with the largest countries in Latin America, which are strongly orientated towards the EU in their AI policy.[13] The cooperation negotiations within the framework of the joint EU-US Trade and Technology Council (TCC) on a more risk-based development of safe AI technology on the basis of the "AI for Public Good EU-U.S. Research Alliance in AI for the Public Good" are much closer, but also more controversial. Since November 2023, cooperation between the USA and the UK has become even closer and more specialised, leading to a memorandum of understanding on the development of joint AI safety tests in April 2024.[14]

In contrast, the fundamentally positive assessment of AI systems appears to be more consistent. UNESCO highlights the benefits in education, science and research, while UN Resolution 2024 sees AI technology as a lever for non-discrimination against countries and societies in the "global south" and for achieving the sustainability goals of the "Agenda 2030". The OECD's more economic assessment of AI is mainly based on economic growth, thereby securing jobs and generally increasing prosperity. This view of the organisation as a whole is reinforced by the almost euphoric assessment of individual member countries. A case in point is the OECD's report on artificial intelligence in Germany, which is based on the claim that Germany could win the international race to implement trustworthy AI with its "hu-

---

13 Li, Cathy: Fourth Industrial Revolution. UN and EU both agree new AI rules, and other digital technology stories you need to know. *World Economic Forum*. 8 April 2024 (Updated 10 September 2024). https://www.weforum.org/agenda/2024/04/artificial-intelligence-technology-news-april-2024/.

14 Bundesministerium für Bildung und Forschung: Kooperation international. *Berichterstattung weltweit*. 8 April 2024. https://www.kooperation-international.de/aktuelles/nachrichten/detail/info/usa-und-vereinigtes-koenigreich-vereinbaren-zusammenarbeit-zu-sicherer-kuenstlicher-intelligenz; World Economic Forum. Accessed 10 May 2024. https://www.weforum.org/agenda/2024/04/artficial-inteligence; EGA: Artificial Intelligence: Latin America's Regulatory and Policy Environment. 4 March 2024. Accessed 15 July 2024. https://www.edelmanglobaladvisory.com/insights/artificial-intelligence-latin-americas-regulatory-and-policy-environment.

man-centred focus" and the involvement of civil society, using the EU as a lever.[15]

NATO's AI Strategy is based on four objectives:

(1) to develop AI responsibly for defence;
(2) to improve interoperability between member nations;
(3) to protect its own AI technology from hostile attack (disinformation campaigns); and
(4) to maintain the West's technological edge.

To this end, NATO relies on a network of sites and extensive cooperation with the private sector and academia, coordinated by a "Defence Innovation Accelerator for the North Atlantic" (DIANA) and a dedicated Science and Technology Facility supported by the NATO Innovation Fund.[16] The Council of Europe, on the contrary, sees the compatibility of AI technology with the principles of democracy, the rule of law and human rights as the goal of the Framework Agreement.

2. Sociopolitical areas

With regard to the scope of AI technology, all AI organisations assume that AI will become a fact of life that will change all areas of life and thus politics in the future, which can hardly be avoided and must therefore be countered with the most comprehensive normative regulation possible. With its rather general AI principles, the OECD specifically targets the protection of privacy and trust in corporate governance in the use of personal data, which is obviously in the interests of the big-data corporations in the US. UNESCO lists eleven policy areas to be covered by the AI regulation, in particular, development cooperation, education, culture, communication, health, work and gender equality. The UN Resolution also emphasises the general accessibility of AI technology through the creation of global structures. NATO also sees the area of application of AI technology beyond the military sector, specifically in the private sector (dual-use technologies). The EU AI Act foresees positive effects for a single AI market, but emphasises the risk-based protection of fundamental rights threatened by AI in different areas of life, from data protection to health and the environment. The AI Act thus becomes a

---

[15] OECD Berlin Centre: *OECD-Bericht zu Künstlicher Intelligenz in Deutschland*. 6 June 2024. (Online edition, OECD-Publishing, Paris). https://doi.org/10.1787/8fd1bd9d-de.

[16] Cheung: Wettrennen um KI. Online ibidem.

cross-cutting regulatory instrument relevant to all areas of competence. The restriction of the areas affected by AI in the Council of Europe appears to be significant. The definition of the affected areas of life and policy led to a dispute with the US, which participated in the negotiations and succeeded in completely excluding the military and security sector and de facto excluding the private sector (i.e. including private companies) from the regulation with a mere opt-in clause for signatory states. The "ground-breaking text", hailed as a success by the Secretary General of the Council of Europe, authorising signatory states to regulate the "private sector" at national level, was strongly criticised by NGOs as a "free pass for corporations and security agencies". This tendency towards "soft" formulation also characterises the other parts of the Convention.[17]

2.1 Ethical justification

All of the AI regulations compared refer to ethical principles that can be derived from the various codifications of fundamental and human rights. However, it is not possible to conclude from this that AI ethics are already being applied globally. The global AI concepts of the UN and UNESCO refer to the principles of the 1948 UN Declaration of Human Rights (including the eight additional conventions), but do not constitute a legally binding and therefore enforceable legal basis. Furthermore, the quoted UNESCO principles have a particularly flexible and dynamic interpretation within the meaning of the definition of "ethics" of the AI 2019: The formulation in the German language allows for constant relativisation and changeability of the ethical standards.[18] This soft ethical formulation also explains why UNESCO's 2019 AI Recommendation was signed by all 193 member states, and although the March 2024 resolution emphasises the protection of fundamental and human rights, including online, the AI panel's proposals for the Global Digital Compact (GDC) see the danger of AI use, mainly in potential electoral manipulation and disinformation. The GDC, which was adopted by the UN General Assembly, echoes the call for the UN Declaration of Human Rights to be applied online and sees acute violations, partic-

[17] Köver: KI-Konvention des Europarates. ibidem.

[18] Ethics of AI: "…eine systematische normative Reflexion, basierend auf einem holistischen, umfassenden, multikulturellen und sich weiter entwickelnden Gerüst wechselseitig voneinander abhängiger Werte, Prinzipien und Handlungen…", in UNESCO: Ethik in Wissenschaft und Technologie, Gesellschaftliche Folgen antizipieren. ibidem.

ularly in the exclusion of one third of the world's population from internet access. The GDC also criticises the use of AI in security and military settings (in particular, autonomous weapons systems) as violating human rights without ensuring human accountability.[19]

The ethical concepts of the (Western) industrialised countries in the OECD, the Council of Europe and the EU are based on regionally applicable codifications of fundamental rights, such as the European Convention on Human Rights of 1950, including the additional protocols of the Council of Europe, and the EU Charter of Fundamental Rights, enshrined in primary law in 2009 (Art. 2, Art. 6 TEU). The AI ethics of the non-European members of the OECD are based on UNESCO Recommendation 2019 and UN Resolution 2024. The position of the Vatican is interesting: at the G7 Summit in Italy in June 2024, the Pope called for an interdenominational and interdisciplinary "ethics of algorithms" in the sense of "digital anthropology", based on Pope Francis' encyclical "Fratelli Tutti" on digitality and the "Rome Call for AI Ethics" drawn up by the Pontifical Academy, other religious communities and IT companies in February 2020. NATO has also established principles for the responsible use of AI technology, based on generally accepted ethical, legal and political commitments, as well as six specific principles for the use of AI in defence, of which the fifth is particularly worthy of a mention here: controllability of operationally necessary human-machine interaction, an aspect that will be addressed in more detail below in the context of the development of unmanned autonomous weapon systems.[20]

2.2 Restrictions/Prohibitions

The negative impact of AI technologies, in particular, the threat to fundamental rights and democracy, is addressed in all AI policies, but the countermeasures mostly amount to general recommendations (notably the OECD's

---

[19] Reiland, Patrick: United Nations. Unpacking the Global Digital Compact. The Intersection of Human Rights and Digital Governance. *Friedrich Naumann Foundation for Freedom*. 19 September 2024. https://www.freiheit.org/human-rights-hub-geneva/unpacking-global-digital-compact.

[20] Communication of the European Union: Consolidated version of the Treaty on the Functioning of the European Union Official Journal 115. 9 May 2008. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:12008E/TXT, 17, 19; Ring-Eifel, Ludwig: KANN. 23.1.2023. Digitale Ethik. Vatikan will beim Thema Künstliche Intelligenz mitreden. Online ibidem; Vatican News: Papst Franziskus nimmt an G7-Sitzung zu KI teil. Online ibidem; NATO: Summary. Online ibidem.

"Should..." principles) to ban or not allow risky AI technologies, with proposals that are sometimes unrealistic (UNESCO: age limit of 13 for AI applications). The UN's Global Digital Compact calls on AI producers to develop appropriate technologies that protect human rights. This basic attitude towards "trustworthy" AI technology is only qualified by the EU's AI Act, which, despite a positive assessment of a future internal AI market, adopts an emphatically risk-based position. The comprehensive understanding of risk is divided into four risk levels according to the degree of risk:[21] (see Figure 1, next page):

(1) **Minimal risk**: "trusted" AI systems that are not subject to the obligations of the AI regulation, but can voluntarily adopt the requirements and codes of conduct. This includes "open source" AI systems.

(2) **Low risk**. AI systems with a risk of manipulation are subject to a transparency obligation (e.g. chatbots).

(3) **High risk**: AI systems with the potential to jeopardise fundamental rights (Art. 1, 8, 11, 12, 21, 29, 47, 48) are listed in two updated annexes (product use cases).
These are divided into two categories:
   a. systems covered by EU manufacturing legislation (toys, aeronautics, medical devices, etc.)
   b. AI systems in specific sectors: infrastructure, education, employment, public and private services, law enforcement, migration, asylum policy, etc.

These high-risk AI systems are subject to a number of compliance obligations, including risk and compliance assessments, data governance, basic human oversight and cybersecurity. These requirements also apply to general-purpose AI systems.
In the event of a violation of fundamental rights, there is a right of appeal to national market surveillance authorities.

(4) **Inacceptable risk** with seven exhaustively listed prohibited applications of AI (ranging from the manipulation of cognitive behaviour and Social Scoring to biometric real-time remote identification and also "subliminal influence").

[21] News: KI-Verordnung final. Online ibidem.

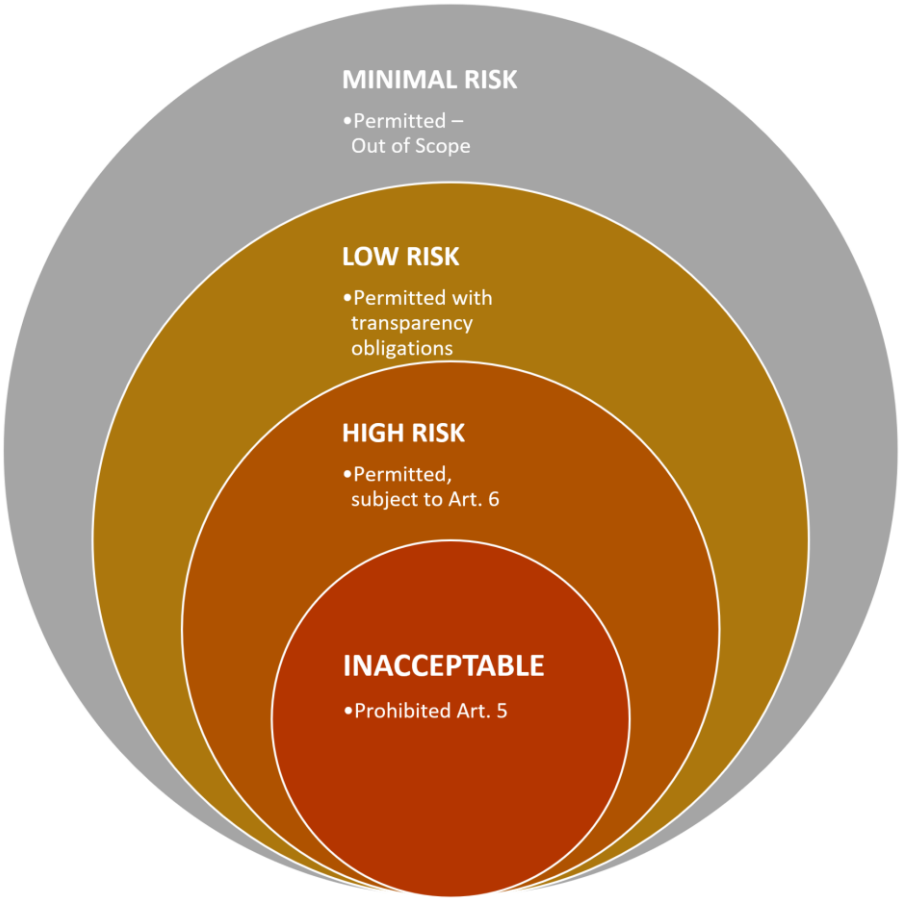


Figure 1: Authors compilation. Source: Eylaw-Ar: AI Act Der EU: Was Start-Ups Rechtlich Wissen Müssen! *EY Law* (blog). 29 July 2024. https://www.eylaw.at/ai-act-der-eu-was-start-ups-rechtlich-wissen-muessen.

However, the EU's highly ethical principle is qualified by the range of exemptions for law enforcement, migration and border control (following judicial approval), research and development of AI, medical and security surveillance and the entire military security sector.[22] The most significant differ-

[22] European Parliament topics: AI Law. First regulation of artificial intelligence. 13 March 2024. Online ibidem; WEKA Business Solutions: The new AI regulation in the EU. 2024. Online ibidem; European Commission: Artificial intelligence Q&A. 12 December 2023. Online ibidem. News: AI regulation final. Online ibidem.

ences to the OECD countries and, in particular, to the US can be seen in the core issue of the risk assessment of AI. The "Guidance for Regulation of Artificial Intelligence Applications" published by the US Office of Management and Budget includes risk assessment and management among the ten basic principles for using AI technology, while the National Institute of Standards and Technology's (NIST) Artificial Intelligence Risk Management Framework explicitly rejects the categorisation of four risk types, arguing that this would only add to the cost of implementing AI regulations and make the path to international AI regulation more difficult. The EU is therefore caught in a dilemma between (too) high ethical standards and the demands of a rapidly growing AI market, which calls for a "Global Partnership on Artificial Intelligence" as defined by the OECD.[23]

2.3 Control and monitoring instruments

The institutionalisation of controls and instruments is analogous to the degree of risk assessment in the use of AI technology. The OECD and UNESCO primarily focus on education, information and training of educators, each coordinated by a Global AI Observatory or Global AI Ethics and Governance Observatory, supported by national institutions of the member states (Austria: Advisory Council on Ethics in AI), while the Council of Europe relies on strengthened legal instruments (guidelines for users on risk mitigation, minimum requirements for impact assessments) to be implemented by national authorities. The UN Resolution also considered its recommendation to be only a prelude to the Summit of the Future held in New York on 22-23 September 2024 as part of the UN General Assembly, where the above-mentioned "digital pact" between the "Global South" and the "Global North" on the fair use of AI technology was concluded. In the end, the only monitoring tools provided were a Scientific Advisory Board and ongoing dialogue between states and private stakeholders. The proposed Digital Human Rights Advisory Service of the UN Office of the High Commissioner was reduced to a monitoring function "upon request".[24] The control measures and instruments of the NATO AI Strategy also do not go beyond general principles such as legality, accountability, transparency, func-

[23] Meltzer/Tielemans: The European Union AI Act. Online ibidem.
[24] Reiland: United Nations. ibidem.

tional reliability and certification following stress tests.[25] The EU AI Act is implemented centrally by a dedicated AI Office and the European Data Protection Supervisor, based on a database for independent high-risk systems, in cooperation with the mandatory national AI authorities (supervisors) of the member states. The extension of the supervisory powers of the AI Office and its sanctioning powers (fines of up to €35 million or 7% of global annual turnover) prevented the softer control proposals (self-assessment by producers, appointment of independent third-party boards) put forward by global data companies during the negotiations on the AI Act. The AI Act is closely linked to other legislation as part of the EU's digital strategy, in particular, the Data Governance Act of April 2022 and the GAIA-X project.[26]

2.4 Conclusion

Apart from the EU's risk-based AI strategy, all current international AI regulations are characterised by consistent optimism surrounding the positive applications of AI in everyday civil society. Protection against the potential dangers of AI and HE technology already in practice, which have long been pointed out by an alarmed international scientific community, is generally found in verbal assurances such as "trustworthy technology", "human-centred control", "transparency of the accountability loop", "ethics of logarithms", etc., whose placebo rhetoric is all too easily overlooked. This is clearly expressed in the proclamations and declarations of the "Global Digital Compact", agreed as part of the UN's "Summit of the Future", which took place on 22-23 September 2024. Although it claims to be a new digital governance order, it does not go beyond general phrases such as "inclusive digital economy" and "inclusive, open, safe and secure digital space".[27]

---

[25] NATO: Summary. Online ibidem.

[26] Meltzer/Tielemans: The European Union AI Act. Online ibidem. European Commission, European Data Protection Supervisor: Artificial Intelligence. accessed 29 June 2024. https://www.edps.europa.eu/data-protection/our-work/subjects/artificial-intelligence_en, News: AI regulation final. Online ibidem. European Commission: European AI Office. 6 June 2024. https://digital-strategy.ec.europa.eu/en/policies/ai-office.

[27] United Nations: Summit of the Future. Accessed 30 September 2024. https://www.un.org/en/summit-of-the-future; United Nation: Dag Hammarskjöld Library, UN documentation: Development. Introduction, 2000-2015. Accessed 10 November 2024. https://research.un.org/en/docs/dev/2000-2015.

3. AI rules in military, defence and security policy

3.1 Special law vs. legal vacuum?

In all of the AI regulations that were compared at the beginning, the area of military security or security policy was either not addressed at all or was deliberately excluded (with the exception of the NATO AI Strategy). This applies both to UN Resolution 2021 (introduced by the US), in which the entire military use of AI was excluded (due to a lack of consensus), and explicitly to the EU AI Act of March 2020, according to which AI systems used exclusively for defence and national security are exempted from the otherwise applicable obligations and prohibitions in the event of unacceptable risks. Against the background of a dense set of AI norms at global and regional European level – which, however, do not cover the entire area of military security and defence policy – there is a need for a specific normative order for this sector, where the massive use of AI technology and the resulting risk potential are currently most evident. The debate at military and civilian level oscillates between the positions of special regulation for the military versus subordination of the military sector to general AI regulations.[28] The UN Convention on Certain Conventional Weapons (CCW), which entered into force in 1983, represented an initial approach to international regulation of AI in the military sector, which subsequently led to years of negotiations by a Group of Governmental Experts (GGE) on the inclusion of "lethal automatic weapons systems" (LAWS). The initial result is the 11 (non-binding) guidelines drawn up in 2019, in which the validity of classical international humanitarian law and the principle of human responsibility and decision-

[28] Ethics and Armed Forces: Controversies in Military Ethics & Security Policy: 2024/1 - AI and Autonomy in Weapons: War and Conflict out of Control? Autonomous Weapons Systems – Current International Discussions. https://www.ethikundmilitaer.de/en/2024/1-ai-and-autonomy-in-weapons-war-and-conflict-out-of-control/autonomous-weapons-systems-current-international-discussions; Ethics and Armed Forces: Controversies in Military Ethics & Security Policy: 2024/1 - AI and Autonomy in Weapons: War and Conflict out of Control? AI for the armed Forces does not need a special morality! A brief argument concerning the regulation of autonomous weapons systems. Accessed 29 June 2024. https://www.ethikundmilitaer.de/en/2024/1-ai-and-autonomy-in-weapons-war-and-conflict-out-of-control/ai-for-the-armed-forces-does-not-need-a-special-morality-a-brief-argument-concerning-the-regulation-of-autonomous-weapons-systems; IFFF: Wiener Konferenz zu autonomen Waffensystemen. 13 May 2024. https://www.wilpf.de/publikationen/.

making authority were established for all automatic (i.e. also AI-based) weapon systems.[29] This consensual mode of negotiation also forms the basis of the mandate of the group of experts appointed by the governments concerned, which was extended in November 2023, but whose results would continue to be considered completely non-binding due to the opposition of certain highly armed states.[30] In contrast to this familiar diplomatic practice, the growing concern of the global community about the potentially lethal effects of the new AI technology led to a joint appeal by the UN Secretary-General and the President of the ICRC in October 2023 to reach a binding agreement on the prohibition of LAWS by 2026. The result of these efforts was UN Resolution 78/241 (introduced by Austria) on autonomous weapons systems, adopted by 164 states at the General Assembly in December 2023. Regional conferences on LAWS have since been held in Latin America, Africa and Europe. The most recent of these was the international conference convened by Austria on 23-24 April 2024 entitled "Humanity at the Crossroads: Autonomous Weapons and the Challenge of Regulation" in Vienna. The aim of the conference, which brought together 130 countries and representatives from business, academia and civil society, was to develop an international regulatory framework to ban or at least restrict autonomous weapons systems (known in the media as "killer robots"). However, the adoption of a common protocol has so far failed to materialise due to the principle of consensus.[31]

---

[29] Wikipedia: Ethics of artificial intelligence. Wikimedia Foundation. 4 December 2024, at 5.29 p.m. (UTC). https://en.wikipedia.org/wiki/Ethics_of_artificial_intelligence.

[30] Wikipedia: Ethics of artificial intelligence, online ibidem; Connolly, Catherine: How are efforts to reach a legally binding agreement on autonomous weapons systems progressing? In: Ethics and Armed Forces: Controversies in Military Ethics & Security Policy. AI and Autonomy in Weapons: War and Conflict out of Control? Accessed 15 July 2024. https://www.ethikundmilitaer.de/en/magazine-datenbank/detail/01-2024/article/autonomous-weapons-systems-current-international-discussions.

[31] Connolly, Catherine: How are efforts to reach a legally binding agreement on autonomous weapons systems progressing? Online ibidem; Bilgeri, Andreas: Autonomous weapons systems - the state of the international debate. In: AI and Autonomy in Weapons: War and Conflict out of Control? ed. Zentrum für ethische Bildung in den Streitkräften (ZEBIS) Ethik und Militär. Kontroversen in Militärethik und Sicherheit. Accessed 15 July 2024. https://www.ethikundmilitaer.de/en/magazine-datenbank/detail/01-2024/article/autonomous-weapons-systems-current-international-discussions; ORF. *News Wien* on 28 April 2024: Konferenz will Regeln für Killerroboter. https://orf.at/stories/3355502/.

3.2 Unity in non-uniformity

In the context of the currently still open debate, the following positions of individual states and alliances of states can be identified at international level, some of which contradictory, some of which cross-cutting:[32]

a) Although a minority of highly armed states participate in the CCW, they reject any prohibition of LAWS through multilateral agreements. In addition to the US, Israel, Australia and India, this includes the Russian Federation, which has been working on the development of LAWS since 2012. The majority of the international community is in favour of a two-tier multilateral regulatory framework that seeks to ban AI and to regulate the controlled use of AI in the military sector. Under international law, this is based on multilateral, non-binding consensus agreements based on the GGE Guiding Principles, which do not in-

[32] Dahlmann, Anja/Hoffberger-Pippan, Elisabeth/Wachs, Lydia: Autonome Waffensysteme und menschliche Kontrolle Konsens über das Konzept. Unklarheit über die Operationalisierung. In SWP-Aktuell 2021/A 31, 14 April 2021. doi:10.18449/2021A31. https://www.swp-berlin.org/10.18449/2021A31/. Bilgeri, S: Autonome Waffensysteme. Online ibidem; Gillen, Erny: Das Militär braucht keine Sondermoral! Ein Zwischenruf zur Regulierung autonomer Waffensysteme. *Ethic and Armed Forces. Controversies in Military Ethics & Security Policy*. 30 October 2024. https://www.ethikundmilitaer.de/en/magazine-datenbank/detail/01-2024/article/ai-for-the-armed-forces-does-not-need-a-special-morality-a-brief-argument-concerning-the-regulation-of-autonomous-weapons-systems; Reitmeier, Gabriele: *Lizenz zum Töten Künstliche Intelligenz in den Waffensystemen und neue Herausforderungen für die Rüstungskontrolle*. ed. Friedrich-Naumann-Stiftung für die Freiheit. October 2020, file:///C:/Users/aborowska/Downloads/Policy%20Paper%20LAWS-final.pdf; Government of the Netherlands. REAIM 2023. Accessed 15 July 2024. https://www.government.nl/ministries/ministry-of-foreign-affairs/activiteiten/reaim; U.S. Department of State, Bureau of Arms Control, Deterrence, and Stability: Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy. 1 November 2023. https://www.state.gov/political-declaration-on-responsible-military-use-of-artificial-intelligence-and-autonomy/; U.S. Department of State (gov): Political Declaration on Responsible Use of Artificial Intelligence and Autonomy. PDF. Accessed 15 July 2024. https://www.state.gov/wp-content/uploads/2023/10/Latest-Version-Political-Declaration-on-Responsible-Military-Use-of-AI-and-Autonomy.pdf, 1-2; EPO, Entwicklungspolitik online Weltpolitik: Autonome Waffensysteme Russland, Indien, Israel und USA blockieren Verbotsverhandlungen. 20 December 2021. https://www.epo.de/index.php?option=com_content&view=article&id=16398:autonome-waffensysteme-russland-indien-israel-und-usa-blockieren-verbotsverhandlungen&catid=50&Itemid=84.

terfere with independent national regulations. These include, for example, the UK and the Netherlands, which together with South Korea organised the REAIM 2023 conference on "Responsible Artificial Intelligence in the Military Domain" in The Hague in February 2023 with the participation of 100 states and representatives from science, research, industry and civil society, and which will be continued in Seoul in September 2024. The "Call for the use of military AI based on national guidelines" agreed by 61 participating states was seen as a setback by civil society NGOs and undermined by the US, with its own "Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy". The programmatic summary of the declaration published by the Department of State is based, on the one hand, on the right of states to develop autonomous weapons systems in order to improve their military capabilities, which ultimately also carry out operations that are no longer controlled by humans. On the other hand, this development should be consistent with ethical standards based on international conventions and traditional international humanitarian law, which should be ensured through training, transparency and oversight of the "senior officer corps". Of the 54 states that signed the declaration by 29 May 2024, 39 already had LAWS.

b) Multilateral agreements at the UN CCW Conference on Disarmament in Geneva, based on the 11 Guiding Principles developed by the Group of Governmental Experts (GGE) in 2019.
    - For most states' parties, these are seen only as non-binding declarations and proposals that do not exclude national regulations, such as France (national AI strategy 2019), Switzerland and Germany, whose government adopted a national AI strategy in 2018, limited to non-military areas, and at the same time decided to ban LAWS, but called on the Bundestag to establish principles for armed drone missions in 2023.[33]
    - A minority (non-aligned states, Austria, Brazil, Japan), on the other hand, calls for the binding effect of the GGE Guiding Principles and the prohibition of LAWS.

[33] Sauer, Frank: Drei Thesen zur nationalen Regulierung von Autonomie in Waffensystemen [Three theses on national regulation for autonomy in weapons systems], in: Lammert, Norbert/Koch, Wolfgang (Eds.): Bundeswehr der Zukunft – Verantwortung und Künstlichen Intelligenz, Konrad-Adenauer-Stiftung, 58-69, 2023.

- Special positions:
  The People's Republic of China occupies a special position. As part of its global power policy, China focuses on unlimited expansion of its armed forces with the development of all AI technologies in order to overcome the existing digital supremacy of the West. In May 2019, China published the "Beijing AI Principles". On the other hand, it is also taking part in the GGE negotiations and, as recently as 2016, was in favour of a ban on LAWS in war operations.[34]
- Japan's current AI policy also appears ambivalent. On the one hand, the G7 member is calling for international regulation of the trustworthy development and use of AI in the civilian sector as part of the Hiroshima AI Process, has established its own AI Safety Institute and is planning a Tokyo Centre for the Global Partnership on AI (GPAI). The Ministry of Defence is also pushing for the widespread use of AI technology in seven core areas of national security as part of a directive issued in July 2024. At the same time, the government issued a directive against the development of LAWS, a position that was also officially submitted to the UN.[35]
- The heterogeneous positions of NATO's 32 member states on the issue of LAWS (ranging from the rejection of any ban to modification of the principle of human-centred control, to government declarations in favour of a ban on autonomous weapons systems) explain why, despite a comprehensive yet very pragmatic catalogue of principles, the Alliance has so far failed to reach a unified position on the regulation of LAWS.

---

[34] Pramudia, Putu Shangrina: China's Strategic Ambiguity on the Issue of Autonomous Weapons Systems. In *Global: Jurnal Politik Internasional*. Vol. 24: No. 1/2000, DOI: 10.7454/global.v24i1.706. https://scholarhub.ui.ac.id/global/vol24/iss1/1/; Schwan, Ben: China stellt Richtlinien für Künstliche Intelligenz auf. Die 'Beijing AI Principles' sollen eine Art Erklärung der Menschenrechte für KI-Anwendungen sein. In: *Heise online, 16 June 2019*. https://www.heise.de/news/China-stellt-Richtlinien-fuer-Kuenstliche-Intelligenz-auf-4442404.html.

[35] Nova, Agenzia: Japan has a policy against the development of lethal fully autonomous weapons: "A human-centric principle should be kept at the center of technological development". 15 July 2024. https://www.agenzianova.com/en/news/il-giappone-adotta-una-politica-contro-lo-sviluppo-di-armi-letali-completamente-autonome/.

- The EU also occupies a special position with its strictly risk-based AI regulation of May 2024, which, however, only applies to the civilian use of AI technologies (and thus also to their commercial use), while the military and security sector is excluded and requires its own regulation. However, there is also resistance within the EU to this "special right" for the military sector.
- The strict separation between civilian and military AI does not exclude its intensive and long-term use as a political and legal agenda of the EU. Any regulation of the production and use of AI technology in the military sector is subject to a permanent conflict of interests within the three decision-making bodies, namely the EU.
    - On the one hand, the need to protect European society from the potentially harmful effects of the use of AI weapon systems.
    - On the other hand, the legitimate security interests of EU member states and the Union itself to protect themselves against attacks and threats from third countries or terrorist groups using all modern weapon systems, including LAWS.
    - In their proposals and concepts, all three decision-makers - the European Commission, the Council (of Foreign Ministers) and the European Parliament - oscillate between an ethically justified strategy of restriction and prohibition and a technology policy that enhances the EU's legitimate defence capabilities.

This applies, in particular, to the position of the European Parliament, which, based on hundreds of submissions and petitions from civil society, had already adopted a resolution in September 2018 calling for a fundamental ban on LAWS. Behind this moral demand, however, there are other security interests of the parliamentarians, who are in favour of the safe development and use of autonomous weapons systems, but in compliance with a clear human-centred legal chain of responsibility or an international governance

control body to be established at UN level.[36] This "soft law" strategy characterises the position of the European Parliament.

This "soft law strategy" also characterises, to a much greater extent, the position of the Foreign Affairs Council (FAC) and the Commission, which avoids all issues and focuses much more on geopolitical and economic aspects, combined with the realisation that the member states are currently not prepared to surrender their sovereignty in these matters in favour of majority decisions. The EU is fully in line with the trend of international AI regulation agreed to date, which is fundamentally characterised by optimistic expectations about the many improvements to everyday life that the new AI technology will bring. This technological "faith" is also widely shared by global civil society, as sociological evidence shows.[37]

3.3 "Human in the loop" vs "Human out of loop"

The decisive criterion for the use or prohibition of LAWS is expressed in two magical keywords: "human in the loop" vs "human out of the loop", i.e. the assured human control of autonomous weapon systems in use - as opposed to operational decisions determined autonomously by the weapon system.[38] However, this apparently ethically and logically clear distinction between applicable and prohibited LAWS loses its selectivity in practice, as the human-centred control principle can also be interpreted pragmatically as "sufficient human control", in which the controlling person exercises overall supervision, but can ultimately be replaced by the autonomous weapon system for operational command.[39]

Transferring the criterion of human-centred control competence to our left-right scheme leads to the following weighting (see Figure 2, next page):

---

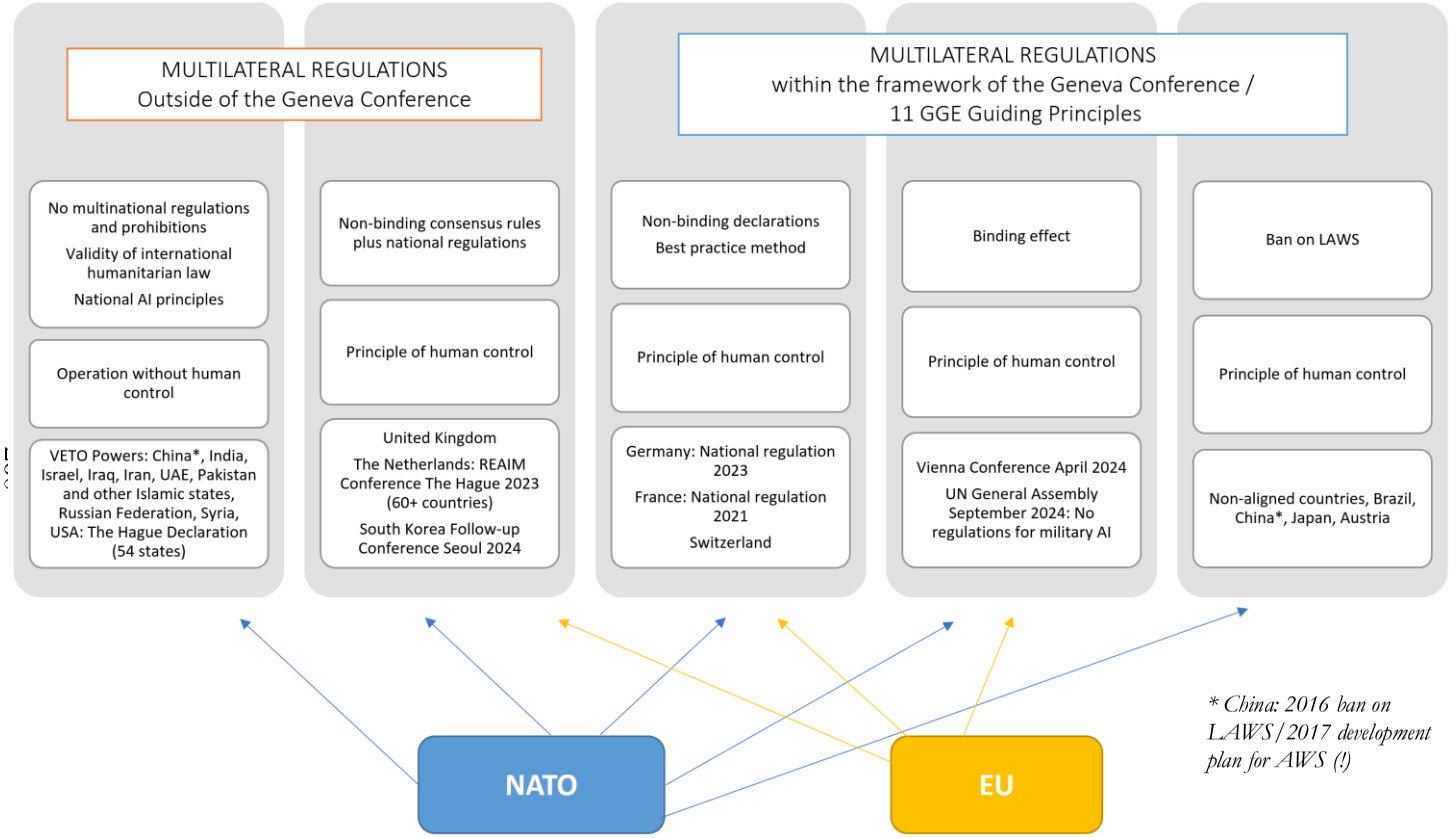
[36] Conn, Ariel: European Parliament Passes Resolution Supporting a Ban on Killer Robots. *Future of Life Institute.* 14 September 2018. https://futureoflife.org/ai/european-parliament-passes-resolution-supporting-a-ban-on-killer-robots/; Santopinto, Federico: *N O T E - The EU, Artificial Military Intelligence and Autonomous Lethal Weapons*. Institut des Relations Internationales et Strategiques. France, 2024. https://coilink.org/20.500.12592/6hdrf4d.

[37] Hofer, Bernhard: Impact. ibidem, Tab. 2 Scenarios.

[38] Wikipedia: Ethics of artificial intelligence. Online ibidem.

[39] Santopinto: The EU. Online ibidem.

Figure 2: Positions of states and alliances of states on LAWS



Authors illustration based on Figure 1. Sources: Dahlmann, Anja/Hoffberger-Pippan, Elisabeth/Wachs, Lydia: Autonomous Weapon Systems and Human Control. Consensus on the concept, uncertainty about operationalisation, SWP-Aktuell 2021/A 31, 14.04.2021, Stiftung Wissenschaft und Politik (SWP) 2021 URL: https://www.swp-berlin.org/10.18449/2021A31/, accessed 15

The minority group against bans and multilateral regulations adopts the interpretation of the U.S. Department of Defense policy that autonomous weapons systems must ultimately be able to select and engage targets without further intervention by a human operator.[40] The International Committee of the Red Cross uses a more drastic definition: "Autonomous weapon systems select targets and apply force without human intervention."[41] In contrast, in 2017, Russia, otherwise a classic veto power, was still in favour of the principle of human-centred control, albeit without an international commitment. For the majority of states that accept multilateral regimes in various forms, the principle of human-centred ultimate control remains, albeit with increasingly different interpretations. The maxim of "meaningful human control" demanded by NGOs and civil society, i.e. the obvious human-control competence as a principle that applies to all phases of decision-making, planning and operational implementation, is replaced by a multitude of semantic modifications from "sufficient human control" to "sufficient human supervision" to "appropriate human judgement". This softening of ethical positions is most evident in the EU institutions: the ban on LAWS adopted by the EP in 2019 was transformed in the resolution of 15 December 2021 into a distinction between pure "killer robots" and normal LAWS, the use of which can only be ordered by authorised persons in specific cases. The Foreign Affairs Council (FAC), which is fundamentally more pragmatic in its approach, is even clearer when it uses the term "sufficient human supervision". With the further differentiation between autonomous air defence systems and LAWS, the promotion of LAWS would, at least in theory, also be possible within the framework of the European Defence Fund.[42]

4. The "Gretchenfrage" (key question) – What do you think about ethics?

The deliberate semantic ambiguities and attempts to interpret key legal and international legal terms have alarmed representatives of ethically and morally oriented groups and civil society bodies to defend principles that were previously beyond dispute. The disputes focus not only on the particularly sensitive area of the application of LAWS, but also on the fundamental prob-

40 Santopinto: The EU. ibidem.
41 Wikipedia: Ethics of artificial intelligence. Online ibidem.
42 Santopinto: The EU. ibidem.

lem of ethically justifiable restrictions on the use of AI and HE technologies, which are ostensibly intended to serve the general progress of humanity or the overall optimisation of human capabilities. In this struggle for "immovable basic principles" and "red lines" that cannot be crossed, we refer to only a few manifestations in the context of the comments in ZEBIS (2024/1) and their main arguments:[43]

- In March 2023, the German Ethics Council spoke out against the ability of machines to act and take responsibility "within the framework of the philosophical theory of action" and insisted on the fundamental human ability to control.
- The interdisciplinary network "Meaningful Human Control – Autonomous Weapon Systems between Regulation and Reflection" calls for "human-centred" interaction between the weapon system and the ultimately responsible human being. This "meaningful human control" would also make it possible to assign responsibility and impose criminal sanctions for misconduct. This is because national and international criminal law, as well as international humanitarian law, are fundamentally human-centred and therefore cannot be applied to the actions of autonomous machines that cause harm.
- From the perspective of Catholic moral theology, "digitalised killing" per se undermines the formal and material dignity of human beings, upon which all modern ethical convictions are ultimately based.
- Beyond religious theology, secular social philosophy also derives its rejection of autonomous AI technology from the principle of inviolable human dignity. Jürgen Habermas and Francis Fukuyama, for example, generally see the physical-psychological "improvement of the species" of humans (illustrated by the figure of the "cyborg") as a loss of their ethical freedom and ultimately a violation of the principle of equality by differentiating between those who have AI and those who do not. The criterion of individual free will for the use of

[43] ZEBIS ed.: Connolly, "Wie geht es mit den Bemühungen [...]"; Gillen: AI for the military ibidem; Beck, Susanne: Humanity in war? The significance of 'Meaningful Human Control' for the regulation of autonomous weapons systems; Koch, Bernhard: Human dignity and 'autonomous' robotics: What is the problem? June 2024. https://www.ethikundmilitaer.de/fileadmin/ethics_and_armed_forces/Ethics-and-Armed_Forces-2024-1.pdf.

(medical, psychological, technical, etc.) HE measures or other AI technologies, which Habermas and others call for as a turning point, extends the debate to the fundamental question of the ethical and legal limits of HE and AI technologies.[44]

Against this phalanx of principle-oriented (deontological) ethicists, to which Bernhard Hofer and Rita Phillips, for example, refer in their contributions, the proponents of "post-humanism" are forming, whose understanding of ethics is measured purely in terms of the success and consequences of the use of weapons, under the maxims of increasing efficiency, global competition and military superiority. According to the maxims of increased efficiency, global competition and military superiority, states (and alliances of states) are entitled not to abandon these potential military advantages without clear evidence of the danger posed by autonomous weapons systems. Civil society, whose acceptance of civil and military AI and HE technology should be demanded, is increasingly being left out of this very "academic" dispute.[45]

This debate is intensified by new innovations and advances in the development of AI capabilities, according to which the problem of the ethical legitimacy of the operation of autonomous AI systems could be solved by the implementation of ethical decision-making capabilities. This would also relativise the chain of human ultimate responsibility in the sense of "machine ethics" (ethics for machines as subjects). If we follow the proponents of this "moral Turing Test", three types of "moral agents" relevant to our debate can be distinguished (James Moor):

---

[44] Hofer, Bernhard: Effects of human enhancement technologies (HET) on society. In chapter SOCIETY in this publication; Tragbar, Lisa/Lagos, Rodrigo: Human Enhancement for military purposes: Ethical considerations. In chapter ETHICS in this publication.

[45] Zajac, Maciek: Burden of Proof in the Autonomous Weapons Debate – Why Ban Advocates Have Not Met It (Yet). Ethics and Armed Forces 01/2024, pp. 34-42 Masuhr, Niklas: KI als militärische Befähigungstechnologie. *CSS Analysen zur Sicherheitspolitik*, nr. 251/October 2019, ed. Merz, Fabien, pp. 1-4. https://css.ethz.ch/content/dam/ethz/special-interest/gess/cis/center-for-securities-studies/pdfs/CSSAnalyse251-DE.pdf; Hofer, Bernhard: Effects of human enhancement technologies (HET) on society. In chapter SOCIETY in this publication. Phillips, Rita: Ethical discourses on autonomous weapon systems. In chapter SOCIETY in this publication.

(1) the implicitly ethical agent, with an ethical dimension traceable to the operator;
(2) the explicitly ethical agent, making ethical decisions based on acknowledged moral information; and
(3) the fully ethical agent, explicitly capable of judgement and reasoning.[46]

However, this would be a clear violation of the principle of responsibility (human in the loop), which has so far been undisputed in international law, and is likely to provoke strong protests and interventions at scientific and social level. On the other hand, it seems appropriate to draw attention to the technological development of military autonomous weapons systems, which is already underway and against which the international community currently imposes no legal restrictions. Joachim Klerx's article "The Future of Human Enhancement in the Military Domain" is a milestone in this respect. In it, the author points to profound changes in warfare between 2025 and 2045, in which a new form of "Artificial Super Intelligence (ASI)" will make the political and strategic decision-making processes of warfare more efficient. This process of "disempowerment" of the individual in the military sector has its counterpart in civilian AI technology: here, the largest global digital platforms (Microsoft, Google, Amazon) are competing to develop and mass-produce a new AI Copilot. This digital agent would be the "intimate" companion of its client, both as a talking advisor and as an independent "legal entity" (signing contracts, managing passwords, etc.). The potential dangers posed by the arbitrary autonomy of AI technology would have to be preemptively averted by internal corporate "ethics committees".[47]

As Rita Phillips explains in detail, certain HE inventions, in particular, augmented reality (AR) and brain-computer interface (BCI), could enable the

---

[46] Wikipedia: Ethik der künstlichen Intelligenz. https://de.wikipedia.org/wiki/Ethik_der_k%C3%BCnstlichen_Intelligenz. Based i.a. on Rath, Matthias/Krotz, Friedrich/ Karmasin, Matthias (eds.): Maschinenethik. Normative Grenzen autonomer Systeme. Springe VS, Wiesbaden 2018; Müller, Vincent C.: Ethics of Artificial Intelligence and Robotics. In: Edward N. Zalta (ed.): The Stanford Encyclopedia of Philosophy (Summer 2021 Edition). https://plato.stanford.edu/archives/sum2021/entries/ethics-ai/.

[47] Klerx, Joachim: The future of human enhancement in the military domain. In chapter TECHNOLOGY in this publication; Jürgens, Johanna: Die Agenten kommen. In: *Die Zeit.* nr. 46. 30 October 2024, 19.

human operator to incorporate more ethical components into the use of AWS. The integration of HE findings into the development of new AWS would strengthen the cognitive and affective capabilities of the human controller in their position "in the loop" (such as the ability to distinguish between civilians and enemies) and thus create greater distance from the killing and destruction of the enemy.[48] However, this new "Icarus" version of a future HE "*Übermensch"* (superhuman) can be countered by the argument of the loss of individual responsibility caused by HE and thus the criminal liability of the actions of "enhanced" persons. The permanent use of biochemical, cybernetic prostheses, etc. HE can lead to the reduction or elimination of the individual's ability to recognise the wrongfulness of the action taken and thus their responsibility and culpability. However, the reduction or loss of the cognitive element of "mens rea" consequently leads to the perpetrator's "enhanced" lack of guilt and impunity. The legal position of the International Criminal Court (ICC), which has been active since 2002, seems to be helpful here. Article 30 of its "Rome Statute", for example, addresses the partial or total lack of responsibility of a perpetrator acting while intoxicated, but not the voluntary use of these means. However, this insight, which is reminiscent of Habermas's postulate of free will, finds its logical limit specifically in the area of military obedience to orders or the ability to act in an emergency. In principle, however, the ultimate responsibility of the state - acting through its executive branch - remains, which, through the use of such means, aims at or tolerates the reduction or elimination of the ethical capabilities of the offending individual.[49]

---

[48] Phillips, Rita: Ethical discourses on autonomous weapon systems. In chapter SOCIETY in this publication.

[49] Harrison Dinniss, Heather A./Kleffner, Jann K.: Soldier 2.0: Military Human Enhancement and International Law. In: International law studies 92/2016, 432-482, here 474-482. https://digital-commons.usnwc.edu/cgi/viewcontent.cgi?article=1695&context=ils; More general on voluntariness and obligation of obedience, see e.g.: Daniel, Giffhorn/Gerndt, Reinhard: "Meaningful Human Control" von autonomen Systemen. In: KI und Autonomie in Waffen: Kriege und Konflikte außer Kontrolle? Ethik und Militär, 01/2024, pp. 68-75. https://www.ethikundmilitaer.de/fileadmin/ethik_und_militaer/Ethik-und-Milit%C3%A4r-2024-1.pdf.

4.1 Military AI ethics perspectives

Two international legal developments are currently decisive for an Austrian HE and AI regulation within the framework of military security policy: The results of the negotiations at the UN Convention on Certain Conventional Weapons (CWW) in Geneva. Despite all verbal commitments to the principles of international law and international humanitarian law, these have a pragmatic-technological orientation (strongly influenced by the USA). This approach of comprehensive international AI regulation is being promoted, in particular, by the Federal Ministry for European and International Affairs, which, following the Vienna Conference on limiting AI in April 2024-2026, anticipates a regulatory proposal from the UN Secretary-General, including the establishment of a regulatory authority. However, the escalation of the current wars means that such a regulation cannot be expected in the foreseeable future.[50]

The EU's AI policy, with its separation of civil and military AI technologies. Although this separation is problematic due to the reality of "dual-use" technologies, it allows the EU to establish a heavily risk-based, binding regulation for the civilian AI sector with a graduated scale of restrictions and prohibitions, while at the same time steering a pragmatic course for the military AI sector that does not abandon the high European ethical standards, but takes into account the security policy sovereignty of the member states and the EU. The fundamental question posed at the beginning of this paper regarding an ethical approach to the use of AI and HE technologies, in AWS in particular, therefore remains unanswered. A European or even international "guideline" is certainly not to be expected. As a result, two technical concepts continue to face each other, with high expectations:

[50] W24: Konferenz in Wien 2024: Regeln für Autonome Waffen. 29 April 2024. https://www.w24.at/News/2024/4/Konferenz-in-Wien-Regeln-fuer-Autonome-Waffen; UN, Regionales Informationszentrum der Vereinen Nationen UNRIC: Summit of the Future. Zukunftsgipfel. https://unric.org/de/un-system/un-zukunftsgipfel-summitofthefuture/; European Commission: The United Nations members adopted a Global Digital Compact shaping a safe and sustainable digital future for all. 23 September 2024, https://digital-strategy.ec.europa.eu/de/node/13020.

- The technocratic perspective, which believes that human responsibility in the operational phase of the mission can be replaced by the further development of fully autonomous weapon systems with the integration of appropriate ethical competences.
- The human-centred counter-position, which believes that the responsibility of the human operator can be preserved even in the final phase of the mission by enhancing human capabilities through the massive use of HE resources.

However, such technologically modified "principled military ethics" based on the European tradition of values can only find their concrete application in conjunction with "situational ethics", which specifically leave room for individual decisions depending on the situation.[51] However, this must not result in the arbitrary generalisation of the predetermined European canon of values, nor must it provide carte blanche for a purely pragmatic and arbitrary interpretation of ethical principles. It should instead be regarded as a guideline for the specific decision-maker on the basis of the proportionality of the infliction of harm and the weighing up of interests. In considering the ethical implications of a machine with "Artificial Super Intelligence", it is essential to recall Clausewitz's enduring insights on "moralische Größen". For Clausewitz, moral values encompass the development of a soldier's character, based on emotional and rational judgement, courage in the face of danger, accountability to external authority and inner conscience. In the context of modern warfare, where machines are increasingly capable of performing tasks traditionally associated with humans, it is crucial to uphold these values to ensure the ethical conduct of operations.[52]

---

[51] Wikipedia: Ethik der künstlichen Intelligenz. https://de.wikipedia.org/wiki/Ethik_der_k%C3%BCnstlichen_Intelligenz.

[52] Sob, Brigitte: 225 Jahre Clausewitz. Teil 3: Friktion und moralische Größen im Krieg. In: Der Soldat. nr. 22/2005, p. 10.

References

Austrian Commission for UNESCO: Ethik in Wissenschaft und Technologie. Gesellschaftliche Folgen antizipieren. https://www.unesco.at/wissenschaft/wissenschafts-und-bioethik/ethik-kuenstliche-intelligenz, accessed 20 April 2024.

Cheung, Sunny; Wettrennen um KI: Chinas Volksbefreiungsarmee und die NATO. In: Friedrich Naumann Stiftung Analyse. November 2023. https://www.freiheit.org/de/taiwan/wettrennen-um-ki-chinas-volksbefreiungsarmee-und-die-nato.

Council of Europe: European Convention on Human Rights. As amended by Protocols Nos. 11 and 14 Supplemented by Protocols Nos. 1, 4, 6, 7, 12, 13 and 16. September 2019. https://edoc.coe.int/en/european-convention-on-human-rights/8363-european-convention-on-human-rights-as-amended-by-protocols-nos-11-and-14-supplemented-by-protocols-nos-1-4-6-7-12-13-and-16.html.

Council of Europe, Secretary General Marija Pejčinović Burić: Artificial Intelligence, Human Rights, Democracy and the Rule of Law Framework Convention. 15 March 2024. https://www.coe.int/en/web/portal/-/artificial-intelligence-human-rights-democracy-and-the-rule-of-law-framework-convention.

Council of the EU: Artificial Intelligence Act: Council calls for promoting safe AI that respects fundamental rights. 6 December 2022. https://www.consilium.europa.eu/en/press/press-releases/2022/12/06/artificial-intelligence-act-council-calls-for-promoting-safe-ai-that-respects-fundamental-rights/.

Dahlmann, Anja/Hoffberger-Pippan, Elisabeth/Wachs, Lydia: Autonomous Weapon Systems and Human Control. Consensus on the concept, uncertainty about operationalisation, SWP-Aktuell 2021/A 31, 14.04.2021, Stiftung Wissenschaft und Politik (SWP) 2021 URL: https://www.swp-berlin.org/10.18449/2021A31/, accessed 15 July 2024.

Europäische Kommission: Gestaltung der digitalen Zukunft Europas. NEWS ARTICLE, 23 September 2024. https://digital-strategy.ec.europa.eu/de/node/13020.

European Commission: Artificial Intelligence – Questions and Answers. 1 August 2024, In: Questions and Answers, 1 August 2024. https://ec.europa.eu/commission/presscorner/detail/en/qanda_21_1683.

European Commission: Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules On Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts. 21 April 2021. https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52021PC0206.

European Parliament: EU AI Act: first regulation on artificial intelligence. The use of artificial intelligence in the EU will be regulated by the AI Act, the world's first comprehensive AI law. Find out how it will protect you. 8 June 2023 (Last updated: 18.06.2024). https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence.

Fokuhl, Josefine: UN-Gremium legt sieben Vorschläge zur Steuerung von KI vor. In: Handelsblatt. 19 September 2024.

Frank, Dorothea: NATO-Zertifizierungsstandard für Künstliche Intelligenz in Behörden. In: Spiegel. 15 February 2023. https://www.behoerden-spiegel.de/2023/02/15/nato-zertifizierungsstandard-fuer-kuenstliche-intelligenz/.

Güell Paule, Laia: UNESCO launches Global AI Ethics and Governance Observatory at the 2024 Global Forum on the Ethics of Artificial Intelligence. 6 February 2024. https://digital-skills-jobs.europa.eu/en/latest/news/unesco-launches-global-ai-ethics-and-governance-observatory-2024-global-forum-ethics.

Hauser, Gunter: Die NATO. Die Grundlage des europäisch-atlantischen Sicherheitsverbundes. In: Hochleitner, Erich P. (ed.): Das Europäische Sicherheitssystem zu Beginn des 21. Jahrhunderts. Edited by Erich P. Hochleitner (Böhlau Verlag, 2000), pp. 267-336.

Hofer, Bernhard: Effects of human enhancement technologies (HET) on society. In chapter SOCIETY in this publication.

KIZUNA: The Hiroshima AI Process: Leading the Global Challenge to Shape Inclusive Governance for Generative AI. 9 February 2024. https://www.japan.go.jp/kizuna/2024/02/hiroshima_ai_process.html.

Kleinwächter, Wolfgang: Entwicklungen im Internet Governance-Umfeld April bis Juli 2024. In: DENIC Internet Governance. 14 August 2024. https://blog.denic.de/entwicklungen-im-internet-governance-umfeld-januar-bis-marz-2024/.

Klerx, Joachim: The future of human enhancement in the military domain. In chapter TECHNOLOGY in this publication

Köver, Chris: KI-Konvention des Europarats. Viel Abkommen um Nichts. In: Netzpolitik Org. 20 March 2024. https://netzpolitik.org/2024/ki-konvention-des-europarats-viel-abkommen-um-nichts/16.

KPMG International: Das bedeutet das EU-Gesetz zur KI. 10 May 2024. https://kpmg.com/de/de/home/themen/2024/03/das-bedeutet-das-eu-gesetz-zur-ki.html.

North Atlantic Treaty Organization: Summary of the NATO Artificial Intelligence Strategy. 22 October 2021. https://www.nato.int/cps/en/natohq/official_texts_187617.html.

OECD: Recommendation of the Council on Artificial Intelligence. OECD/LEGAL/0449.

OECD: Artificial Intelligence in Society. OECD Publishing, Paris 2019. https://doi.org/10.1787/eedfee77-en, accessed 20 April 2024.

Österreichische UNESCO-Kommission: UNESCO Talk: Ethische KI in Europa. 23 June 2023. https://www.unesco.at/wissenschaft/artikel/article/unesco-talk-ethische-ki-in-europa.

Oswald, Fiene: Erste UN-Resolution zu Künstlicher Intelligenz. In: Basecamp. 2 April 2024, pp. 1-6. https://www.basecamp.digital/ki-verstehen-erste-un-resolution-zu-kuenstlicher-intelligenz/.

Phillips, Rita: Ethical discourses on autonomous weapon systems. In chapter SOCIETY in this publication.

Research Institute AG Wien: Vereinte Nationen und Europarat beschließen Regelwerke zu künstlicher Intelligenz. 28 March 2024. https://researchinstitute.at/vereinte-nationen-und-europarat-beschliessen-regelwerke-zu-kuenstlicher-intelligenz/.

Ring-Eifel, Ludwig: Digitale Ethik Vatikan will beim Thema Künstliche Intelligenz mitreden. 23 January 2023. https://www.herder.de/hk/aktuell/vatikan-will-beim-thema-kuenstliche-intelligenz-mitreden.

Tragbar, Lisa/ Lagos, Rodrigo: Human Enhancement for military purposes: Ethical considerations. In chapter ETHICS in this publication.

UNESCO: Regierungen müssen generative KI in Schulen schnell regeln. 7 September 2023. https://www.unesco.de/wissen/ethik/kuenstliche-intelligenz/regierungen-generative-ki-schulen-altersgrenze-13-jahre, accessed 20 April 2024.

United Nations, Office for Digital and Emerging Technologies: Global Digital Compact 1. 15 October 2024. https://www.un.org/global-digital-compact/sites/default/files/2024-09/Global%20Digital%20Compact%20-%20English_0.pdf.

UNRIC-Regionales Informationszentrum der Vereinten Nationen: UN-Zukunftsgipfel. https://unric.org/de/un-system/un-zukunftsgipfel-summitofthefuture/, accessed 30 October 2024.

Vatican News: Pope Francis to participate in G7 session on AI. 26 April 2024. https://www.vaticannews.va/en/pope/news/2024-04/pope-francis-g7-summit-italy-artificial-intelligence.html.

Wikipedia: Human Enhancement, last modified 24 November 2024 at 11.32 PM (UTC). https://en.wikipedia.org/wiki/Human_enhancement.