

# UNESCO

# Recommendation

# on the Ethics

# of Artificial

# Intelligence

Conditions for the Implementation  
in Germany



**unesco**

German Commission  
for UNESCO

# UNESCO Recommendation on the Ethics of Artificial Intelligence. Conditions for the Implementation in Germany

Prof Dr Matthias C. Kettemann

# Table of Contents

---

Abbreviations	6
Preface	8
Summary and Policy Recommendations	11

---

I. The New UNESCO Recommendation on the Ethics of Artificial Intelligence – Global Framework for AI Ethics	14
--	----

---

II. Added Value of the UNESCO Recommendation on the Ethics of AI Compared to Other International Initiatives	20
--	----

---

II.1. OECD Recommendation on Artificial Intelligence	22
--	----

---

II.2. Council of Europe: CAHAI	23
--------------------------------	----

---

II.3. Draft Regulation of the European Union	25
--	----

---

II.4. UNESCO Recommendation on the Ethics of AI: Approach and Added Value	26
---	----

---

II.5. Summary	27
---------------	----

---

III. Conditions for the Implementation of the UNESCO Recommendation on the Ethics of AI in Germany	28
--	----

---

III.1. Introduction	30
---------------------	----

---

III.2. Ethical Impact Assessment (Policy Area 1)	30
--	----

---

III.3. Ethical Governance and Stewardship (Policy Area 2)	34
---	----

---

III.4. Data Policy (Policy Area 3)	36
------------------------------------	----

---

III.5. Development and International Cooperation (Policy Area 4)	37
--	----

---

III.6. Gender (Policy Area 6)	39
-------------------------------	----

---

IV. Conclusion	44
----------------	----

---

Imprint	48
---------	----

# Abbreviations

4IR	Expert National Task Force on the Fourth Industrial Revolution
AI	Artificial intelligence
AIGO	OECD Expert Group on Artificial Intelligence
BMFSFJ	Bundesministerium für Familie, Senioren, Frauen und Jugend/ Federal Ministry for Family Affairs, Senior Citizens, Women and Youth
BMWi	Bundesministerium für Wirtschaft und Energie/ Federal Ministry for Economic Affairs and Energy
CAHAI	Ad hoc Committee on Artificial Intelligence of the Council of Europe
CAHAI-LFG	CAHAI Legal Frameworks Group
CDEP	Committee on Digital Economy Policy
DEK	Datenethikkommission der Bundesregierung/ Data Ethics Commission of the Federal Government
DIN	Deutsches Institut für Normung/ German Institute for Standardization
DKE	Deutsche Kommission Elektrotechnik Elektronik Informationstechnik/ Expertise Centre for Electrotechnical Standardisation in Germany
DUK	Deutsche UNESCO-Kommission/ German Commission for UNESCO
EIA	Ethical Impact Assessment

EKKI	Enquete-Kommission Künstliche Intelligenz des Bundestages/ Enquete Commission on Artificial Intelligence of the German Bundestag
EMRK	Europäische Menschenrechtskonvention/ European Convention on Human Rights
EU	European Union
EU-AI	Proposal for an EU Regulation on Artificial Intelligence
GIZ	Gesellschaft für Internationale Zusammenarbeit/ German Corporation for International Cooperation
AI HLEG	High-level Expert Group on Artificial Intelligence of the European Commission
LDCs	Least Developed Countries
LLDCs	Landlocked Developing Countries
LMICs	Low- and Middle-Income Countries
NKIS	Nationale KI-Strategie der Bundesregierung/ Artificial Intelligence Strategy of the German Federal Government
OECD	Organisation for Economic Cooperation and Development
SDGs	Sustainable Development Goals
SIDS	Small Island Developing States

“Even though much is still open, one thing is clear: we as humanity still have the opportunity to guide the future development of AI and to ethically frame it in a humanistic, human rights perspective. UNESCO has taken up this issue as part of its mandate.”  
— Prof Dr Maria Böhmer, President of the German Commission for UNESCO

# Preface



Prof. Dr. Maria Böhmer,  
President of the German  
Commission for UNESCO

Artificial intelligence (AI) has found its way into many areas of our lives in recent years and fundamentally changed them. It provides personalised search results on the World Wide Web, facial recognition on smartphones and semi-automated driving. The use of self-learning algorithms is a megatrend of our time, but at the same time the long-term consequences of the use of AI can hardly be predicted. Thus, the current discourse on AI and its effects spans a wide range, from utopias of a better and fairer world to dystopias of an uncontrollable superintelligence that dominates humans. Even though much is still open, one thing is clear: we as humankind still have the chance to steer the future development of AI and to ethically frame it within humanistic and human rights-based boundaries.

UNESCO has taken up this issue as part of its mandate. After a long negotiation process involving all 193 Member States and a wide range of stakeholders, UNESCO adopted a landmark Recommendation on the Ethics of Artificial Intelligence in November 2021. The UNESCO Recommendation is the first globally negotiated response to one of the greatest challenges of our time. It convinces with its strong human rights language, a clear ethical foundation and concrete tasks for action for policymakers. Thus, it fulfils all requirements to become a global compass for responsible AI development and use.

The Recommendation defines ethical guidelines as well as concrete tasks for the implementation in policy areas in which AI is already widely discussed today. Beyond that, the inclusive negotiation process has brought areas into focus that have so far played a rather minor role in the AI discourse, but are at the core of UNESCO's mandate. The UNESCO Recommendation contains tasks for action in eleven different policy areas, including education and science, communication, health and the environment. Noteworthy is the Recommendation's focus on previous "blind spots" of AI regulation, such as AI and gender, AI and sustainability, AI and the Global South, and AI and education. In addition, it specifies clear procedural and transparency obligations, such as a multi-stakeholder-based, open and responsive design of relevant processes.

The German Commission for UNESCO, as a link between the UNESCO Secretariat, national politics and civil society, supports the implementation of this important international law document in Germany. This study serves as a substantive foundation for this task. It takes a look at the status quo and clarifies which processes, instruments and laws already exist in Germany that support AI development in accordance with the recommendation.

At the same time, it also analyses where regulation or action is still needed and offers concrete proposals for the implementation of the UNESCO Recommendation in Germany. In addition, it elaborates on the added value of the Recommendation in comparison to other initiatives, such as the OECD Recommendation or the European Commission's proposed Regulation. Just two examples for this added value: The UNESCO Recommendation is the only legal document to refer to cultural aspects that may be affected by the use of AI systems. Furthermore, it strongly emphasises the need to protect marginalised groups in the context of AI development.

We would like to extend our sincerest thanks to the author of the study, Professor Matthias C. Kettemann. We need every effort from politics and administration to ensure that AI development proceeds in an ethical and human rights-based manner. AI dystopias must remain science fiction!





# Summary and Policy Recommendations

## Summary: The Added Value of the UNESCO Recommendation on the Ethics of AI

→ The UNESCO AI Ethics Recommendation is the first international legal text on AI ethics that has a global reach and was negotiated in an open and inclusive multi-stakeholder process. It includes globally accepted ethical standards for AI-based technologies and applications in full respect of international law, in particular human rights. Thus, human rights can now play a central role in the development of AI-related standards around the world. 193 Member States have committed to the Recommendation. The open process over two years, in which non-state actors were also able to contribute their views, represents a significant added value in terms of the Recommendation's legitimacy.

→ Remarkable in terms of content is the Recommendation's holistic approach. At a general level, ethical AI development and use is defined as a central challenge and obligation for the whole of society. On the level of AI governance, the Recommendation formulates concrete tasks for action in eleven policy areas, including education and science, communication, health and the environment.

→ The focus on ethical rules, which are, however, closely linked with and justified by human rights, is deliberately chosen. Clearly tailored ethical rules and principles relate to law in various ways; they can help in the development and implementation of policies and in the interpretation of legal norms by providing orientation points for technological development.

→ Noteworthy is also the Recommendation's focus on normative "blind spots" in other areas, such as AI and gender, AI and sustainability, AI and the Global South (development), AI and education, as well as on clear procedural and transparency obligations: Specifically, it calls for ethical impact assessments and a multi-stakeholder-based, open and responsive design of the corresponding processes.





# Policy Recommendations for AI Actors

## 1 AI must respect, protect and promote human rights

Worldwide, and especially in Germany, every decision on AI – as a contingent technology – must be aligned with ethical values, human rights and international law in the process of planning and deployment.

## 2 Ethical (and legal) guidelines must be followed throughout the whole life cycle of AI systems

AI is no deus ex machina. Therefore, all AI stakeholders must take the entire life cycle of AI systems into account – a process that begins with the formulation of the requirements for an AI system and the composition of the development team that decides on training data and the conditions for machine learning.

## 3 AI data must be as open to the public as possible

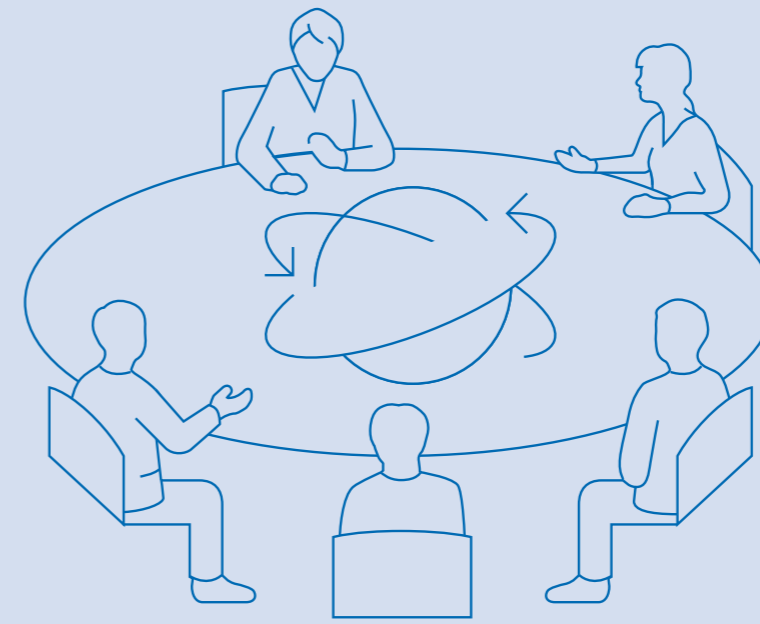
The German government must ensure that the scientific community has access to AI research data in a privacy-conscious manner and that this data is not monopolised by large companies. In addition to the scientific community, this access must also be guaranteed for selected civil society actors who follow different logics and interests than the scientific community and can assume a public watchdog function.

## 4 AI must contribute to a diverse society

Gender, inclusiveness and diversity mainstreaming throughout the entire life cycle of AI applications are crucial to prevent risks of discrimination. In order to ensure that the content of training data is free of stereotypes and bias, documentation obligations must be created for AI stakeholders in Germany. Developers and research teams must also be more diverse. But non-discrimination and equal treatment are systemic issues: politics and administration in Germany should therefore take measures in all policy areas to actively promote equity and equality for women and diverse people.

## 5 AI must be publicly monitored

The German government must develop and establish clearer frameworks for the implementation of ethical impact assessments and monitoring mechanisms. These must include a right to disclosure of transparency protocols for users. Only internal monitoring mechanisms are not sufficient. Corporate AI stakeholders must ensure that the results of AI decision-making processes are comprehensible, explicable and justifiable. Human beings must not feel at the mercy of AI. On the contrary: All AI actors are obliged to increasingly introduce citizens to the role and potential of AI by means of modern knowledge formats.



## 6 AI must become globally accessible

AI stakeholders, especially the German government, should help build AI capacity in Low- and Middle-Income Countries (LMICs) and Least Developed Countries (LDCs) to correct contextual biases of 'Western AI'. It is precisely the lack of access to big data, which is crucial in AI development, that will deepen existing gaps unless deliberate countermeasures are taken. In addition to public development cooperation, the role of the private sector is particularly important. The private sector should act as a partner and should increasingly held accountable by the federal government.

## 7 AI must contribute to the realisation of the SDGs

Digitisation must be sustainable, and sustainability must be digitally designed. AI must always be assessed with a view to promoting the achievement of the Sustainable Development Goals (SDGs). In general, the use of AI must (also) be classified in terms of ecological sustainability.

## 8 AI standards must be negotiated in inclusive processes

The German government and all involved AI companies and standardisation bodies must enable greater civil society engagement in standardisation processes as industry interests are traditionally over-represented here.

As many AI stakeholders as possible should be involved, especially representatives of civil society, data protection authorities and consumer protection organisations as well as representatives of vulnerable groups.



**Introduction: The New UNESCO  
Recommendation on the Ethics of Artificial  
Intelligence – Global Framework for AI Ethics**





# I. Introduction

The UNESCO Recommendation on the Ethics of Artificial Intelligence is a contribution to a global approach to digital human rights. For too long the protection of human rights in digitalization has been sought within national normative logics by too many states. The German Federal Government's 14th report on its human rights policy<sup>1</sup> includes the strengthening of human rights protection "in the face of digital transformation" as a priority in the "Human Rights Action Plan" for the years 2021–2022. This also includes the protection of human rights in the context of the development of Artificial intelligence (AI).

Given the diversity of economic, social, cultural, civil and political rights whose respect, protection and enforcement are affected by AI, there is a need for principles and rules to realise the potential of AI while protecting individual freedoms and guaranteeing social cohesion. Against this background, UNESCO has developed its Recommendation on the Ethics of Artificial Intelligence (hereinafter referred to as: "Recommendation").<sup>2</sup>

The German Government's Coalition Agreement 2021–2025 does not mention human rights in the context of AI, while referring to AI as a "future sphere". At the same time, the use of the potentials of AI is stipulated and the importance of an intensive transatlantic dialogue on data sovereignty, web freedom and AI is emphasised.<sup>3</sup> This shows that a debate on the human rights and ethical aspects of AI development and use is crucial. Building on this, the definition of clear ethical principles and operationalisable tasks for relevant groups of actors is necessary.<sup>4</sup> In this context, the UNESCO Recommendation on the Ethics of Artificial Intelligence is a contribution to a globally conceived, common digital human rights policy.

1 Federal Foreign Office, Fourteenth Human Rights Report of the German Government (Reference period: 1 October 2018 until 30 September 2020), <https://www.auswaertiges-amt.de/blob/2422644/3f981cf30f610babfd16d0eb63ee542c/201202-mrb-14-download-data.pdf>. (All online sources last accessed on 10 December 2021).

2 UNESCO, Recommendation on the Ethics of Artificial Intelligence, SHS/BIO/REC-AIETHICS/2021, 2021 ("The Recommendation").

3 Koalitionsvertrag 2021–2025 zwischen SPD/Bündnis 90/Die Grünen/FPD, [https://www.spd.de/fileadmin/Dokumente/Koalitionsvertrag/Koalitionsvertrag\\_2021-2025.pdf](https://www.spd.de/fileadmin/Dokumente/Koalitionsvertrag/Koalitionsvertrag_2021-2025.pdf), p.19, 21, 154.

4 The heavily used term "ethics" in the current technology debate also requires reflection. What is important is the independent consideration of non-legal normative systems, the added value that lies in their formal non-binding nature and in their significance for the flexibility in dealing with changing technology. However, it is important to point out the undeniable advantages of legal systems of order (which can be enforced by coercive force, if necessary).

The Recommendation, adopted on 23 November 2021, provides an "ethical guiding compass and a global normative bedrock allowing to build strong respect for the rule of law in the digital world."<sup>5</sup> This study focusses on what this can mean for Germany.

First, the study looks at the added value of the UNESCO Recommendation in comparison to other international initiatives. This is necessary against the background of the considerable number of 173 sets of rules on AI ethics contained in AlgorithmWatch's "AI Ethics Guidelines Global Inventory", which was updated in mid-2020.<sup>6</sup>

Principles must become relevant in practice. Therefore, the study analyses – with a view to specific policy areas – the conditions for the implementation of the Recommendation in Germany. An important output of the study is a summary with policy recommendations for political practitioners and multipliers.

**“The UNESCO Recommendation on the Ethics of Artificial Intelligence is a contribution to a global digital human rights policy, which unfortunately is still thought of only nationally in too many states.”**

5 UNESCO, Recommendation on the Ethics of Artificial Intelligence, 2021, <https://en.unesco.org/artificial-intelligence/ethics>.

6 Cf. AlgorithmWatch, AI Ethics Guidelines Global Inventory, <https://inventory.algorithm-watch.org>.

## Characteristics

The Recommendation is the first globally negotiated text under international law in the field of AI ethics. It is not only global, but also holistic in terms of its approach. The Recommendation offers the 193 UNESCO Member States a framework for action in this important future field (by contrast, the Council of Europe's AI rules to be discussed only apply to 47 states; 46 states were involved in the OECD AI standards). UNESCO Member States include not only states with a solid rule of law, such as Germany, in which – via the European level – AI standards are already introduced or currently negotiated. They also encompass states such as China, which produces the lion's share of AI and use it in a challenging way from a human rights perspective. In addition, there are also those Member States that have neither fundamental rights protection against AI nor national AI capacities. Observers found it particularly noteworthy that China, which was critical of certain human rights provisions in the Recommendation during the negotiations, ultimately supported the consensus.<sup>7</sup> Although the largest user of AI (and home to most AI-based start-ups), the USA, is not a UNESCO Member State at the moment, it is known that the USA followed the process of drafting the Recommendation closely within the scope of their observer status.

## Process

The UNESCO Recommendation was developed over two years in an intensive and sometimes controversial intergovernmental negotiation process. This multi-stakeholder approach, which was enriched by several regional stakeholder consultations, already offers significant added value in terms of the Recommendation's legitimacy. By involving various stakeholder groups, the Recommendation thus followed a "best practice" model for international standard-setting.

<sup>7</sup> Lina Rusch, „Künstliche Intelligenz: Soft Law statt Regulierung?“, Tagesspiegel Background Digitalisierung, 16 November 2021, <https://background.tagesspiegel.de/digitalisierung/kuenstliche-intelligenz-soft-law-statt-regulierung>.

## Content

With specific reference to eleven policy areas, including education and science, communication, health and the environment, the Recommendation translates principles for a dignity-sensitive use of AI into policy-making tasks.

The aim of the Recommendation is to align AI with human rights. It gives AI an ethical foundation that not only protects human rights and human dignity, but addresses all three dimensions of human rights protection: respect, protection and promotion.

## Ethics and Rights

The focus on ethical rules, and not primarily on human rights obligations, is deliberately chosen. On the one hand, UNESCO has a special responsibility with regard to ethical considerations in important social agendas and discourses. On the other hand, clearly defined ethical values and principles relate to law in various ways; they can support the development and implementation of policies and the interpretation of legal norms by, as the Recommendation puts it, "providing guidance with a view to the fast pace of technological development". Ethics is not "lesser" than human rights, but an alternative (albeit congruent in terms of protective intent in many areas). Ethical rules are conceived and structured differently. They are not centrally controlled, and enforced in another way than human rights, regularly not by coercive force. States that enter into human rights obligations are bound by them under international law. States that commit themselves to ethical obligations can only – but at least – be induced by international pressure to behave in accordance with their obligations. While for some states the obligation arising from human rights treaties is limited to signing them, it is obvious that ethical commitments have to be fulfilled through 'striving efforts' by states. Especially in a fairly new field such as AI, legal rules have not yet reached consensus – especially internationally. Ethical rules have an important function here.



## Aim

The Recommendation aims to define globally accepted ethical standards for AI technologies in full respect of international law, in particular human rights, which can then be key to the development of AI-related norms around the world.

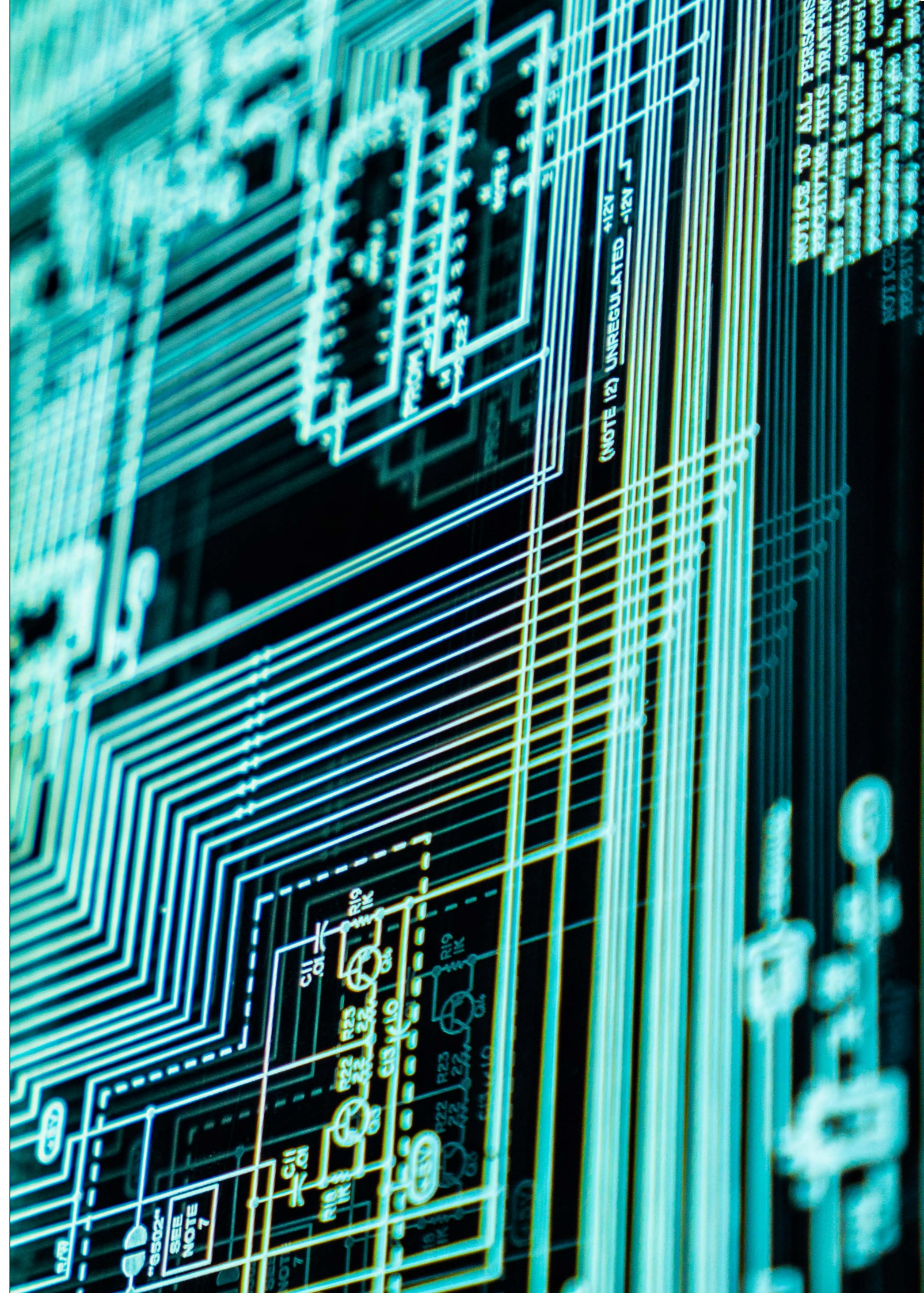
The Recommendation is further characterised by a holistic focus on the different policy fields, an awareness that each field requires different regulatory approaches and the focus on "blind spots" of previous AI regulations. These include environmental protection, the sustainable and resource-efficient use of AI, and the use of AI in education, with full recognition of the right to education for all.

## Normative technique

Depending on the field of action, the recommended measures have a different degree of obligation. Although the Recommendation as a whole is uniformly "soft law", i.e. not formally binding law, it was negotiated in such a detailed manner that its quality differs largely from simple resolutions or declarations. The Recommendation cannot be enforced in courts, but it becomes effective towards states. Among other things, it calls for concrete governance measures, e.g. an Ethical Impact Assessment (EIA) for AI systems or a network of independent AI ethics officers to monitor the EIA. Beyond this, the Recommendation calls for international cooperation and research in the field of AI, and a reconsideration of the usually means-driven selection of measures by states to ensure ethical AI use.



## Added Value of the UNESCO Recommendation on the Ethics of AI Compared to Other International Initiatives



NOTICE TO ALL PERSONS RECEIVING THIS DRAWING: This drawing is only valid for the purpose of reference and neither receipt nor possession thereof confers any right in, or creates any obligation, subject matter



# II. Added value of the UNESCO Recommendation on the Ethics of AI compared to other international initiatives

In order to make a sound assessment of the added value of the UNESCO Recommendation, it is necessary to first look at other international ethics-based AI initiatives. This shows in which policy and governance areas the UNESCO Recommendation makes new, further-reaching proposals, re-structures existing ones or introduces new topics, instruments or approaches to governance. Furthermore, it becomes clear to what extent the Recommendation also provides structural added value besides content-related aspects.

In addition to the UNESCO Recommendation, other AI initiatives of particular relevance to Germany come from the Council of Europe and its Ad hoc Committee on Artificial Intelligence (CAHAI), the European Union and the OECD. They all have issued guidelines and laws on the ethical use of AI or are currently in the process of negotiating them.

## II.1. OECD Recommendation on Artificial Intelligence

The Organisation for Economic Cooperation and Development (OECD) has recognised the need to develop international policy instruments in the context of the widespread development and deployment of AI systems. In the course of this, the OECD Committee on Digital Economy Policy (CDEP), with the involvement of governments, industry, civil society and trade unions in an expert group (OECD Expert Group on Artificial Intelligence, AIGO) developed recommendations which ultimately resulted in the adoption as the “Recommendation of the OECD Council on Artificial Intelligence” on 22 May 2019.<sup>8</sup>

<sup>8</sup> OECD, Recommendation of the Council on Artificial Intelligence, OECD/LEGAL/0449, 22 May 2019, <https://legal-instruments.oecd.org/en/instruments/OECD-LEGAL-0449>.

The OECD Recommendation begins by defining some terms [I.]: The definition of AI systems is so broad, however, that the question arises as to what the difference between AI systems and other software systems should be. The term “machine-based system”<sup>9</sup> does not help much either. It is noteworthy here that the life cycle of an AI system, including its phases, is explicitly defined. In particular the phase of design, data collection and processing and modelling of the system is described as a “context-dependent sequence”.<sup>10</sup> After in its own reflection first intergovernmental standardisation initiative on AI<sup>11</sup>, the OECD Recommendation sets out the following principles for the responsible governance of trustworthy AI in the first section: inclusive growth, sustainable development and quality of life as the goal of the use of AI systems [IV.1.1.], human-centred values and fairness [IV.1.2.], transparency and explainability [IV.1.3.], robustness and security of systems [IV.1.4.], and accountability of “AI actors” [IV.1.5.].<sup>12</sup>

In its second section on national measures and international cooperation, the OECD Recommendation provides for the promotion of a “digital ecosystem”, which should include the digital technologies, infrastructures and mechanisms such as a data trust model necessary for an adequate exchange of “AI knowledge” [V.2.2].<sup>13</sup> This exemplifies the adoption of a rather business-oriented perspective in the document. This impression is reinforced by the call for an “enabling policy environment for AI” [V.2.3].<sup>14</sup> A quick transition should be achieved from the research and development phase to the implementation and operation of trustworthy AI systems [V.2.3.a)].<sup>15</sup>

Furthermore, countries are encouraged to review and, if necessary, adapt their policies and regulatory frameworks to foster AI-related innovation and competition [V.2.3.b)].<sup>16</sup> Thus, a more economy-oriented position in favour of the use of AI in principle cannot be dismissed in the OECD Recommendation.

<sup>9</sup> Ibid., p. 7.

<sup>10</sup> Ibid.

<sup>11</sup> As in the preceding background information, see *ibid.*, p. 3 ff.

<sup>12</sup> Ibid., p. 7 ff.

<sup>13</sup> Ibid., p. 8.

<sup>14</sup> Ibid., p. 9.

<sup>15</sup> Ibid.

<sup>16</sup> Ibid.

Furthermore, the OECD Recommendation assumes that an increasing implementation of AI systems in the economy, corresponding to a change in the world of work and society is necessary, so that humans must acquire the skills needed for the effective use of and interaction with AI systems [V.2.4.].<sup>17</sup> However, such assessments are rightly viewed critically, as they could possibly obscure other causes of social and economic problems as well as related counter-measures.<sup>18</sup> Overall, the OECD Council’s Recommendation is a concise document that draws first lines for the use of AI from a more economy-oriented perspective. On the one hand, the values listed there are of such a universal and obvious nature that most stakeholders would probably not resist agreeing to them. On the other hand, they remain on an abstract level, so that the question of concrete forms, policy and governance measures remains open. This is a significant difference to the UNESCO Recommendation, which does not remain abstract, but instead sets out normatively graded, policy field-specific demands at various points, which put different societal AI actors under obligation.

## II.2. Council of Europe: CAHAI

The Council of Europe is currently working intensively on Artificial Intelligence and its regulation in its bodies. On 11 September 2019, the Committee of Ministers mandated an Ad Hoc Committee on Artificial Intelligence (CAHAI) to conduct a study on the basis of multi-stakeholder consultations and the Council of Europe’s human rights standards. This study was to determine the feasibility of a legal framework and its possible elements.<sup>19</sup> In December 2020, this study was finally published,<sup>20</sup> supplemented by a publication summarising the stakeholder perspectives contributed so far.<sup>21</sup>

<sup>17</sup> Ibid.

<sup>18</sup> Cf. economic-historically: Aaron Benanav, *Automatisierung und die Zukunft der Arbeit* (Übers. von Felix Kurz, Suhrkamp 2021), p. 17 ff.

<sup>19</sup> Lisa Peets/Marty Hansen/Sam Jungyun Choi/Marianna Drake, *AI Update: The Council of Europe Publishes Feasibility Study on Developing a Legal Instrument for Ethical AI*, Global Policy Watch, 14 January 2021, <https://www.globalpolicywatch.com/2021/01/ai-update-the-council-of-europe-publishes-feasibility-study-on-developing-a-legal-instrument-for-ethical-ai>.

<sup>20</sup> CAHAI, Feasibility Study, CAHAI(2020)23, 17 December 2020, <https://rm.coe.int/cahai-2020-23-final-eng-feasibility-study-/1680a0c6da>.

<sup>21</sup> See Council of Europe, *Towards Regulation of AI Systems*, DGI(2020)16, December 2020, <https://rm.coe.int/prems-107320-gbr-2018-compli-cahai-couv-texte-a4-bat-web/1680a0c17a>.

The process initiated in 2019/2020 continues to progress. Currently, the Legal Frameworks Group of CAHAI (CAHAI-LFG) has, after multi-stakeholder consultations and plenary meetings of CAHAI<sup>22</sup> developed a first feasibility study of the possible elements of legally binding instruments in the context of AI, human rights, democracy and the rule of law, which is now to be reviewed by CAHAI and subsequently submitted to the Committee of Ministers.<sup>23</sup>

AI technologies, here characterised as services and products, are seen as having the potential to promote welfare as well as individual and societal prosperity. On the other hand, potential negative effects and dangers, especially with regard to the rights protected by the European Convention on Human Rights (ECHR), are also recognised [marginal no. 2].<sup>24</sup> Examples of human rights risks identified in the study include AI systems that undermine the right to equality and non-discrimination by perpetuating prejudices and stereotypes (e.g. in employment), and AI-driven surveillance and tracking applications that threaten individuals’ rights to freedom of assembly and expression.<sup>25</sup> Furthermore, CAHAI explicitly recognises the lack of a uniform, universally recognised definition of the term Artificial Intelligence and classifies it as a collective term that should be approached in a technology-neutral manner [marginal no. 8-10].<sup>26</sup>

The associated CAHAI feasibility study considers a wide range of instruments for regulating AI, including international legal instruments such as the ECHR and the EU Charter of Fundamental Rights [marginal no. 68 ff]<sup>27</sup>, ethical guidelines for AI, including those developed by private companies and public sector organisations [marginal no. 74 ff]<sup>28</sup>, and national instruments and policies for AI [marginal no. 78 ff].<sup>29</sup>

<sup>22</sup> See in this regard CAHAI, *Abridged Meeting Report and List of Decisions*, CAHAI(2021)10, 7 July 2021, <https://rm.coe.int/cahai-2021-10-5th-plenary-abridged-report-2776-1003-8532-v-2/1680a31d48>.

<sup>23</sup> Council of Europe Communication of 03 November 2021, <https://www.coe.int/en/web/artificial-intelligence/-/the-cahai-lfg-holds-its-5th-and-final-meeting>.

<sup>24</sup> CAHAI, Feasibility Study, CAHAI(2020)23, 17 December 2020, <https://rm.coe.int/cahai-2020-23-final-eng-feasibility-study-/1680a0c6da>, p. 2 ff.

<sup>25</sup> Lisa Peets/Marty Hansen/Sam Jungyun Choi/Marianna Drake, *AI Update: The Council of Europe Publishes Feasibility Study on Developing a Legal Instrument for Ethical AI*, Global Policy Watch, 14 January 2021, <https://www.globalpolicywatch.com/2021/01/ai-update-the-council-of-europe-publishes-feasibility-study-on-developing-a-legal-instrument-for-ethical-ai>.

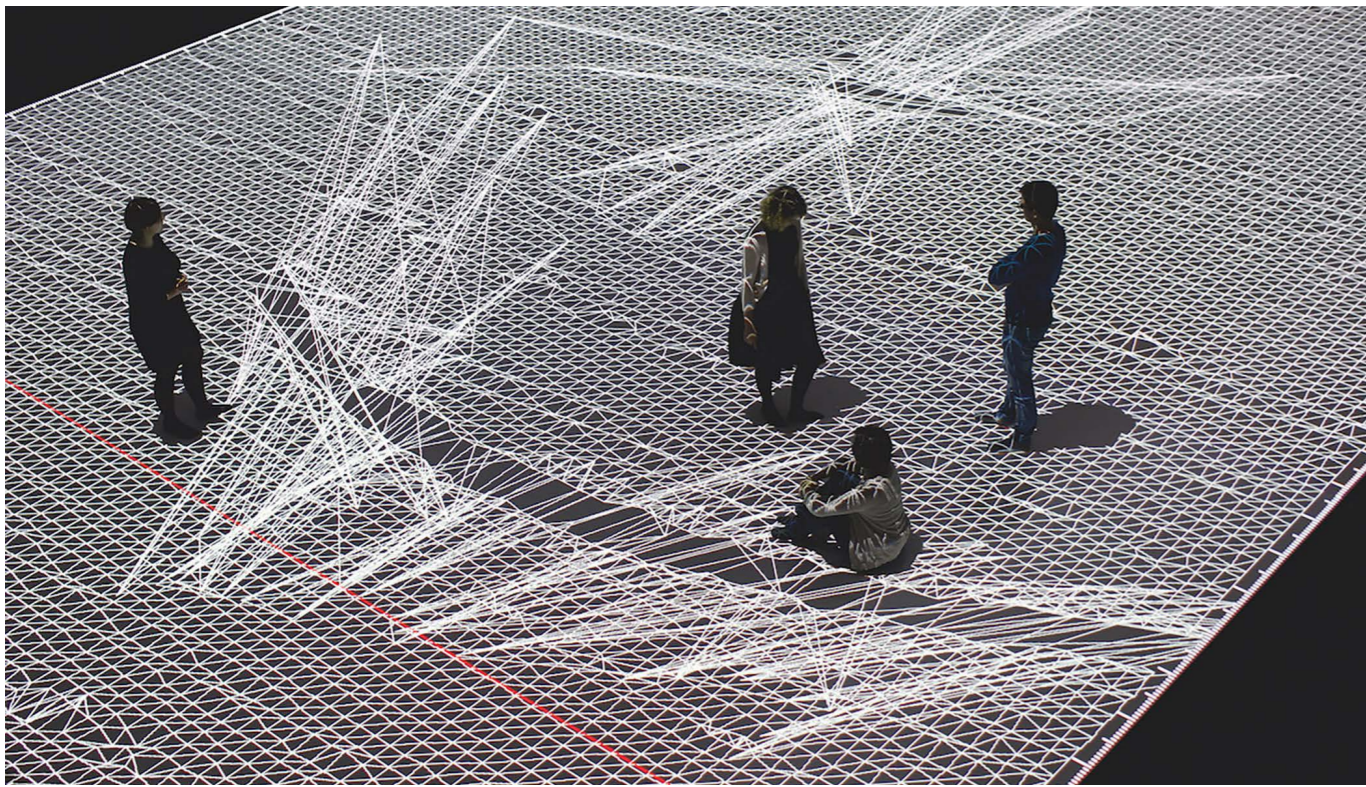
<sup>26</sup> CAHAI, Feasibility Study, CAHAI(2020)23, 17 December 2020, <https://rm.coe.int/cahai-2020-23-final-eng-feasibility-study-/1680a0c6da>, p. 4.

<sup>27</sup> Ibid., p. 18 ff.

<sup>28</sup> Ibid., p. 20 ff.

<sup>29</sup> Ibid., p. 21 ff.





German Commission for UNESCO

CAHAI's feasibility study weighs up the advantages and disadvantages of these measures and finds that there is currently no binding international legal instrument specifically tailored to the challenges posed by AI systems [marginal no. 82 ff].<sup>30</sup> In addition, the many already existing instruments that partly overlap and reinforce each other make consistent and comprehensive interpretation difficult. As a result, no uniform level of protection can be guaranteed while the CAHAI study also concludes that contemporary human rights protection as a whole has gaps in this respect [marginal no. 83 ff].<sup>31</sup>

Among other factors, special emphasis is put on the need for sufficient human control and oversight<sup>32</sup> of the technical robustness of AI applications, as well as transparency<sup>33</sup> and explainability.<sup>34</sup>

In response to the challenges, the CAHAI study sets out the core elements of a legal framework in the form of principles, rights and obligations [marginal no. 95 ff].<sup>35</sup> In addition to these key elements, two overarching measures aim to ensure that human rights are taken into account in the development and use of AI: Firstly, by ensuring that those affected can invoke their human rights in the AI context. Secondly, by describing the requirements that developers and users of AI systems must meet in order to protect human rights. CAHAI identifies seven essential principles to ensure the integration and protection of human rights in and by AI, including human dignity [marginal no. 98 ff.], the requirement of non-discrimination, gender justice, fairness and diversity [marginal no. 105 ff].<sup>36</sup>

The Recommendation of the CAHAI study to the Council of Europe represents a legally binding instrument which, based on the above principles, could serve as a basis for national legislation which should follow a risk-based and -differentiated approach [marginal no. 95 ff].<sup>37</sup> In addition, further soft and hard law instruments are recommended, which should be precise and tailored to the AI sector-specific needs and peculiarities.<sup>38</sup>

30 Ibid.

31 Ibid.

32 See also Riikka Koulu, "Proceduralizing Control and Discretion: Human Oversight in Artificial Intelligence Policy", *Maastricht Journal of European and Comparative Law* 27.6 (2020).

33 See, among others Thomas Wischmeyer, "Artificial Intelligence and Transparency: Opening the Black Box", in: Thomas Wischmeyer/Timo Rademacher (ed.), *Regulating Artificial Intelligence* (Springer International Publishing 2020); Margot E Kaminski, "Understanding Transparency in Algorithmic Accountability", in: Woodrow Barfield (ed.), *The Cambridge Handbook of the Law of Algorithms* (Cambridge University Press 2020).

34 Lisa Peets/Marty Hansen/Sam Jungyun Choi/Marianna Drake, AI Update: The Council of Europe Publishes Feasibility Study on Developing a Legal Instrument for Ethical AI, *Global Policy Watch*, 14 January 2021, <https://www.globalpolicywatch.com/2021/01/ai-update-the-council-of-europe-publishes-feasibility-study-on-developing-a-legal-instrument-for-ethical-ai>.

35 CAHAI, Feasibility Study, CAHAI(2020)23, 17 December 2020, <https://rm.coe.int/cahai-2020-23-final-eng-feasibility-study-1680a0c6da>, p.18 ff.

36 Ibid., p.27 ff.

37 Ibid.

38 Lisa Peets/Marty Hansen/Sam Jungyun Choi/Marianna Drake, AI Update: The Council of Europe Publishes Feasibility Study on Developing a Legal Instrument for Ethical AI, *Global Policy Watch*, 14 January 2021, <https://www.globalpolicywatch.com/2021/01/ai-update-the-council-of-europe-publishes-feasibility-study-on-developing-a-legal-instrument-for-ethical-ai>.

While the CAHAI study is an expert study, the UNESCO Recommendation is the result of an intergovernmental negotiation process between UNESCO Member States. The CAHAI study therefore did not have to seek consensus. It is also noticeable that the study deliberately refers to the necessary development of legal regulations. With the elaboration of the AI-related human rights acquis the study is highly valuable but, however, different from the policy field-specific Recommendation of UNESCO with its broad ethically based approach.

### II.3. Draft Regulation of the European Union

The European Commission's legislative proposal for an AI regulation can be described as ambitious and will be binding to EU Member States after its adoption. On 21 April 2021, the European Commission presented the Draft Regulation<sup>39</sup> as a first, essential step in a potentially lengthy and complex legislative process. This was preceded by the establishment of a High-Level Expert Group on AI, which produced a comprehensive report<sup>40</sup> and the publication of the Commission's "White Paper on Artificial Intelligence – A European Approach to Excellence and Trust".<sup>41</sup> However, those two processes were not seamlessly interconnected.

The structure and approach of the European Commission's Draft Regulation is guided by established measures in the field of product safety.<sup>42</sup> It follows – as recommended by the CAHAI study – what it calls a "well-defined", risk-based and regulatory approach.<sup>43</sup> Art. 3 introduces a comprehensive catalogue of definitions, including a definition of AI as "software that is developed with one or more of the techniques and approaches listed in Annex I and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with".<sup>44</sup> By combining functional and descriptive elements in the

39 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, COM(2021) 206 final, 21 April 2021, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206>.

40 High-Level Expert Group on AI, Policy and Investment Recommendations for Trustworthy AI, 26 June 2019, p.48, <https://digital-strategy.ec.europa.eu/en/library/policy-and-investment-recommendations-trustworthy-artificial-intelligence>.

41 European Commission, White Paper On Artificial Intelligence – A European approach to excellence and trust, COM(2020) 65 final, 19 February 2020, [https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020\\_en.pdf](https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf).

42 Michael Veale/Frederik Zuiderveen Borgesius, "Demystifying the Draft EU Artificial Intelligence Act – Analysing the Good, the Bad, and the Unclear Elements of the Proposed Approach", *Computer Law Review International* 22.4 (2021), p.97 f.

43 European Commission, Artificial Intelligence Act, p.3.

44 Ibid., p.39 ff.

definition, the regulatory scope of the Commission's draft is broad. While not all types of software are covered, some characteristics of most systems are.<sup>45</sup>

The Commission differentiates between four different risk levels of AI, namely unacceptable, high, limited and minimal risks.<sup>46</sup> Thus, certain practices of Artificial Intelligence are explicitly prohibited in Art. 5, albeit with certain reservations of permission.<sup>47</sup> In Chapter 2, Art. 8 ff. of the draft, certain requirements are defined for high-risk AI systems to be permitted, including transparency and information obligations (Art. 13) and human supervision (Art. 14).<sup>48</sup> Chapter 3, Art. 16 ff. stipulates the obligations of providers and users as well as other parties involved in the use of high-risk AI systems.<sup>49</sup> In this context, there has already been criticism that the text of the draft is ambiguous with regard to the definition of high-risk AI systems and therefore needs to be improved.<sup>50</sup> Art. 56 ff. envisage the establishment of a "European Artificial Intelligence Board" to advise and assist the Commission.<sup>51</sup> The application and implementation of the Regulation is assigned in Art. 59 to national supervisory authorities to be designated or established by the Member States.<sup>52</sup> Art. 60 provides for the establishment of an EU database for stand-alone high-risk AI systems by the Commission,<sup>53</sup> which can be assessed as a very reasonable effort.<sup>54</sup>

It would go beyond the scope of this paper to describe and critically assess the entire regulatory structure of the Commission's draft. As points of criticism, however, it is worth mentioning here – as an example – that the Commission's draft does not contain sufficient complaint and legal protection mechanisms for individuals and communities affected by AI systems, and that the enforcement system exhibits a certain degree of incoherence.<sup>55</sup> However, the draft as a whole represents the world's first attempt to horizontally regulate AI systems and thus fits into the image of the EU as a big player of digital policy that is unrivalled in its critical approach towards the normative challenges of digitisation – which could admittedly entail risks.

45 See critically Michael Veale/Frederik Zuiderveen Borgesius, "Demystifying the Draft EU Artificial Intelligence Act – Analysing the Good, the Bad, and the Unclear Elements of the Proposed Approach", *Computer Law Review International* 22.4 (2021), p.109.

46 Ibid., p.98.

47 European Commission, Artificial Intelligence Act, p.43 ff.

48 Ibid., p.46 ff.

49 Ibid., p.52 ff.

50 Luciano Floridi, "The European Legislation on AI: A Brief Analysis of Its Philosophical Approach", *Philosophy & Technology* 34 (2021), p.215, 219.

51 European Commission, Artificial Intelligence Act, p.72 ff.

52 Ibid, p.73 ff.

53 Ibid, p.74.

54 Michael Veale/Frederik Zuiderveen Borgesius, "Demystifying the Draft EU Artificial Intelligence Act – Analysing the Good, the Bad, and the Unclear Elements of the Proposed Approach", *Computer Law Review International* 22.4 (2021), p.111.

55 Ibid.



## II.4. UNESCO Recommendation on the Ethics of AI: Approach and Added Value

For obvious reasons, the UNESCO Recommendation on the Ethics of AI<sup>56</sup> does not pursue a strong regulatory claim comparable to the European Commission's draft. As a supplementary and fundamental normative instrument [see marginal no. 5 ff.], the Recommendation attempts to set its own accents in a comprehensive document by supporting Member States in establishing legal and regulatory frameworks and by serving all "AI actors" as a basis for Ethical Impact Assessments along the life cycle of AI systems [marginal no. 4]. This life cycle perspective is to be welcomed, since upstream and downstream effects and potential challenges of AI systems are considered [marginal no. 2]. The Recommendation refers to systems that have the capacity to process data and information in a way that resembles intelligent behaviour – without defining the concept of intelligence – and typically include aspects of reasoning, learning, perception, prediction, planning or control [marginal no. 2].

In the following, the UNESCO Recommendation first outlines the ethical values affected by AI [III.1, marginal no. 13 ff] and specifies their connection to AI in principles [III.2, marginal no. 25 ff]. These, in turn, are operationalised in eleven areas of policy action [marginal no. 48 ff]. Besides rather evident areas, such as (Ethical) Impact Assessments [marginal no. 50 ff.] and the establishment of strong governance mechanisms [marginal no. 54 ff.], areas that have so far been rather neglected are explicitly and extensively referred to, e.g. the protection of the environment and ecosystems [marginal no. 84 ff.], gender [marginal no. 87 ff.], culture [marginal no. 94 ff.] and education and science [marginal no. 101 ff.]. It should be emphasised that the Recommendation opposes the idea that AI technologies are per se some kind of miracle cure to ensure the well-being of humans, the environment or ecosystems as a whole [marginal no. 25]. Further, the Recommendation demands human and public oversight of AI systems and a final human responsibility and accountability for their results [marginal no. 35 f.].

<sup>56</sup> UNESCO, Recommendation on the Ethics of Artificial Intelligence, SHS/BIO/REC-AIETHICS/2021, 2021 ("The Recommendation").

Moreover, various aspects of the UNESCO Recommendation deserve particular attention, such as the recognition that the use of AI systems also influences the human mind, the ways in which people think, interact and make decisions, and thus the image of humanity. In its global claim, the UNESCO Recommendation also emphasises the special characteristics and challenges of Low – and Middle-Income Countries (LMICs) – among which so-called Least Developed Countries (LDCs), Landlocked Developing Countries (LLDCs) and Small Island Developing States (SIDS) are highlighted – and thus addresses especially the Global South.<sup>57</sup> The cultures, values and knowledge of the indigenous population, whose recognition, protection and promotion are needed for the development of sustainable digital economies, are particularly emphasised and outlined in the areas of political action [cf. e.g. marginal no. 85].<sup>58</sup>

The UNESCO Recommendation thus points to unique contexts in states and societies that are highly relevant for the development and implementation of AI systems. So far, this aspect has been mentioned at best on the fringes of ethical discussions, especially with regard to the international initiatives described above.

This approach of those initiatives is partly understandable as particularly the CAHAI Recommendation and the European Commission's Draft Regulation address the European context. However, the international effects of those initiatives must be considered: Especially the regulatory approach of the European Union is being observed very closely internationally, with considerable effects on respective policy and governance initiatives in other states; the EU's General Data Protection Regulation is a prominent example of this.

The UNESCO Recommendation is particularly legitimised by the multi-stakeholder approach used in its preparation: input from representatives of different societal groups was deliberately sought. Furthermore, in this respect, the UNESCO Recommendation generates added value as its development process can serve as an example for other international standard-setting processes.<sup>59</sup>

<sup>57</sup> Ibid., p.2.

<sup>58</sup> Ibid., p.15.

<sup>59</sup> Ibid., p 20 ff.

## II.5. Summary

When it comes to AI, many ethical guidelines overlap in the propagation of some conventional principles, but diverge in the next step when it comes to practical implementation.<sup>60</sup> A profound understanding of AI systems and the consequences of their implementation requires a sound knowledge of society and its processes. This knowledge presupposed, the need for context-sensitive, long-term perspectives on the social embedding of AI technologies could be formulated. Many of the impacts caused by these are determined less by the technology itself than by the conditions in which it is used and the broader economic dynamics with which it is associated. This requires governments to take on a broader and more guiding role.<sup>61</sup> UNESCO's value-based Recommendation will provide substantial support in this regard.

A clear added value of the UNESCO Recommendation can also be seen in the fact that it refers to the cultural aspects that can be affected by the use of AI systems, and to the special economic dynamics that can be different in LMICs and could potentially be much more vulnerable. In this context, it is particularly significant that explicit reference is also made to the substantial contributions that LMICs provide to fairer AI use and better AI regulation.<sup>62</sup> The fact that the UNESCO Recommendation also places emphasis on the importance of marginalised groups and the need to protect them should also be seen as an important impetus to ensure an adequate level of protection for such groups in other legislative contexts, especially on the EU level.

<sup>60</sup> See already CAHAI, Feasibility Study, CAHAI(2020)23, 17 December 2020, <https://rm.coe.int/cahai-2020-23-final-eng-feasibility-study-/1680a0c6da>, marginal no. 74 ff., p.20.

<sup>61</sup> According to the Dutch Scientific Council (WRR) in the English summary of its report "Opgave AI. De nieuwe systeemtechnologie", see WRR, Mission AI, The New System Technology, 11 November 2021, [https://www.wrr.nl/binaries/wrr/documenten/rapporten/2021/11/11/opgave-ai-de-nieuwe-systeemtechnologie/Summary+WRRreport\\_Mission+AI\\_The+New+System+Technology\\_R105.pdf](https://www.wrr.nl/binaries/wrr/documenten/rapporten/2021/11/11/opgave-ai-de-nieuwe-systeemtechnologie/Summary+WRRreport_Mission+AI_The+New+System+Technology_R105.pdf), p.39.

<sup>62</sup> See for example Jason Edward Lewis (ed.), Indigenous Protocol and Artificial Intelligence Position Paper, Honolulu, Hawaii: The Initiative for Indigenous Futures and the Canadian Institute for Advanced Research (CIFAR), 15 April 2020, <https://doi.org/10.11573/spectrum.library.concordia.ca.00986506>.



Conditions for the Implementation of  
the UNESCO Recommendation on the Ethics  
of AI in Germany





# III. Conditions for the Implementation of the UNESCO Recommendation on the Ethics of AI in Germany

## III.1. Introduction

In Germany, AI is receiving a lot of attention in both public and political discourse, especially with regard to regulation and promotion of AI. This chapter describes the conditions for the implementation of the AI Recommendation in Germany in the following policy areas:

- Ethical Impact Assessment (Policy Area 1)
- Ethical Governance and Stewardship (Policy Area 2)
- Data Policy (Policy Area 3)
- Development and International Cooperation (Policy Area 4)
- Gender (Policy Area 6)

In describing the status quo in these policy areas, the following questions will be addressed:

To what extent do (political and regulatory) processes, instruments and governance structures already exist in Germany and at the EU level that address AI in these policy areas in accordance with the Recommendation (laws, platforms, funding programmes, bodies, testing mechanisms, standards etc.)? What contents of the Recommendation are still in great need of regulation or action in Germany? If there is a need for regulation or action, has it already been addressed in other contexts (e.g. by the Enquete Commission on AI of the German Bundestag)?

In each policy area, the description of the status quo is followed by a brief, exemplary comparison with rules or initiatives in the Global South (especially Africa).

## III.2. Ethical Impact Assessment (Policy Area 1)

In the following section, the contents of the UNESCO Recommendation regarding ethical consequences of AI systems and its current implementation will be examined. As reference documents, the proposal for the AI Act of the EU<sup>63</sup> (EU-AI) and the National AI Strategy of the German Federal Government<sup>64</sup> (NKIS) were taken into consideration in particular. In addition, the Ethical Guidelines for Trustworthy AI of the EU

63 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, 2021/0106 (COD), 2021. Cited as EU-AI.

64 Artificial Intelligence Strategy of the German Federal Government, 02 December 2020, [https://www.ki-strategie-deutschland.de/files/downloads/Fortschreibung\\_KI-Strategie\\_engl.pdf](https://www.ki-strategie-deutschland.de/files/downloads/Fortschreibung_KI-Strategie_engl.pdf). Cited as NKIS.

Commission's High-Level Expert Group on Artificial Intelligence<sup>65</sup> (HEG-KI), the report of the Enquete Commission on Artificial Intelligence of the German Bundestag<sup>66</sup> (EKKI) and the expert opinion of the Federal Government's Data Ethics Commission<sup>67</sup> (DEK) were examined.

### Number 50 of the UNESCO Recommendation: Ethical Impact Assessments

No. 50 of the Recommendation calls on Member States to introduce an "Ethical Impact Assessment", which aims to identify and assess the benefits, concerns and risks of AI systems, and to propose appropriate measures for risk prevention, mitigation and monitoring, among other safety mechanisms. The EU-AI, on the other hand, focusses in §§ 30 ff. on the classification of an AI system as high-risk AI, depending on which hazards are to be expected. Ethical aspects such as the impact on fundamental rights are also included here. At the national level, the DEK recommends that the state should be involved in the development of ethical standards for AI, which should then serve as a reference point for impact assessment tools.<sup>68</sup> The DEK also recommends involving "civil society actors, data protection authorities, consumer protection experts or spokespersons for organisations representing the parties affected"<sup>69</sup> in the standardisation process. The EKKI formulates similar claims.

The Federal Government is therefore obliged to create the basis for an Impact Assessment of AI.<sup>70</sup> No. 50 of the UNESCO Recommendation has thus been largely taken up and only needs to be implemented in practice across the board.

65 High-Level Expert Group on Artificial Intelligence, Ethics Guidelines for Trustworthy Artificial Intelligence, 8 April 2019, <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>. Cited as HEG-KI.

66 Report of the Enquete Commission on Artificial Intelligence – Social Responsibility and Economic, Social and Ecological Potentials (orig. Bericht der Enquete-Kommission Künstliche Intelligenz – Gesellschaftliche Verantwortung und wirtschaftliche, soziale und ökologische Potenziale), BT-DrS 19/23700, 28 October 2020, <https://dserver.bundestag.de/btd/19/237/1923700.pdf>. Cited as EKKI.

67 Opinion of the Data Ethics Commission, October 2019 [https://www.bmi.bund.de/SharedDocs/downloads/EN/themen/it-digital-policy/datenethikkommission-abschlussgutachten-lang.pdf?\\_\\_blob=publicationFile&v=4](https://www.bmi.bund.de/SharedDocs/downloads/EN/themen/it-digital-policy/datenethikkommission-abschlussgutachten-lang.pdf?__blob=publicationFile&v=4). Cited as DEK.

68 DEK, p.76.

69 Ibid.

70 EKKI, p.121.

### Number 51 of the UNESCO Recommendation: Due Diligence and Oversight Mechanisms

No. 51 of the Recommendation calls on Member States and private sector companies to develop due diligence and oversight mechanisms to identify the impact of AI systems on the compliance with human rights, the rule of law and inclusive societies, to prevent or mitigate any negative consequences, and to be accountable for how they deal with them.

The EU-AI provides for a tiered risk management system for AI in Art. 9. This includes measures to safeguard fundamental rights. According to Art. 9(6), tests must be "suitable for fulfilling the intended purpose of the AI system". Similar provisions are found in the NKIS. It calls for an appropriate control structure and verifiability of AI applications.<sup>71</sup> The HEG-KI emphasises that "human values are central to the way in which AI systems are developed, deployed, used and monitored" which should also serve the respect of fundamental rights.<sup>72</sup>

In contrast to the UNESCO Recommendation, the national AI Strategy and the planned EU-AI Regulation do not provide for mandatory measures to analyse the socio-economic impact of AI on poverty. The HEG-KI and the EKKI recognise the potential of AI to bridge the gap between rich and poor<sup>73</sup> and to promote economic equality.<sup>74</sup> However, concrete measures to explore and use this potential are not mentioned.

Entitlement to the issuance of transparency protocols: The HEG-KI sees transparency as one of the criteria that a trustworthy AI must fulfil.<sup>75</sup> Adequate documentation must allow ex-post controls by authorities.<sup>76</sup> According to Art. 23 of the EU-AI, responsible authorities can demand the release of the protocols according to Art. 20 if this is necessary for the compliance of the AI system with the obligations of the Regulation. The EU-AI does not provide for a right to release of transparency protocols for organisations or (private) individuals.

In contrast, the UNESCO Recommendation emphasises the need to ensure access to information, including information of public interest, held by private entities. This is not only to assess the impact of the socio-economic status on AI use, but also to ensure that the gap between people living in wealth and poverty, as well as the digital divide between and within countries, is not widened by the massive use of AI technologies now and in the future. Member States, private companies and civil society should study the sociological and psychological impact of AI-based findings on the decision-

71 NKIS, p.23.

72 HEG-KI, p.37.

73 Ibid., p.9.

74 EKKI, p.12.

75 HEG-KI, p.8.

76 Ibid., p.24.

making autonomy of humans. AI systems identified as posing potential risks to human rights should be widely tested in the context of Ethical Impact Assessments. Where appropriate, assessments should also include tests under real-world conditions.

The National AI Strategy also recognises transparency as an important building block for AI.<sup>77</sup> The EKKI report goes one step further: It states that transparency could also be enforced by means of regulatory measures in order to contain risks.<sup>78</sup> People should also be able to defend themselves against discrimination by AI. This requires a right to transparency.<sup>79</sup> The DEK takes a similar view, considering a right to transparency of the data used, depending on the risk category of the AI system, necessary.<sup>80</sup>

The sociological and psychological effects of AI-based recommendation systems need to be researched: The EU-AI prohibits in Art. 5 para. 1 lit. a any AI system "that deploys subliminal techniques beyond a person's consciousness in order to materially distort a person's behaviour in a manner that causes or is likely to cause that person or another person physical or psychological harm". However, just what causes psychological harm is not answered in the proposed Regulation. Nor does it contain any obligation to research this field. The HEG-KI does not deal with the consequences of recommendation algorithms either.

The EKKI describes interdisciplinary research into the effects of recommendation algorithms as an urgent task.<sup>81</sup> The DEK agrees with this assumption, however it does not limit its assessment to the effects of recommendation algorithms. The impact of any form of AI should be further researched.<sup>82</sup>

The EU-AI provides in Art. 5 for comprehensive testing obligations over the entire life cycle of AI. This also includes tests before an AI system is introduced into the market. According to Art. 9, para. 6 of the EU-AI, the test procedures "shall be suitable to achieve the intended purpose of the AI system". This means that testing in a real environment is not mandatory. The national AI Strategy recognises the importance of tests, especially tests in regulatory sandboxes ("Reallabore").<sup>83</sup> The German government supports digital test fields.<sup>84</sup>

77 NKIS, p.23.

78 EKKI, p.464.

79 Ibid., p.64.

80 DEK, p.186.

81 EKKI, p.461.

82 DEK, p.174.

83 NKIS, p.20.

84 Ibid.



**Number 52 of the UNESCO Recommendation:  
Monitoring all phases of the AI life cycle**

No. 52 of the Recommendation provides that, as part of the Ethical Impact Assessment, Member States and companies should take appropriate measures to monitor all stages of the life cycle of an AI system, including the functioning of the algorithms used for decision-making, the data, and the AI actors involved in the process, especially in public services and where direct interaction with end-users is required. The human rights obligations of Member States should be part of the ethical aspects of the evaluation of AI systems.

Art. 9, para. 2 of the EU-AI stipulates that the risk management of high-risk AI is to be understood as a continuous process throughout the entire life cycle of an AI system. The HEG-KI also recommends such continuous monitoring, e.g. by public authorities.<sup>85</sup>

According to Art. 7 para. 1 lit. b of the EU-AI, the violation of fundamental rights is one aspect of the risk assessment of AI. The national AI Strategy declares the protection of fundamental rights to be the goal of all norms and standards for AI. The EKKI sees the Basic Law and the Charter of Fundamental Rights of the European Union as the yardstick for the political shaping of AI.<sup>86</sup> Therefore, the promotion and regulation of AI should be oriented towards these rights and values.<sup>87</sup> The DEK supports the approach that Europe's AI Strategy should be strictly

oriented towards "European values and fundamental rights, in particular those enshrined in the European Union's Charter of Fundamental Rights and the Council of Europe's Convention for the Protection of Human Rights and Fundamental Freedoms".<sup>88</sup> This also includes a ban on AI with "incursions into personal privacy and integrity that are incompatible with fundamental rights".<sup>89</sup>

**Number 53 of the UNESCO Recommendation:  
Legal Framework for Impact Assessments**

No. 53 of the Recommendation encourages governments to adopt a legal framework that establishes a process (especially for public authorities) to conduct Ethical Impact Assessments for AI systems in order to predict consequences, mitigate risks, avoid harmful consequences, facilitate citizen participation and address societal challenges. The assessment should also establish appropriate oversight mechanisms, including verifiability, traceability and explainability, that allow for the assessment of algorithms, data and design processes and include external review of AI systems. Ethical Impact Assessments should be transparent and, where appropriate, made available to the public. Such assessments should also be multidisciplinary, multi-stakeholder, multicultural, pluralistic and inclusive. Public authorities should be required to monitor the AI systems they implement and/or use by putting in place appropriate mechanisms and tools.

85 HEG-KI, p.24.  
86 EKKI, p.83.  
87 Ibid., p.137.

88 DEK, p.226.  
89 Ibid., p.97.

Article 64, para. 1 of the EU-AI regulates the powers of the market surveillance authorities. They are given unrestricted access to training, validation and test data sets. In the case of specific violations of fundamental rights, Art. 64 para. 3 provides for special measures. The assessment of an AI as a risk system is also the responsibility of the authorities and is carried out according to the guidelines set out in Title III, Chapter 1.<sup>90</sup>

The National AI Strategy stipulates the promotion and further development of practical approaches to risk management.<sup>91</sup> Furthermore, supervision and enforcement of requirements should primarily be assigned to the supervisory authorities as they would already have competence in this field.<sup>92</sup> Ethical Impact Assessments should be transparent and, where appropriate, publicly available.

Neither the Recommendation of the HEG-KI nor the EU-AI provide for the publication of the Ethical Impact Assessment. According to the NCIS, one of the requirements for AI is "a level of transparency (...) that adequately reflects the risks".<sup>93</sup> However, it does not provide for specific obligations to publish Ethical Impact Assessments.

**Conclusion**

There is still a lot to catch up on the part of German legislators and other AI actors when it comes to implementing the UNESCO Recommendation's provisions on Ethical Impact Assessments. In this regard it would be a promising approach to consider the existing ethical recommendations and documents of the OECD, the Council of Europe and the EU jointly and to distil the respective added values. The UNESCO Recommendation offers a global and holistic framework that proposes important measures across policy areas. The forthcoming EU-AI Regulation, on the other hand, has the advantage of being binding for a small group of states. In a next step, the respective added values should therefore be identified in a comparative review, whereby the UNESCO Recommendation, as the thematically broadest, offers the best basis.

While not included in the draft EU-AI Regulation, other institutions besides UNESCO recommend the right to the publication of transparency protocols for users. Research into the sociological and psychological consequences of recommendation algorithms is also not required at European level, but at national level. Furthermore, no obligation to test AI in a real-world environment

90 EU-AI, Art. 30 ff.  
91 NKIS, p.30.  
92 Ibid.  
93 Ibid., p.23.

at European level exists. Germany could lead the way in implementing the UNESCO Recommendation and demand transparency from AI providers. The digital divide within Germany between the formally educated and the formally less educated, between rich and poor, must also be overcome. The development of an Ethical Impact Assessment tool for AI could be headed by the Office of Technology Assessment at the German Bundestag.

A clear legal framework for an ethical assessment of AI systems is still needed, partly because the EU-AI Regulation is yet to be adopted. The future framework must be able to pursue different objectives in a sensitive way, namely both to introduce effective monitoring mechanisms and to (simultaneously) monitor them again. In particular, public authorities should be required to monitor their own AI in use: The DEK recommends an obligation for public authorities to monitor the AI they use<sup>94</sup> and to appoint a contact person for the use of AI.<sup>95</sup> The EKKI also recommends that public authorities monitor the AI they use and that public authorities' competences and resources are strengthened for this purpose.<sup>96</sup> These points are also put forward in the Recommendation, but extended by more concrete requirements.

A brief, comparative look at India: The world's largest democracy is one of the few developing and emerging countries with its own AI Strategy.<sup>97</sup> Among other things, this strategy stipulates that AI must be subjected to an ethical impact assessment. This is intended in particular to prevent or limit discrimination through AI. The introduction of an ethics council for AI is also recommended.<sup>98</sup> A less concrete initiative in the sense of an ethical change of perspective is the focus on Buddhist values in the ethical impact assessment of AI.<sup>99</sup>

94 DEK, p.198.  
95 Ibid., p.29.  
96 EKKI, p.194.  
97 Eugenio V. Garcia, "The International Governance of AI: Where is the Global South?", The Good AI, 28 January 2021, <https://thegoodai.co/2021/01/28/the-international-governance-of-ai-where-is-the-global-south>.  
98 National Strategy for Artificial Intelligence, June 2018, [https://www.niti.gov.in/writereaddata/files/document\\_publication/NationalStrategy-for-AI-Discussion-Paper.pdf](https://www.niti.gov.in/writereaddata/files/document_publication/NationalStrategy-for-AI-Discussion-Paper.pdf).  
99 Soraj Hongladarom, "What Buddhism can do for AI ethics", MIT Technology Review, 6 January 2021, <https://www.technologyreview.com/2021/01/06/1015779/what-buddhism-can-do-ai-ethics>.



### III.3. Ethical Governance and Stewardship (Policy Area 2)

The self-assessments by public authorities with regard to their use of AI, as suggested in No. 57 and 58 of the UNESCO Recommendation, are only planned in Germany. Thus, no established practice by and within public authorities exists.<sup>100</sup> With regard to securing a regulatory framework aimed at accountable and responsible AI, as outlined in No. 68-70 of the Recommendation, there is currently no general regulatory framework in Germany. However, there are partial or sector-specific regulations, e.g. the labelling obligation for so-called social bots in the State Media Treaty (§ 18 para. 3 of the Medienstaatsvertrag, MStV)<sup>101</sup> or the media law-specific transparency requirements for algorithmic recommendation systems (§§ 93 ff. MStV).<sup>102</sup>

The national AI Strategy put forward in No. 56 of the UNESCO Recommendation was originally adopted by Germany's Grand Coalition in November 2018<sup>103</sup> and most recently updated with an advanced report on the strategy in December 2020<sup>104</sup> in order to "strengthen the development and expansion of the quality infrastructure on the basis of an appropriate regulatory framework (...) for secure and trustworthy AI" and to "support civil society networking and involvement in the development and use of public good-oriented AI".

With regard to the forms of soft governance of AI as stipulated in No. 56 of the Recommendation, standardisation initiatives have already been launched along the so-called AI standardisation roadmap.<sup>105</sup> However, these initiatives are currently shaped by a primarily technical perspective and incorporate at best incompletely the multi-level review systems envisaged in the Recommendation, including ethical standards and procedural consultation and review

obligations with respect to those affected. The Recommendation can therefore serve as an impetus for further practical initiatives in the future, such as experiments with institutionalised users or stakeholder participation in the context of AI governance.

In a comparable way, European standard-setting initiatives are also influenced by an overemphasis on technical issues and respective stakeholders that take part in them. Beyond, the draft European AI Regulation<sup>106</sup> provides no answer to the question of compatibility of AI use with fundamental and human rights and the participation models required to assure this compatibility as put forward in No. 61-63 of the UNESCO Recommendation.<sup>107</sup> It seems likely that, due to the incentive regime of the AI Act, standardisation authorities, in particular the European Committee for Electrotechnical Standardisation, will be granted a strong practical role in "on-the-ground" rulemaking within the scope of the AI Act.<sup>108</sup> Standardisation processes tend to be structurally inaccessible to those affected by AI products and services and to NGOs. This entails the risk of overemphasis on industry interests.

Nevertheless, the draft for the EU-AI Regulation reflects some of the basic conditions with regard to ensuring AI systems that comply with human rights and participation requirements as stipulated in the UNESCO Recommendation. In particular, the draft provides for a procedure to report serious incidents and malfunctions (Art. 62) and establishes risk-based transparency obligations (Art. 52). With regard to the practical conditions for implementation, it is still unclear whether the Committee on AI to be established will be sufficiently equipped to fulfil its role effectively and to protect the public and individual interests of those affected.<sup>109</sup> Besides the EU-AI draft Regulation, the promotion of AI-based innovation on EU level is tackled by other initiatives, in particular by the Data Governance

Act,<sup>110</sup> the Open Data Directive<sup>111</sup> and further measures under the EU Data Strategy<sup>112</sup>. All those initiatives collectively pursue the goal to create trustworthy mechanisms and services for the re-use, sharing and aggregation of data.

The integration of international organisations in the development of AI strategies and standards, as stipulated in No. 64 of the UNESCO Recommendation, should be further expanded in Germany. For example, the German government's AI Strategy so far only calls for a study on the feasibility and possible elements of a legal framework under international law for the development, design and application of AI, without mentioning clear participation models for its development and realisation.<sup>113</sup> Besides, there have been no solid proposals on the inclusive composition of AI development teams and data sets (No. 67). In the context of the EU-AI Regulation, only the voluntary codes of conduct according to Art. 69 contain respective provisions.

#### Conclusion

The currently prevailing approach of stewardship of AI processes in Germany is clearly expandable to ensure more diversity and participation of stakeholders in the conception and development of AI systems. Especially with regard to the latter (and also in view of the UNESCO Recommendation), a political and legal definition of which models of participation should be tested seems appropriate. In this respect, experiments with (new) institutional structures must be undertaken.

This observation coincides in part with the Enquete Commission's proposal to expand transfer and cooperation mechanisms between science, business, politics and civil society in the field of AI research. The Enquete Commission suggested testing various participation models within the frameworks of citizen sciences, real labs and other inclusive approaches.<sup>114</sup> Beyond, the commission suggests to integrate adapted forms of (analogue) employee participation into AI governance systems in the area of "work".<sup>115</sup>

100 Cf. however the statement of the Federal Government on the "White Paper on AI" of the European Commission, with the proposal for the establishment of a central body for certification or conformity assessment of the use of AI systems by security authorities: Federal Government, Statement on the White Paper on Artificial Intelligence (orig. Stellungnahme zum Weißbuch zur Künstlichen Intelligenz), 2020, [https://www.ki-strategie-deutschland.de/files/downloads/Stellungnahme\\_BRReg\\_Weissbuch\\_KI.pdf](https://www.ki-strategie-deutschland.de/files/downloads/Stellungnahme_BRReg_Weissbuch_KI.pdf), p.18.

101 Lena Isabell Löber/Alexander Roßnagel, „Kennzeichnung von Social Bots: Transparenzpflichten zum Schutz intergrer Kommunikation“, Multimedia und Recht (2019), p.493.

102 In particular § 93 MStV, cf. in this respect also Art. 29 of the draft Digital Services Act, available at <https://eur-lex.europa.eu/legal-content/en/TXT/?uri=COM%3A2020%3A825%3AFIN>; comprehensive on the classification of the relationship of national platform regulation to the AI Act: Jan Christopher Kalbhenn, „Designvorgaben für Chatbots, Deepfakes und Emotionserkennungssysteme: Der Vorschlag der Europäischen Kommission zu einer KI-VO als Erweiterung der medienrechtlichen Plattformregulierung“, Zeitschrift für Urheber- und Medienrecht (2021), p.663–674, here: p.663.

103 Bundesregierung, KI-Strategie, 2018, [www.ki-strategie-deutschland.de](http://www.ki-strategie-deutschland.de).

104 NKIS.

105 DIN, Standardization Roadmap AI, <https://www.din.de/en/innovation-and-research/artificial-intelligence/ai-roadmap>.

106 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, Brüssel COM(2021) 206 final 2021/0106 (COD), 2021, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206>.

107 For a comprehensive joint statement by numerous NGOs on this see European Digital Rights, <https://edri.org/wp-content/uploads/2021/01/EDRI-open-letter-AI-redlines.pdf>, 12 January 2021.

108 Michael Veale/Frederik Zuiderveen Borgesius, "Demystifying the Draft EU Artificial Intelligence Act — Analysing the Good, the Bad, and the Unclear Elements of the Proposed Approach", Computer Law Review International 22.4 (2021), p.14.

109 European Data Protection Board and the European Data Protection Supervisor, Joint Opinion on the Artificial Intelligence Act, 18 June 2021, [https://edpb.europa.eu/system/files/2021-06/edpb-edps\\_joint\\_opinion\\_ai\\_regulation\\_en.pdf](https://edpb.europa.eu/system/files/2021-06/edpb-edps_joint_opinion_ai_regulation_en.pdf).

110 Proposal for a Regulation of The European Parliament and of the Council on European data governance (Data Governance Act), COM(2020) 767, 25 November 2020, <https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%3A52020PC0767>.

111 Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information (PE/28/2019/REV/1, ABL. L 172 of 26.6.2019, p.56).

112 European Commission, "A European strategy for data", COM(2020) 66 final, 19 February 2020, [https://ec.europa.eu/info/sites/default/files/communication-european-strategy-data-19feb2020\\_en.pdf](https://ec.europa.eu/info/sites/default/files/communication-european-strategy-data-19feb2020_en.pdf).

113 NKIS, p.30.

114 EKKI, p.102.

115 Ibid., p.317 f., p.320 f.

Such participation-oriented institutional and procedural models of stakeholder participation should also be tested beyond the area of "work". This seems particularly advisable in view of the current overemphasis on standardisation procedures in regulatory approaches (see above). Of course, they have an epistemic added value but they cannot replace further stakeholder participation in AI governance.<sup>116</sup>

A comparison of the regulatory measures described above with existing regulations or initiatives in Africa reveals that only a few formalised national AI strategies have been introduced there.<sup>117</sup> However, some countries have set up task forces and committees that have positioned themselves on the outlined topics, such as Uganda's "Expert National Task Force on the Fourth Industrial Revolution" (4IR),<sup>118</sup> in whose final report<sup>119</sup> the realisation of the benefits of AI as a tool for progress is valued higher than potential dangers. Also, the question on appropriate (institutional) structures for participation and networking in the field of AI is taken up in the Global South: The final report of the Ugandan 4IR refers, for example, to initiatives for "Centres of Excellence" of the National Association of Software Companies in India, which bring potential stakeholders for the development or use of certain AI applications together. Uganda wants to transfer this model to the agricultural sector in order to achieve improvements there through a public-private partnership.<sup>120</sup>

116 For an analysis of the EU AI proposal in this respect, emphatically Michael Veale/Frederik Zuiderveen Borgesius, "Demystifying the Draft EU Artificial Intelligence Act — Analysing the Good, the Bad, and the Unclear Elements of the Proposed Approach", Computer Law Review International 22.4 (2021), p.15.

117 As of March 2021, Mauritius was said to be the only country in Africa with a "fully formalised" national AI Strategy, cf. Arthur Gwagwa/Patti Kachidza/Kathleen Siminyu/Matthew Smith, Responsible Artificial Intelligence in Sub-Saharan Africa: Landscape and general state of play, [https://irci.org/wp-content/uploads/2021/03/AI4D\\_Report\\_Responsible\\_AI\\_in\\_SSA.pdf](https://irci.org/wp-content/uploads/2021/03/AI4D_Report_Responsible_AI_in_SSA.pdf), 2021, p.16 f., with reference to A. Sey, „Sub-Saharan Africa: Regional Analysis“, in: Oxford Insights, & International Development Research Centre (IDRC). Government Artificial Intelligence readiness index, 2020, <https://www.oxfordinsights.com/government-ai-readiness-index-2020>.

118 Xinhuanet, "Uganda Prepares to Harness Opportunities of 4th Industrial Revolution", 4 September 2019, [http://www.xinhuanet.com/english/2019-04/09/c\\_137963317.htm](http://www.xinhuanet.com/english/2019-04/09/c_137963317.htm).

119 Cf. 4IR Task Force, "Uganda's National 4IR Strategy: A continental 4IR hub that enables a smart and connected Ugandan Society", 2021, <https://ict.go.ug/wp-content/uploads/2020/10/Executive-Summary-Ugandas-National-4IR-Strategy.pdf>.

120 Ibid., p.14, with references to the facilities of the Indian National Association of Software and Service Companies (NASSCOM), cf. <https://coe-dsai.nasscom.in> and <https://www.coe-iot.com>.

### III.4. Data Policy (Policy Area 3)

The German Federal Government adopted its strategy on Artificial Intelligence in November 2018.<sup>121</sup> The strategy, its concrete measures and current developments are available on the website “www.ki-strategie-deutschland.de”. The strategy explicitly lists “data use, data security, law and ethics” as one of the fields of action. Its overall intention is to provide an “overview of goals, foundations and fields of action for a national strategy on Artificial Intelligence”. Under point 3.8 (“Making data available and facilitating its use”), some topics are addressed in accordance with the UNESCO Recommendation. In particular, the quality of data should be increased and access to data should be facilitated.<sup>122</sup> In this context, data protection regulations and the right to informational self-determination should always be respected.<sup>123</sup> Beyond, the creation of high-quality training data sets should be promoted.<sup>124</sup> In addition, an exchange with data protection supervisory authorities on the development of AI systems should occur.<sup>125</sup>

In November 2019, an interim report on the strategy’s implementation was published.<sup>126</sup> In the area of “data policy” the report initiated the data infrastructure “GAIA-X”, regular exchanges on AI data protection issues with data protection supervisory authorities and an open data platform.<sup>127</sup> In addition, the Data Ethics Commission was established in the meantime and delivered its report on the strategy in October 2019.<sup>128</sup> This report emphasises respect for the rights of individuals involved, use and sharing of data, data quality, information security and transparency.<sup>129</sup> It also describes in detail, among other things, data rights and corresponding data obligations<sup>130</sup>, requirements for the use of personal data<sup>131</sup> and access to personal data.<sup>132</sup> In December 2020, the update of the AI Strategy was published, which formulates the further steps for the implementation of the strategy.<sup>133</sup> Also the final report of the Enquete Commission on Artificial Intelligence from October 2020 deals in detail with AI and data.<sup>134</sup>

121 Bundesministerium für Wirtschaft und Energie (BMWi), Strategie Künstliche Intelligenz der Bundesregierung, 16 November 2018, <https://www.bmwi.de/Redaktion/DE/publikationen/technologie/strategie-kuenstliche-intelligenz-der-bundesregierung.html> (translations by the author).

122 Ibid., p. 33.

123 Ibid.

124 Ibid., p. 35.

125 Ibid., p. 37.

126 Bundesregierung, Zwischenbericht ein Jahr KI-Strategie, available under “Downloads” on the website [www.ki-strategie-deutschland.de](http://www.ki-strategie-deutschland.de).

127 Ibid., p. 7.

128 DEK.

129 Ibid., p. 81 ff.

130 Ibid., p. 85 ff.

131 Ibid., p. 95 ff.

132 Ibid., p. 124 ff.

133 NKIS.

134 EKKI.

“Lernende Systeme” (“Learning Systems”) is yet another important AI platform in Germany.<sup>135</sup> Within the focus area “Data Ecosystems” it published a brochure which defines access to data and data quality as important basis requirements for responsible AI use.<sup>136</sup> In addition, the German Institute for Standardisation (DIN) and the Expertise Centre for Electrotechnical Standardisation in Germany (DKE), together with the Federal Ministry for Economic Affairs and Energy (BMWi), have published a roadmap on norms and standards for Artificial Intelligence.<sup>137</sup> The roadmap focusses on data protection<sup>138</sup> and secure data<sup>139</sup> under the main topic of “IT security in AI systems”. The KI-Bundesverband is another body that also regularly deals with “data policy” in the field of application of AI in position papers and press releases.<sup>140</sup>

With regard to funding programmes, the establishment of AI service centres should be highlighted.<sup>141</sup> Those centres should provide access to data sets and at the same time guarantee data sovereignty. A complete list of funding programmes can be found at “www.ki-strategie-deutschland.de”.<sup>142</sup>

At the European level, a European Commission agenda was published in April 2018, which emphasises, among other things, access to data and an appropriate ethical and legal framework.<sup>143</sup> In spring 2020, the “White Paper on Artificial Intelligence – A European Approach to Excellence and Trust”<sup>144</sup> was presented. It also addresses the risks to personal data protection and security.<sup>145</sup> In addition, the draft EU-AI Regulation addresses data and data governance in Art. 10 and describes the quality management

135 See Homepage Plattform Lernende Systeme, <https://www.plattform-lernende-systeme.de/home-en.html>.

136 Lernende Systeme, Creating Value from Data, July 2020, [https://www.plattform-lernende-systeme.de/files/Downloads/Publikationen\\_EN/PLS\\_Booklet\\_Creating\\_Value\\_from\\_Data\\_2.pdf](https://www.plattform-lernende-systeme.de/files/Downloads/Publikationen_EN/PLS_Booklet_Creating_Value_from_Data_2.pdf).

137 DIN and DKE, German Standardization Roadmap Artificial Intelligence, November 2020, <https://www.din.de/resource/blob/772610/e96c34dd-6b12900ea75b460538805349/normungsmroadmap-en-data.pdf>.

138 Ibid., p. 95.

139 Ibid., p. 98 ff.

140 Position papers and statements of the German AI Association, see <https://ki-verband.de/en/>.

141 Bundesministerium für Bildung und Forschung, Richtlinie zur Förderung zum Aufbau von KI-Servicezentren, 1 October 2021, <https://www.bmbf.de/bmbf/shareddocs/bekanntmachungen/de/2021/10/2021-10-01-Bekanntmachung-KI.html>.

142 The German federal states have also published AI strategies with a “data policy” focus. They are not presented here. However, they are available at [www.ki-strategie-deutschland.de](http://www.ki-strategie-deutschland.de).

143 European Commission, Communication from the Commission – Artificial Intelligence for Europe, COM(2018) 237 final, 25 April 2018, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2018%3A237%3AFIN>.

144 European Commission, White Paper On Artificial Intelligence – A European approach to excellence and trust, COM(2020) 65 final, 19 February 2020, [https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020\\_en.pdf](https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf).

145 Ibid., p. 10 ff.

system in Art. 17.<sup>146</sup> Data protection is a central cross-cutting theme of the draft.<sup>147</sup> In addition, the European Commission’s High-Level Expert Group on AI also published recommendations on policies regarding data economy, data access and protection of affected individuals in a paper in 2019.<sup>148</sup> The ethical guidelines developed by this group also contain requirements on technical security, privacy, transparency and fairness.<sup>149</sup>

In addition, an overview of relevant documents and blog posts by stakeholders can be found on the homepage of the European AI Alliance.<sup>150</sup>

#### Conclusion

There is a need for action in Germany in the coordination of data protection regulations with AI strategies. This is particularly the case with regard to No. 74 of the UNESCO Recommendation, which advises that already existing data protection regulations be strengthened in order to protect personal data, including particularly sensitive data. In addition, science and business should exchange with each other and work together more intensively and not predominantly conduct projects in parallel. In this respect, AI companies should, in accordance with No. 77 of the Recommendation, be encouraged to share the data they collect in order to promote research and innovation.

The initiative “FAIR Forward – Artificial Intelligence for All with Africa and Asia” addresses access to training data and framework conditions for better data protection.<sup>151</sup> The “Global South – AI4COVID” initiative also deals with framework conditions for data processing, but with a focus on COVID-19.<sup>152</sup> In addition, the initiative “Artificial Intelligence for Development in Africa (AI4D Africa)” addresses framework conditions for data use of AI.<sup>153</sup> The focus of these initiatives is primarily on ensuring that the Global South is not left behind on AI issues. In addition, AI should be developed by the countries themselves and not only for them. According to OECD.AI, only very few African countries have established AI initiatives.<sup>154</sup>

146 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, COM(2021) 206 final, 21 April 2021, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206>.

147 Ibid., Marginal Number 36.

148 High-Level Expert Group on AI, Policy and Investment Recommendations for Trustworthy AI, 26 June 2019, <https://digital-strategy.ec.europa.eu/en/library/policy-and-investment-recommendations-trustworthy-artificial-intelligence>.

149 HEG-KI.

150 See European AI Alliance, <https://futurium.ec.europa.eu/en/european-ai-alliance/pages/about?language=en>.

151 See FAIR Forward, Toolkit Digitalisierung, <https://toolkit-digitalisierung.de/en/fair-forward/>.

152 See Global South AI4COVID Program, <https://covidsouth.ai/about>.

153 See AI4D Africa, <https://africa.ai4d.ai/about-ai4d>.

154 See OECD.AI Policy Observatory, <https://oecd.ai/en/dashboards>.

### III.5. Development and International Cooperation (Policy Area 4)

Policy area 4 of the UNESCO Recommendation is dedicated to development and international cooperation. Provision No. 78 addresses UNESCO Member States and transnational corporations and calls on them to prioritise AI ethics by including respective topics in relevant international, intergovernmental and multi-stakeholder fora. To some extent, this has already happened. In the German government’s AI Strategy from 2018, “national and international networking” is mentioned as a separate field of action under point 3.11.<sup>155</sup> In concrete terms, cooperation and exchange with EU institutions as well as other economic areas should be strengthened.<sup>156</sup> Beyond, capacities and knowledge on AI should be built up in developing countries.<sup>157</sup> The interim report on the AI Strategy focusses on the German government’s monitoring and support at the European level.<sup>158</sup> It also emphasises Germany’s active role in the international economy via engagement in G7, G20 and OECD processes.<sup>159</sup> In addition, the report lists further measures that have already been implemented, such as international future laboratories or the initiative “FAIR Forward – Artificial Intelligence for All with Africa and Asia”.<sup>160</sup>

The updated version of the AI Strategy also explicitly mentions support for the creation of framework conditions for AI in the Global South.<sup>161</sup> This is in line with No. 79 of the UNESCO Recommendation, which calls on Member States to ensure that the use of AI in key sectors of development cooperation is in line with the values and principles set out in this Recommendation (these key sectors include education, science, culture, communication and information, health-care, agriculture and food supply, environment, natural resource management and infrastructure, economic planning and growth). The Data Ethics Commission’s report welcomes the choice of a European path that is different from other value systems, cultures and social models.<sup>162</sup> It is emphasised that the fundamental paradigms of this path should be incorporated even more in global and non-European discourses and contexts.<sup>163</sup> The final report of the Enquete Commission further mentions cooperation between research, business and society as one of the central recommendations for action.<sup>164</sup>

155 Bundesministerium für Wirtschaft und Energie (BMWi), Strategie Künstliche Intelligenz der Bundesregierung, 16 November 2018, <https://www.bmwi.de/Redaktion/DE/Publikationen/Technologie/strategie-kuenstliche-intelligenz-der-bundesregierung.html>.

156 Ibid., p. 42 f.

157 Ibid., p. 43.

158 Bundesregierung, Zwischenbericht ein Jahr KI-Strategie, November 2019, p. 8, available under “Downloads” on the website [www.ki-strategie-deutschland.de](http://www.ki-strategie-deutschland.de).

159 Ibid.

160 Ibid., p. 9.

161 NKIS.

162 DEK.

163 Ibid.

164 EKKI, p. 107.



A working group of the platform “Lernende Systeme” deals with the development field “Health & Care”.<sup>165</sup> In addition, a chapter in the German standardisation roadmap is dedicated to “AI in Medicine”.<sup>166</sup> Besides this, the German AI Association published a position paper entitled “How Artificial Intelligence can promote climate protection and sustainability” in February 2021.<sup>167</sup> Among other things, this addresses the creation of sustainability platforms for interdisciplinary cooperation as well as the development fields of energy and agriculture.<sup>168</sup> Furthermore, the funding programme “Digital GreenTech – Environmental Technology Meets Digitalisation” also focusses on the application fields of resource efficiency, sustainable agriculture and geotechnology.<sup>169</sup>

Germany must, however, increase its commitment to AI research cooperation, especially with LDCs (No. 81 and 82 of the UNESCO Recommendation). In more concrete terms, this commitment should include establishing research and innovation centres and networks with strong participation and leadership of researchers from the Global South.

The “White Paper on Artificial Intelligence – A European Approach to Excellence and Trust” calls for cooperation between the public and the private sector as one of the most important measures.<sup>170</sup> The European Commission’s High-Level Expert Group on AI also calls for stronger multi-stakeholder cooperation.<sup>171</sup> Beyond, it also cites the use of ethical guidelines for trustworthy AI in the development cooperation fields of climate protection and sustainable infrastructure, health and education.<sup>172</sup>

An overview of further relevant documents and blog posts by stakeholders can be found on the homepage of the European AI Alliance.<sup>173</sup>

165 See Plattform Lernende Systeme, <https://www.plattform-lernende-systeme.de/home-en.html>.

166 DIN and DKE, German Standardization Roadmap Artificial Intelligence, November 2020, <https://www.din.de/resource/blob/772610/e96c34dd-6b12900ea75b460538805349/normungsroadmap-en-data.pdf>, p.129 ff.

167 KI Bundesverband, Wie Künstliche Intelligenz Klimaschutz und Nachhaltigkeit fördern kann, 22 February 2021, <https://ki-verband.de/wp-content/uploads/2021/02/KIBV-Klima-Positionspapier-1.pdf>.

168 Ibid.

169 See FONA, [https://www.fona.de/de/massnahmen/foerdermassnahmen/DigitalGreenTech.php?utm\\_source=hootsuite](https://www.fona.de/de/massnahmen/foerdermassnahmen/DigitalGreenTech.php?utm_source=hootsuite).

170 European Commission, White Paper On Artificial Intelligence – A European approach to excellence and trust, COM(2020) 65 final, 19 February 2020, [https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020\\_en.pdf](https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf), p.3.

171 High-Level Expert Group on AI, Policy and Investment Recommendations for Trustworthy AI, 26 June 2019, <https://digital-strategy.ec.europa.eu/en/library/policy-and-investment-recommendations-trustworthy-artificial-intelligence>, p.48.

172 HEG-KI, p.32 ff.

173 See European AI Alliance, <https://futurium.ec.europa.eu/en/european-ai-alliance/pages/about?language=en>.

## Conclusion

In Germany, there is still a need for action with regard to the policy area “Development and International Cooperation”, especially when it comes to the cooperation of and with private enterprises. In most cases, politics and science are described as cooperation partners, but the private sector is hardly mentioned, especially with regard to development cooperation with the Global South. According to No. 82 of the UNESCO Recommendation, this would be urgently necessary in order to develop AI for specific cultures and contexts, among other things. Cooperation generally focusses on the European continent, which is particularly evident in the mentioning of the “European way”. Yet, individual initiatives and strategies mention the global character of cooperation initiatives, especially with emerging and developing countries. The focus here should also be on the inclusion of researchers from developing countries themselves, as stipulated in No. 81 of the Recommendation. In addition, there is still a need for regulation or action regarding the overcoming of geotechnological boundaries, as emphasised in No. 83.

The above-mentioned initiative “FAIR Forward – Artificial Intelligence for All with Africa and Asia” is an example of the active inclusion of the Global South.<sup>174</sup> The “Global South – AI4COVID” initiative supports interdisciplinary research on COVID-19 in developing countries.<sup>175</sup> In addition, the Artificial Intelligence for Development in Africa (AI4D Africa) initiative focusses on the Sustainable Development Goals (improving the food chain, education, health and climate change).<sup>176</sup> While in the European context the location of AI businesses is a key topic of the discourse, stakeholders in the Global South increasingly focus on the AI’s potential to foster sustainable development of the society and the environment. Cooperation is usually initiated by organisations and countries that are not part of the Global South.

An important commitment is made at the end of the section in No. 83 of the UNESCO Recommendation. It recalls that all cooperation is based on ethical values and “in full respect of international law”.

174 See FAIR Forward, Toolkit Digitalisierung, <https://toolkit-digitalisierung.de/en/fair-forward/>.

175 See Global South AI4COVID Program, <https://covidsouth.ai/about>.

176 See AI4D Africa, <https://africa.ai4d.ai/about-ai4d>.

## III.6. Gender (Policy Area 6)

There are few (political and regulatory) processes, instruments and governance structures in the policy area of gender (and diversity) in Germany and at the EU level that actively address AI. The term “gender” appears exactly once in the German Federal Government’s AI Strategy of November 2018<sup>177</sup>, namely in the area of risk impact assessment of AI in the context of gainful employment. Implicitly, this policy area is negotiated under the premise of non-discrimination. Thus, the general prohibition of discrimination is recalled at various points within the strategy.<sup>178</sup> “AI-based decisions, services and products” should be made reviewable with regard to potential inadmissible discrimination, according to the strategy.<sup>179</sup> In addition, “transparency, traceability and verifiability of AI systems” should ensure “effective protection against distortions, discrimination, manipulation or other abusive uses, especially in the use of algorithm-based forecasting and decision-making systems”,<sup>180</sup> whereby “discrimination against disadvantaged population groups in developing

177 Bundesministerium für Wirtschaft und Energie, Strategie Künstliche Intelligenz der Bundesregierung, 16 November 2018, <https://www.bmwi.de/Redaktion/DE/Publikationen/Technologie/strategie-kuenstliche-intelligenz-der-bundesregierung.html>.

178 Cf. *ibid.*, p.39.

179 Cf. *ibid.*, p.38 (translation by the author).

180 Cf. *ibid.*, p.39 (translation by the author).

countries” is also explicitly mentioned.<sup>181</sup> To ensure these imperatives, “the Federal Government is examining the establishment or expansion of government agencies and private audit institutions to monitor algorithmic decisions”.<sup>182</sup> This includes establishing auditing standards and standards for (technological) Impact Assessments. It also involves creating a structure in which all elements of the AI process are disclosed transparently and comprehensibly for the examination of fundamental rights criteria.<sup>183</sup>

No. 87 of the UNESCO Recommendation also obliges Member States to ensure that the potential of digital technologies and AI contributes to the achievement of gender equality and is fully exploited. Member States must ensure that the human rights and fundamental freedoms of girls and women, as well as their security and integrity, are not violated at any stage of the life cycle of AI systems. Furthermore, the previously described Ethical Impact Assessment should include a transversal gender perspective.

The strategy, measures and current developments are available on the website “www.ki-strategie-deutschland.de”. Also on this online platform, which describes the current political status in Germany, non-discrimination, gender and diversity is not considered a separate field of

181 Cf. *ibid.* (translation by the author).

182 Cf. *ibid.*, p.40 (translation by the author).

183 Cf. *ibid.*





action and hardly receives any attention, with the exception of promoting young female AI scientists. The interim report on the implementation of the strategy<sup>184</sup> published in November 2019 does not contain any statements on gender as a policy field either. This is also because neither the Federal Anti-Discrimination Agency<sup>185</sup> nor the Federal Ministry for Family Affairs, Senior Citizens, Women and Youth (BMFSFJ)<sup>186</sup> which is responsible for gender, were actively involved in the implementation of the strategy. Recently, the BMFSFJ has been coordinating three projects focusing on gender equality in the labour market under the umbrella of the AI Strategy. The update of the AI Strategy from 2020 also remains largely vague.<sup>187</sup> In its report on the strategy, however, the Data Ethics Commission, which has been established in the meantime, points out the dangers of discrimination through algorithms as well as the lack of regulatory approaches and effective law enforcement.<sup>188</sup> The final report of the Enquete Commission on Artificial Intelligence<sup>189</sup> argues in a similar way. In its recommendations for action,

184 Federal Government, Interim Report One Year AI Strategy (orig. Zwischenbericht ein Jahr KI-Strategie), November 2019, available under "Downloads" on the website [www.ki-strategie-deutschland.de](http://www.ki-strategie-deutschland.de). While the BMBF-funded Plattform Lernende Systeme already published a study on the topic in 2019: Artificial Intelligence and Discrimination, <https://www.plattform-lernende-systeme.de/publikationen-details/kuenstliche-intelligenz-und-diskriminierung-herausforderungen-und-loesungsansatze.html>; just as the KI Observatory funded by the BMAS focusses on the issues of gender and discrimination: <https://www.ki-observatorium.de/indikatoren#tab-638>.

185 Federal Anti Discrimination Agency, <https://www.antidiskriminierungsstelle.de/EN/> – They have certainly kept an eye on the issue and published a study on the topic of discrimination and algorithms in 2019: Carsten Orwat, Risks of Discrimination through the Use of Algorithms, [https://www.antidiskriminierungsstelle.de/SharedDocs/downloads/EN/publikationen/Studie\\_en\\_Diskriminierungsrisiken\\_durch\\_Verwendung\\_von\\_Algorithmen.html](https://www.antidiskriminierungsstelle.de/SharedDocs/downloads/EN/publikationen/Studie_en_Diskriminierungsrisiken_durch_Verwendung_von_Algorithmen.html). Here, as part of their conclusions, they demand, among other things: "The catalogues of protected characteristics laid down in the GG and AGG are to be examined to see whether new methods of analysis, in particular with Artificial Intelligence algorithms, for the automated identification of personality characteristics require their extension. This also makes characteristics identifiable and accessible for differentiation, which can be used to identify and use the dependency on a good, resource or position in order to establish or increase structural advantage. The still largely unknown connections between technically feasible and potentially endangered protection objectives should be researched and their legitimacy should be assessed from a societal perspective." (p.75) Further recommendations in the direction of regulation follow from p.89 ff.

186 BMFSFJ, <https://www.bmfsfj.de/bmfsfj/meta/en>. They too, for example, promote diversity in AI development, such as the initiative <https://www.jumpp.de/diversitaet-KI>; this topic is also strongly addressed in the Third Gender Equality Report of the Federal Government of 2021 (see e.g. p.17 ff.) <https://www.bmfsfj.de/resource/blob/184544/665a7070dbc68f9984fe968dc05fd139/dritter-gleichstellungsbericht-bundestagsdrucksache-data.pdf>. Here, too, there is definitely criticism of the orientation of the AI Strategy: "Other digitisation-related strategies such as the Federal Government's 'Artificial Intelligence Strategy' (2018) or the BMBF's digital strategy 'Digital Future: Learning. Research. Knowledge' (2019) lack a gender and women's policy approach (...). The discussions held with other digital bodies in the course of preparing the report confirm this finding of the German Women's Council." (translation by the author).

187 NKIS.

188 DEK.

189 EKKI.

it very clearly calls for "AI-supported decisions to be regularly reviewed on their non-discriminatory nature."<sup>190</sup>

On the European level, a first agenda on AI was published in April 2018.<sup>191</sup> It addresses the issue of diversity and non-discrimination mainly in the context of the development phase of AI systems. However, in this very phase, the document defines the problem in a broad and proactive way: "The development of AI needs to involve more women and people from different backgrounds, including people with disabilities, starting with inclusive AI education and training to ensure that AI is non-discriminatory and inclusive". In the context of AI at the EU level, gender is thus not only considered in terms of the outcome (discrimination) and a simple solution (equal treatment), but more broadly: as a complex of various, intersectional discriminations with different legal, societal and educational policy connecting factors.

Under the roof of fundamental and human rights, the European AI agenda classifies the principle of non-discrimination as being part of privacy and consumer protection rights. Besides regulation aimed at transparency, traceability and accountability of algorithms, there is also a call for more research to uncover actual patterns of discrimination to prove them scientifically. The European Commission's "White Paper on Artificial Intelligence – A European Concept for Excellence and Innovation" from 2020 takes up this issue and develops it further.<sup>192</sup>

190 Ibid., p.195 (translation by the author).

191 European Commission, Communication from the Commission – Artificial Intelligence for Europe, COM(2018) 237 final, 25 April 2018, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52018DC0237>; more background information on the complex also here: [https://ec.europa.eu/commission/presscorner/detail/en/QANDA\\_21\\_1683#2](https://ec.europa.eu/commission/presscorner/detail/en/QANDA_21_1683#2).

192 European Commission, White Paper On Artificial Intelligence – A European approach to excellence and trust, COM(2020) 65 final, 19 February 2020, [https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020\\_en.pdf](https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf); The White Paper is widely supported by the Federal Government: [https://www.ki-strategie-deutschland.de/files/downloads/Stellungnahme\\_BReg\\_Weissbuch\\_KI.pdf](https://www.ki-strategie-deutschland.de/files/downloads/Stellungnahme_BReg_Weissbuch_KI.pdf) – The White Paper, on the other hand, is often criticised by civil society. This is often because biometric mass surveillance technologies are not explicitly banned, and it remains unclear what legislative limits should be set for AI-based systems that violate fundamental rights. It is also not really clear on what basis the current list of high-risk applications was drawn up and what can be considered a valid reason for a future risk classification. Cf. e.g. AlgorithmWatch, AlgorithmWatch's response to the European Commission's proposed regulation on Artificial Intelligence – A major step with major gaps, 22 April 2021, <https://algorithmwatch.org/en/response-to-eu-ai-regulation-proposal-2021>; red lines for the use of AI that threatens fundamental rights are also demanded by a multitude of other actors. Cf. e.g. EDRI, Civil society calls for AI red lines in the European Union's Artificial Intelligence proposal, 12 January 2021, <https://edri.org/our-work/civil-society-call-for-ai-red-lines-in-the-european-unions-artificial-intelligence-proposal/>; The assessment of the German Women Lawyers Association (Deutscher Juristinnenbund), Opinion: 21-14, 30 June 2021, <https://www.djb.de/presse/stellungnahmen/detail/st21-14> is also very informative in this context.

The White Paper mentions the dangers of discrimination by AI with regard to gender already in the first paragraph. This awareness can be observed, though in parts rather superficially, throughout the entire document: "The use of AI can affect the values on which the EU is founded and lead to breaches of fundamental rights, including the rights to freedom of expression, freedom of assembly, human dignity, non-discrimination based on sex, racial or ethnic origin, religion or belief, disability, age or sexual orientation, as applicable in certain domains, protection of personal data and private life, or the right to an effective judicial remedy and a fair trial, as well as consumer protection."<sup>193</sup> It explicitly recognises that discrimination on the basis of gender or race has already been taken place by AI software that should predict recidivism of offenders or that is used in the field of facial recognition.<sup>194</sup>

An often-cited study commissioned by the Council of Europe also identifies a large number of fundamental rights that could be affected by the use of AI.<sup>195</sup> The Commission also supports the core requirement of "diversity, non-discrimination and fairness".<sup>196</sup>

It also makes clear that AI developers and users are already subject to existing European legislation on fundamental rights, but points out that these may need to be revised in the face of AI.<sup>197</sup> The focus of the EU's regulatory framework is on avoiding "the most serious harm", which explicitly includes the disrespect of fundamental rights, especially with regard to the principle of non-discrimination.<sup>198</sup> With regard to training data for AI systems, for example, the White Paper recommends the establishment of requirements that oblige developers to use "comprehensive data sets that are sufficiently representative".<sup>199</sup> Moreover, there is the mandatory documentation of the "programming and training methodologies (...) avoiding bias that could lead to prohibited discrimination".<sup>200</sup>

193 Ibid., p.10 ff.

194 Ibid., p.12.

195 Council of Europe, Algorithms and Human Rights – Study on the human rights dimensions of automated data processing techniques (in particular algorithms) and possible regulatory implications, March 2018, <https://rm.coe.int/algorithms-and-human-rights-en-rev/16807956b5>.

196 European Commission, White Paper On Artificial Intelligence – A European approach to excellence and trust, COM(2020) 65 final, 2020, [https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020\\_en.pdf](https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf), p.9; and also in the German Standardisation Roadmap Artificial Intelligence this passage appears emphatically: <https://www.din.de/resource/blob/772610/e96c34dd6b12900ea75b460538805349/normungsroadmap-en-data.pdf>, p.54f.

197 European Commission, White Paper On Artificial Intelligence – A European approach to excellence and trust, COM(2020) 65 final, 2020, [https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020\\_en.pdf](https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf), p.10 ff.

198 Ibid., p.11.

199 Ibid., p.18.

200 Ibid., p.19.

Furthermore, according to No. 88 of the UNESCO Recommendation, Member States must provide earmarked public funds to finance gender-equitable programmes and ensure that national digital policies include a gender equality action plan.

In its No. 87–93 on gender policy, the UNESCO Recommendation<sup>201</sup> has many similarities with the standards envisaged by the EU, some of which have already been launched nationally. However, the Recommendation goes much further. It includes the special protection of fundamental and human rights as well as of marginalised groups under the precept of non-discrimination. It also stipulates the promotion of diversity in the AI (development) sector, especially in terms of representation and relevance. Discriminatory biases, inherited through technology, should be prevented and the strengthening of inclusive, diverse AI development should be promoted. In its Recommendation, UNESCO lays a strong focus on representation and yet remains on the surface with its demands and measures for action.

The UNESCO Recommendation puts great emphasis on maintaining gender equality in AI development and use; Germany still has a lot of work to do in this regard. There is considerable room for improvement with regard to gender equality and diversity in AI development teams in Germany (No. 19). The promotion of women and increasingly also of diverse people in the field of natural sciences and STEM has been repeatedly articulated at the federal level for years and supported by various funding projects in the field of education and vocational training (addressed in No. 87–89 and 91–93 of the UNESCO Recommendation). However, Germany is still poorly positioned when it comes to gender equality in the development and research of AI as well as in the establishment of interdisciplinary teams.

In Germany only 16% of AI professionals are female (the global share is 22%), and women in the AI sector in Germany earn on average 25% less than their male colleagues.<sup>202</sup> To change this it needs, besides stereotype-sensitive early childhood education, explicit support for women and diverse people to sustainably and successfully enter the predominantly white, male technology development sector. Also more extrinsic incentives are needed and should be communicated: It is verified that more diverse teams lead to more diverse and also more successful products. In addition, there is a need for more visibility of women in AI as role models, for example through the dissemination of success stories.<sup>203</sup>

201 Cf. UNESCO, Recommendation on the Ethics of Artificial Intelligence, SHS/BIO/REC-AIETHICS/2021, 2021, <https://unesdoc.unesco.org/ark:/48223/pf0000380455>, p.15 f., Policy Area 6: Gender.

202 Cf. Hedda Nier, „Wie weiblich ist die IT?“, statista, 20 March 2018, <https://de.statista.com/infografik/13283/frauen-in-der-tech-branche>.

203 Cf. HIIG AI Lab, Positionspapier zum Roundtable „KI und Frauen\*“, December 2020, [https://www.hiig.de/wp-content/uploads/2020/12/Positionspapier-KI-und-Frauen-WEB\\_V2.pdf](https://www.hiig.de/wp-content/uploads/2020/12/Positionspapier-KI-und-Frauen-WEB_V2.pdf), p.7 ff.





Alongside accessible education for all ages, there is a need for a politically driven, serious structural change in the (business) logic of IT and AI companies.<sup>204</sup> The current imbalance in representation will not be fixed through education alone; strong incentives for structural and systemic change within the organisational cultures of companies are needed. Furthermore, a rethinking of AI development in an intersectional perspective is needed to make different patterns of exclusion visible and to overcome them.<sup>205</sup>

## Conclusion

Member States should ensure that AI does not exacerbate the large gender gaps that already exist in the analogue world in various fields, but rather eliminate these differences. These gaps include: the gender pay gap; unequal representation in certain professions and activities; lack of representation in executive positions, supervisory boards or research teams; the education gap; the gap in access to digital content; the unequal distribution of unpaid work and care responsibilities. Member States should also, according to No. 90 of the UNESCO Recommendation, ensure that gender stereotypes and discriminatory biases are not incorporated

<sup>204</sup> Ibid., p. 11.

<sup>205</sup> Cf. *ibid.*, p. 12.

into AI systems, but that they are identified and proactively eliminated. As one group of experts recently formulated: “AI development must be thought of intersectionally in order to break through manifest systems of exclusion and make rigid category systems visible. This requires not least an intensive examination of the question of access to material and immaterial resources, opportunities for access and participation, opportunities for education and involvement.”<sup>206</sup>

According to the OECD, there are initiatives on AI and gender in six African countries.<sup>207</sup> Of course, there are individual initiatives, partly also funded by Germany, that push the issue, such as “FAIR Forward – Artificial Intelligence for All with Africa and Asia”.<sup>208</sup> However, these initiatives only deal marginally with the topic of gender, diversity and non-discrimination. In the initiative “Artificial Intelligence for Development in Africa” (AI4D Africa), the promotion of inclusive, diversity-sensitive AI is granted a stronger focus, but has not yet been translated into a regulatory framework.<sup>209</sup>

<sup>206</sup> *Ibid.* (translation by the author).

<sup>207</sup> Cf. OECD, Dashboards, <https://oecd.ai/en/dashboards>.

<sup>208</sup> FAIR Forward, Toolkit Digitalisierung, <https://toolkit-digitalisierung.de/en/>.

<sup>209</sup> Cf. Artificial Intelligence for Development in Africa, <https://africa.ai4d.ai/blog/ai4d-africa-launch-driving-inclusive-ai-innovation-in-africa>.

With regard to the commitment that gender equality must not be violated at any stage of the life cycle of AI systems, Germany is also still at the very beginning (No. 90 of the Recommendation): Ensuring that gender stereotypes and discriminatory biases are not incorporated into AI systems but identified and proactively eliminated, will lead to substantial changes in the state’s approach to AI-based decision-making processes.

Equal treatment also relates to data policy, as many data with a gender dimension or those related to sexual orientation are very sensitive. In No. 74 of the Recommendation, Member States commit to providing special protection for sensitive data. These include: Data on offences, criminal proceedings, convictions and related security measures; biometric, genetic and health data; personal data such as data on race, colour, descent, sex, age, language, religion, political opinion, national origin, ethnic origin, social origin, economic or social circumstances of birth or disability and other characteristics.

Concerning “remote biometric identification” it becomes at least implicitly apparent how highly sensitive these technologies are in relation to other fundamental rights. Clearly, any infringement of these fundamental rights must be urgently prevented.<sup>210</sup> Even more concrete is the European Commission’s proposal for an AI Regulation.<sup>211</sup> Here, Articles 13, 15, 17, 33, 35–39, 44, 45 and 47 take up in great detail what is already laid down in the White Paper: namely, the unconditional prevention of the reinforcement of (implicit) biases by AI used in highly sensitive areas and trained with datasets that are not free of intersectional discrimination.<sup>212</sup>

<sup>210</sup> *Ibid.*, p. 25.

<sup>211</sup> 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, COM(2021) 206 final, 21 April 2021, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206>, more background information and timelines here: <https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=COM:2021:205:FIN>.

<sup>212</sup> *Ibid.*, similar approaches with regard to preventing biases in data sets and AI, the more diverse composition of developer teams and the transparency factor can also be found in the High-Level Expert Group on AI, e.g. here: Policy and Investment Recommendations for Trustworthy AI, 26 June 2019, <https://digital-strategy.ec.europa.eu/en/library/policy-and-investment-recommendations-trustworthy-artificial-intelligence> and here: Ethics guidelines for trustworthy AI, 8 April 2019, <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>: “Diversity, non-discrimination and fairness: Unfair bias must be avoided, as it could have multiple negative implications, from the marginalisation of vulnerable groups, to the exacerbation of prejudice and discrimination. Fostering diversity, AI systems should be accessible to all, regardless of any disability, and involve relevant stakeholders throughout their entire life cycle.” As well as here: European AI Alliance, <https://futurium.ec.europa.eu/en/european-ai-alliance/>.

Surveillance technologies endanger marginalised groups all over the world. Their use and the conditions of their use must always be critically questioned with regard to law enforcement, the welfare state and all other areas of society. They must be continuously evaluated independently, comprehensively and intersectionally against the background of postcolonial power relations and under the umbrella of fundamental and human rights.<sup>213</sup>

<sup>213</sup> Cf. netzforma\* e. V., Wenn KI – dann feministisch, Berlin 2020, [https://netzforma.org/wp-content/uploads/2021/01/2020\\_wenn-ki-dann-feministisch\\_netzforma.pdf](https://netzforma.org/wp-content/uploads/2021/01/2020_wenn-ki-dann-feministisch_netzforma.pdf), Kapitel Überwachung, p. 15–30 and p. 206 ff.



Conclusion





## IV. Conclusion

Most ethical guidelines on AI propagate general principles and do not aim at practical implementation. Yet, the UNESCO Recommendation on the Ethics of AI is the first international legal text negotiated in an open and inclusive process that translates ethical principles into concrete tasks for policy-making. The Recommendation includes globally accepted ethical standards for AI technologies with full respect for human rights.

It is remarkable that all 193 UNESCO Member States have committed to the Recommendation. The preceding two-year development process, in which non-state actors were also able to contribute their voices, added significant value in terms of legitimacy. As regards content, the holistic approach of the Recommendation stands out. On a general level, it becomes clear that the Recommendation considers AI ethics a task for society. In order to fulfil it, citizens must be made aware of the potential and risks of AI. On the policy level, the Recommendation also defines AI ethics as a directly applicable set of normative standards. Thus, it formulates concrete tasks for national action on AI development and use in eleven policy areas, including education and science, communication, health and the environment. The focus is not on abstract demands, but on concrete improvements.

The Recommendation makes use of the advantages of ethics over law, particularly with regard to policy development and implementation, and in clarifying legal norms, such as equal treatment obligations, which are extended to AI applications. The reference to ethics as well as the soft law status of a Recommendation allowed both a strong human rights language and a focus on practical implementation to be embedded in the document. This would not have been possible with the objective of a “hard” global AI law, as normative orders are often only just being created. The consideration of previous “blind spots” in AI law, such as the impact of AI on gender, sustainability, development policy or education is also noteworthy.

**“It is now up to the Member States to fulfil their obligation under international law and implement the Recommendation.”**

Furthermore, the Recommendation defines procedural and transparency obligations: Specifically, it calls for Ethical Impact Assessments for the development and use of AI systems, which in turn should occur in a multi-stakeholder-based, open and responsive process.

A further added value of the UNESCO Recommendation is its reference to cultural aspects possibly affected by the use of AI systems. These include potentially positive effects such as the preservation of endangered languages, but also negative effects like new access barriers on cultural digital platforms. For the first time, the Recommendation also focusses on the particular dynamics that countries of the Global South may face in the use and development of AI systems. It not only points out the potentially higher vulnerability of developing countries, but also highlights their ability to contribute to fairer AI use and better AI regulation. By emphasising the importance of certain marginalised groups and the need for their protection, the UNESCO Recommendation can provide an important impetus for the EU’s legislative process. A sufficient level of protection can thus be ensured through suitable regulatory instruments and mechanisms.

The UNESCO Recommendation on the Ethics of AI is the first global consensus on the ethical development and use of Artificial Intelligence. It opens a unique opportunity to guide AI development in an ethical manner and thus to actively shape it in the sense of human (rights). At this point, it is now up to the Member States to fulfil their obligation under to international law and implement the Recommendation. The laws, programmes and funding instruments that are now adopted on the basis of the Recommendation will, in retrospect, be the yardstick for its success.

# Imprint

## Editor

---

Deutsche UNESCO-Kommission e.V./  
German Commission for UNESCO  
Martin-Luther-Allee 42  
53175 Bonn

### Authorised to represent:

Prof Dr Maria Böhmer (President)  
Prof Dr Christoph Wulf (First Vice-President)  
Prof Dr Hartwig Carsten Lüdtke (Second Vice-President)  
Dr Roman Luckscheiter (Secretary-General)  
Dr Lutz Möller (Special Representative in accordance with § 30 BGB)

Legal Form: Registered Association  
Headquarters: Bonn; Register court: Amtsgericht Bonn; Association register number: VR 4827

## Editors and translation

---

Jeannine Hausmann (responsible),  
Dr. Maximilian Müngersdorff,  
Anne Diessner

## Design and Layout

---

Panatom, Berlin

## Copyright

---

The texts of this publication are licensed under the Creative Commons License Attribution 3.0 Germany (CC BY 3.0 NC DE) ■ <https://creativecommons.org/licenses/by-sa/3.0/deed.de>. The photos are excluded from the licence.

## ISBN

---

978-3-947675-28-9

## Bibliographic Information of the German National Library

The German National Library lists this publication in the German National Bibliography; detailed bibliographic data are available on the Internet at ■ <http://dnb.d-nb.de>.

## Authors

---

Matthias C. Kettemann is Professor for Innovation, Theory and Philosophy of Law and head of the Department of Legal Theory and Future of Law at the University of Innsbruck, head of a research programme on private online rule-making at the Leibniz Institute for Media Research | Hans-Bredow Institute, head of the Research Group “Global Constitutionalism and the Internet” at the Humboldt Institute for Internet and Society, Berlin, and of the Research Group “Platform and Content Governance” at the Sustainable Computing Lab at the Vienna University of Economics and Business.

Philipp Jaud is a researcher and Lisa-Maria Riedl a student assistant at the Department for Theory and Future of Law at the University of Innsbruck.

Martin Fertmann, Vincent Hofmann, Katharina Mosene and Johannes Schmees are research assistants (Junior Researchers) at the Leibniz Institute for Media Research | Hans-Bredow-Institut.

## Photo credits

---

p.10 © Martin Hieslmair  
p.15 © Gerd Altmann  
p.19 © iStock, metamorworks  
p.21 © Adi Goldstein  
p.24 © Martin Wagenhan  
p.29 © This is Engineering  
p.32 © iStock, Peter Howell  
p.42 © iStock, Keattisak A  
p.45 © vog.photo



