# AI CENTRE MGIMO
EMPOWER INNOVATIONS

# MGIMO UNIVERSITY

February 2022

## Anna Abramova
## Anastasia Ryzhkova
## Iuliia Tserekh

# AI ETHICS ASSESSMENT AT NATIONAL AND INTERNATIONAL LEVELS. THE APPROACH TO INDEX FRAMEWORK AND METHODOLOGY

## AI FOR DEVELOPMENT research paper collection

приоритет2030^
Лидерами становятся

MOSCOW STATE INSTITUTE OF INTERNATIONAL

RELATIONS

(UNIVERSITY) OF THE MINISTRY OF FOREIGN AFFAIRS OF RUSSIA

MGIMO Centre For AI

Anna Abramova, Anastasia Ryzhkova, Iulia Tserekh

## "AI ETHICS ASSESSMENT AT NATIONAL AND INTERNATIONAL LEVELS. THE APPROACH TO INDEX FRAMEWORK AND METHODOLOGY"

**Research paper**

Moscow,2022

**Authors:**

**Anna Abramova**, PhD, Director MGIMO Centre for AI, Head of the Department of Digital Economy and Artificial Intelligence of the ADV group at MGIMO-University

**Anastasia Ryzhkova**, PhD, researcher MGIMO Centre for AI

**Iulia Tserekh**, junior researcher MGIMO Centre for AI

**Abstract**

Artificial intelligence (AI) ethics become one of the essential elements of soft law in regulating national and international market. December 2021 UNESCO adopted the Recommendation on the ethics of artificial intelligence that provides the approaches for international soft regulation putting ethics in the heart. The Global AI Ethics index framework could be the basis for ethical impact assessment in alliance with the Recommendation and OECD AI Principles and the framework for AI classification. Methodology places the human in the center and includes all the key stakeholders along the AI system lifecycle. The data for the index could be taken from the existing databases of UNESCO, OECD, UNCTAD, WBG. But specificity of the topic enlarges on clarification and elaboration of the data that could be added to national statistics tables at macrolevel and also from the surveys, that cover the microlevel.

**"AI ethics assessment at national and international levels. The approach to index framework and methodology"** is the first publication in AI FOR DEVELOPMENT research paper collection

Cover photo: canva.com

**Moscow, 2022**

# Contents

## List of abbreviations

| | |
|---|---|
| AI | Artificial intelligence |
| ICT | Information-communication technology |
| IP | Intellectual property |
| NGO | Non-governmental organization |
| OECD | Organisation for Economic Co-operation and Development |
| SDGs | Sustainable Development Goals |
| UNCTAD | United Nations Conference on Trade and Development |
| UNESCO | United Nations Educational, Scientific and Cultural Organization |
| R&D | Research and development |
| WEF | World Economic Forum |
| WBG | World Bank Group |
| WHO | World Health Organisation |

# Introduction

*"We need international and national policies and regulatory frameworks to ensure that these emerging technologies benefit humanity as a whole.*

*We need a human-centered AI. AI must be for the greater interest of the people, not the other way around"*

UNESCO, Recommendation on the ethics of artificial intelligence, 2021

Digital technologies are developing rapidly and are applied almost in every sphere of human life. Artificial intelligence (AI) is one of the key ones having extremely high potential for future economy development. Innovative policy solutions in AI sector should not only be determined[1] but already implemented and measured.

AI is one of the core elements of the Fourth Industrial Revolution, which is capable for fundamental structural changes in the majority of aspects of life, from production to interpersonal communication. For example, the World Bank Group emphasizes the importance of non-cognitive skills[2], highlighting AI contribution to social and economic progress. Intelligent control systems for equipment, automated systems for the provision of public services, intelligent systems in the field of financial technologies, artificial neural networks for a wide range of tasks, etc. are being actively introduced and effective measurement tools are deeply necessary for business and society[3].

---

[1] OECD: The OECD Development Pathways series [E-resource]. Available at: URL https://www.oecd-ilibrary.org/development/oecd-development-pathways_23087358

[2] World Bank Group: Socioemotional Skills Development in Highly Violent Contexts Measurements and Impacts [E-resource]. Available at: URL https://documents1.worldbank.org/curated/en/573591646665166695/pdf/Socioemotional-Skills-Development-in-Highly-Violent-Contexts-Measurements-and-Impacts.pdf

[3] OECD Digital Economy Papers [E-resource]. Available at: URL https://www.oecd-ilibrary.org/science-and-technology/tools-for-trustworthy-ai_008232ec-en

Ethics and principles of soft law regulation in AI development are widely discussed within all the group of stakeholders. At international level the discussion run under agenda of UNESCO, OECD, UNCTAD, the Council of Europe and the list of organizations and initiatives is enlarging from year to year. The UNESCO currently is one of the key organizations in terms of AI regulation system development considers this technology as a service with ethical challenges to be answered in a short run in their "Recommendation on the ethics of artificial intelligence".

All the organizations mentioned above put human in the heart of the AI regulation system that is on the way now. For example, UNESCO in the "Recommendation on the ethics of artificial intelligence" emphasizes the importance of a human-centred AI approach.[4] The OECD Framework for the Classification of AI Systems also puts an emphasize on human-centric and AI-trustworthy approach[5]. Moreover, the Council of Europe, UNCTAD, the European Commission, and the WHO also make a significant contribution initiating the discussions on ethics in AI being supported with relevant research such as Digital economy Report 2021, UNCTAD[6].

AI ethics assessment contribute the achievement of Sustainable Development Goals (SDGs) achievement, especially goal 3 (Good Health and Well-being), goal 4 (Quality Education), goal 5 (Gender Equality), goal 8 (Decent Work and Economic Growth), goal 9 (Industry, Innovation and Infrastructure), goal 10 (Reduced Inequality), goal 11 (Sustainable Cities and Communities), goal 16 (Peace and Justice Strong Institutions), goal 17 (Partnerships to achieve the Goal).

---

[4] UNESCO: Recommendation on the ethics of artificial intelligence [E-resource]. Available at: URL https://en.unesco.org/artificial-intelligence/ethics

[5] OECD 2022. FRAMEWORK FOR THE CLASSIFICATION OF AI SYSTEMS [E-resource]. Available at: URL https://www.oecd-ilibrary.org/docserver/cb6d9eca-en.pdf?expires=1646822229&id=id&accname=guest&checksum=0D6117C31817EA6B9FFDB7A65AACAFCF

[6] UNCTAD 2021: Digital economy report [E-resource]. Available at: URL https://unctad.org/webflyer/digital-economy-report-2021

Assessment through indexes is used for AI sector analysis covering technical and economic aspects but AI ethics metrics are the gap for the current state of AI research (Government AI Readiness Index[7], AI Watch Index[8], AI Index Report[9] and etc)

In this regard, we propose the framework for possible global AI ethic assessment through index putting in the center the following stakeholders: government, business, civil society, research centres/ think tanks. The framework contains relevant groups of indicators showing ethical aspects both for developers and users.

The structure of the research is follows – the first section is introduction, the second is devoted to methodology and the final one is focused on the discussion covering the most challenging issues in practical implementation of the framework.

The authors are grateful to the leaders and coordinators of the National Priority 2030 project for making it possible to conduct the research.

[7] OXFORD 2021. GOVERNMENT AI Readiness Index 2021. [E-resource]. Available at: URL https://www.oxfordinsights.com/government-ai-readiness-index2021

[8] EUROPEAN COMISSION 2021. AI watch index. [E-resource]. Available at: URL https://op.europa.eu/en/publication-detail/-/publication/15568192-a95f-11eb-9585-01aa75ed71a1/language-en/format-PDF/source-209026200

[9] STANFORD 2021. THE AI INDEX REPORT. [E-resource]. Available at: URL https://aiindex.stanford.edu/report/

# Methodology

Global AI ethics assessment framework requires multidisciplinary approach both for indicators and results analysis that implement quantitative and qualitive insights connected with the country's engagement in AI industry development. The general framework for AI ethics assessment relies on UNESCO Recommendation[10], OECD AI principles and OECD AI classification[11]. Recommendation on the ethics of artificial intelligence that provides the approaches for international soft regulation put ethics in the heart. OECD AI principles and OECD AI classification could be applied as the essential background for AI market developments assessment[12].

The parts of the frameworks were developed through analysis of the key reports and research papers from UNESCO, OECD, UNCTAD, WEF, case studies from the leading companies and expert consultations with the representatives of different groups of stakeholders that are interconnected with all stages of AI lifecycle. The framework provides the basis for global AI ethics index construction that support the ability of the society to understand the capacity building needs with regard to AI transformation of all the parts of the society. AI ethics index provides understanding of the average level of the countries' readiness to face the wide range of challenges (incl. societal ones, regulatory, risks mitigation) putting in the center humans and society. This approach provides the opportunity to indicate the level of digital inequality in connection with AI development and application, development challenges at a short and long run.

---

[10] UNESCO 2021. Recommendation on the ethics of artificial intelligence. [E-resource]. Available at: URL https://unesdoc.unesco.org/ark:/48223/pf0000380455

[11] OECD 2022. FRAMEWORK FOR THE CLASSIFICATION OF AI SYSTEMS [E-resource]. Available at: URL https://www.oecd-ilibrary.org/docserver/cb6d9eca-en.pdf?expires=1646822229&id=id&accname=guest&checksum=0D6117C31817EA6B9FFDB7A65AACAFCF

[12] OECD 2022. FRAMEWORK FOR THE CLASSIFICATION OF AI SYSTEMS [E-resource]. Available at: URL https://www.oecd-ilibrary.org/docserver/cb6d9eca-en.pdf?expires=1646822229&id=id&accname=guest&checksum=0D6117C31817EA6B9FFDB7A65AACAFCF

Global AI ethics index framework (fig.1) contains 19 groups of indicators. These proposed groups are deeply interconnected with each other and the stakeholders along AI lifecycle such as GOVERNMENT, BUSINESS, CIVIL SOCIETY, RESEARCH CENTRES. Also we include in the framework three subindexes that are influencing all the actors mentioned above – AI literacy, R&D investments and ICT infrastructure development.

The framework should be read line by line, accept common to all or a part of stakeholders' indicators.

A mathematical apparatus of the assessment consists of an Index of ethical development (InED), which is actually based on each stakeholder's Indicator of significance and Indicators of AI literacy, R&D investments and ICT infrastructure development, AI applications for education.

The key indicator by which Global AI ethics assessment is formed should be considered is InED.

We propose to calculate this Index as

$$InED^i = \sum \gamma_r (r_j^i + r_{j,lit}^i + r_{j,R\&D}^i + r_{j,ICT}^i + r_{j,edu}^i) \qquad (1)$$

where

InED - meaning of Index of ethical development for specific country,

i - country identifier,

j – stakeholder's identifier,

r – indicator's identifier,

$\gamma_r$ - weight factor for each Indicator (considers national highlights of AI ethical development strategies)

$r_j^i$ – meaning of Indicator of significance of each stakeholder (business, government, etc.) for specific country,

$r_{j,lit}^i$ — meaning of Indicator of AI literacy of each stakeholder (business, government, etc.) for specific country

$r_{j,R\&D}^i$ — meaning of Indicator of R&D investments of each stakeholder (business, government, etc.) for specific country

$r_{j,ICT}^i$ — meaning of Indicator of ICT infrastructure development of each stakeholder (business, government, etc.) for specific country

$r_{j,edu}^i$ meaning of Indicator of AI applications for education of each stakeholder (business, government, etc.) for specific country

Indicators of significance should be described as one in formula 2.

$$r_i = \alpha_i^j \beta_i^j \qquad (2)$$

where

i - country identifier,

j — stakeholder's identifier,

$r_i$ — meaning of Indicator of significance,

$\alpha_i^j$ - weight factor for each component, that describes stakeholders' activity and national strategy's interests in specific sphere.

$\beta_i^j$ - meaning of each component (sub indicator), that describes specific stakeholders' activity and it's contribution into AI ethic development.

*More details about calculation operations components of each Index of significance and InED are described in subsequent research papers of MGIMO Centre for AI.*

**Figure 1. Global AI ethics index framework**
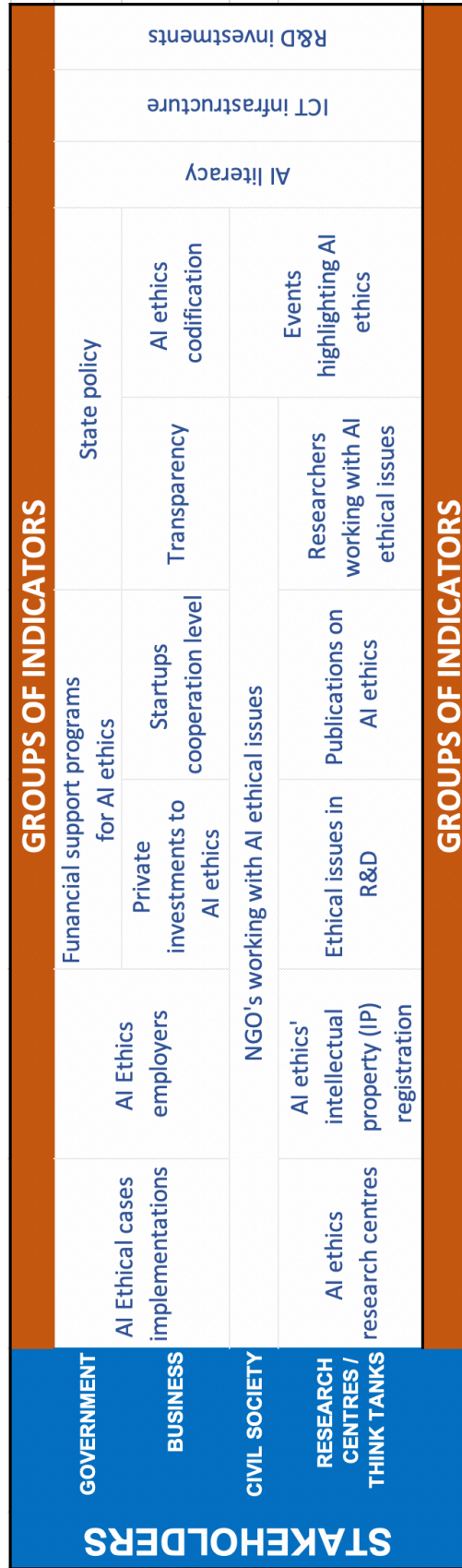


Source: MGIMO Centre for AI, 2022

## GOVERNMENT:

AI ETHICAL CASES IMPLEMENTATION shows the quantity and quality of ethical aspects application examples of the use of AI

AI ETHICS EMPLOYERS - a group of indicators reveals the characteristics of key employers in the field of AI ethics, and shows aspects of their application of ethical approaches and development

FINANCIAL SUPPORT PROGRAMS FOR AI ETHICS discloses possible and available state support programs for AI ethics and considers the volumes and directions for finance at the national level.

STATE POLICY – a complex group of indicators, that contains quantity and quality of state initiatives in the field of law, codification, as well as other related aspects of the development of AI ethics.

## BUSINESS:

AI ETHICAL CASES IMPLEMENTATION shows the quantity and quality of ethical aspects application examples of the use of AI

AI ETHICS EMPLOYERS - a group of indicators reveals the characteristics of key employers in the field of AI ethics and reflects the level of ethical approaches& development already implemented

PRIVATE INVESTMENTS TO AI ETHICS reveals the volumes and qualitative composition of private investments (including venture capital) in the field of AI ethics.

STARTUPS COOPERATION LEVEL indicates the qualitative and quantitative composition of the start-up solutions implementation in large companies and the processes of M&A in field of AI ethics.

TRANSPARENCY discloses information about business understanding of the ethical aspects of AI solutions work. It also reveals business approaches to developing the transparency of both implemented and planned AI solutions.

AI ETHICS CODIFICATION reveals approaches to the implementation of ethical principles in the field of AI ethics by business with real-life case disclosing.

## CIVIL SOCIETY:

NGO's working with AI ethical issues shows the level of involvement of national and international NGO's in prevention and mitigation of AI ethics violations.

EVENTS WITH HIGHLIGHTING AI ETHICS revels the annual activities for AI ethics promotion for different groups of stakeholders. It makes possible for AI developers and consumers get better understanding of the main concerns and risks arising from enlarging use of AI applications

## RESEARCH CENTRES / THINK TANKS:

AI ETHICS RESEARCH CENTRES group of indicators could be split in two groups: those who are AI ethics focused and those having AI ethics as a part of applied or fundamental research.

AI ETHICS' IP REGISTRATION covers the indicators that are relevant to all the forms of registration connected with AI ethic challenges mitigation

ETHICAL ISSUES IN R&D show the state of ethical concerns among the researchers and developers both in ICT and other sectors where AI applications are used

PUBLICATIONS ON AI ETHICS reveals the volume of peer-reviewed publications in English and national languages for both national data-bases and international ones

EVENTS HIGHLIGHTING AI ETHICS covers national and international conferences and forums on AI Ethics that could provide the chance for all the stakeholders in exchanging the view on further research challenges

RESEARCHERS WORKING WITH AI ETHICAL ISSUES include those who are focused on their activities in the field of AI ethics analysis


COMMON TO ALL STAKEHOLDERS:

ICT INFRASTRUCTURE shows the software and hardware development that is necessary for implementation and progress of AI ethics on national level for each stakeholder.

R&D INVESTMENTS reveals the volumes and qualitative composition of investments into R&D by state institutions, business, civil organizations and research centers separately.

AI LITERACY reveals the efforts of the public and private sector in ICT education, focusing on AI for different groups of stakeholders and for different age groups

AI APPLICATIONS FOR EDUCATION covers the metrics from schools and universities in using AI for education programs delivery (as a tool) and use of AI applications for arranging the social life (incl. technical assistance, marketing etc.)

# Discussion

Methodology places the human in the center of the assessment and includes all the key stakeholders along the AI system lifecycle. The data for the index could be taken from the existing databases of UNESCO, OECD, UNCTAD, ITU, WBG. These international organizations provide the databases on economic and social developments, incl. ICTs infrastructure, education, progress in regulation at national level etc. But still all this data is not enough for the analysis of digital economy development[13]. AI ethics is rather narrow topic that needs additional matrix of data for international comparisons. The main problem that in this case we face the challenge of analyzing AI subsectors and products that pushes us to microlevel. The specificity of the topic enlarges on clarification and elaboration of the data that could be added to national statistics tables at macrolevel and also from the surveys, that cover the microlevel.

The other aspect is the cultural perception of AI and its role in social life. Understanding of this issue also brings additional complexity making it difficult to find the consensus between all the actors.

Also, national sovereignty must be taken into consideration. As not all the segments of AI sector could be covered by the surveys because of national security concerns. And the first steps in national cloud initiatives are showing the complexity of international cooperation in digital economy.

---

[13] https://unctad.org/meeting/working-group-measuring-e-commerce-and-digital-economy-second-meeting

# References

**AI Ethics: Another Step Closer to the Adoption of UNESCO's Recommendation.** [E-resource]. Available at: URL https://en.unesco.org/news/ai-ethics-another-step-closeradoption-unescos-recommendation-0

**EUROPEAN COMISSION 2021. AI watch index.** [E-resource]. Available at: URL https://op.europa.eu/en/publication-detail/-/publication/15568192-a95f-11eb-9585-01aa75ed71a1/language-en/format-PDF/source-209026200

**Multistakeholder group discusses ten building blocks towards creating inclusive AI policies.** [E-resource]. Available at: URL https://en.unesco.org/news/multistakeholder-group-discusses-ten-building-blocks-towards-creating-inclusive-ai-policies

**Multistakeholder group discusses ten building blocks towards creating inclusive AI policies.** [E-resource]. Available at: URL https://en.unesco.org/news/multistakeholder-group-discusses-ten-building-blocks-towards-creating-inclusive-ai-policies

**OECD 2022. FRAMEWORK FOR THE CLASSIFICATION OF AI SYSTEMS** [E-resource]. Available at: URL https://www.oecd-ilibrary.org/docserver/cb6d9eca-en.pdf?expires=1646822229&id=id&accname=guest&checksum=0D6117C31817EA6B9FFDB7A65AACAFCF

**OECD 2022. FRAMEWORK FOR THE CLASSIFICATION OF AI SYSTEMS** [E-resource]. Available at: URL https://www.oecd-ilibrary.org/docserver/cb6d9eca-en.pdf?expires=1646822229&id=id&accname=guest&checksum=0D6117C31817EA6B9FFDB7A65AACAFCF

**OXFORD 2021. GOVERNMENT AI Readiness Index 2021.** [E-resource]. Available at: URL https://www.oxfordinsights.com/government-ai-readiness-index2021

**STANFORD 2021. THE AI INDEX REPORT.** [E-resource]. Available at: URL https://aiindex.stanford.edu/report/

**UNESCO 2019. Beijing Consensus on Artificial Intelligence and Education.** [E-resource]. Available at: URL https://unesdoc.unesco.org/ark:/48223/pf0000368303

**UNESCO 2021. Forum on AI and Education engages international partners to ensure AI as a common good for education.** [E-resource]. Available at: URL https://en.unesco.org/news/unesco-forum-ai-and-education-engages-international-partners-ensure-ai-common-good-education

**UNESCO 2021. Forum on AI and Education engages international partners to ensure AI as a common good for education.** [E-resource]. Available at: URL https://en.unesco.org/news/unesco-forum-ai-and-education-engages-international-partners-ensure-ai-common-good-education

**UNESCO 2021. Recommendation on the ethics of artificial intelligence.** [E-resource]. Available at: URL https://unesdoc.unesco.org/ark:/48223/pf000038

**MGIMO Centre for AI** was established to enhance international cooperation and support collaboration with all the actors of digital economy both at national and international levels. Our multidisciplinary research is focused on international cooperation agenda, national policies for AI and business opportunities. International trade and trade policy (prioritising digital trade), sustainable development, AI ethics are the key areas of our activities.

On the basis of MGIMO-University we promote an international AI expert platform with regular conferences and round tables, peer-revied articles and research papers. Our enlarging network of strategic partneships makes it possible to provide AI consulting and policy solutions both for business and goverment agencies.

**The Centre was founded in October, 2021**

**MGIMO Centre for AI research paper collection**
- Artificial intelligence in education

Artificial intelligence in education -Common aspects. February,2022
- Artificial intelligence for development

Global AI ethics assessment – the approach to index framework and methodology. February,2022

**Monthly digest**

 Digital economy bulletin

**Reports**

Discussions on Artificial Intelligence Ethics: Development Tracks by Key Groups of Actors,2021

**Annual conference AI Global dimension**

**MGIMO Digital Discussion club**

**Round tables&workshops**

**These documents& other Centre updates are available at https://aicentre.mgimo.ru/activities/research/papers**

**We hope to develop cooperation and we are open to any partnership offerings!**

**Our contacts**

143007, Moscow Region, Odintsovo,
Novo-Sportivnaya street, 3
https://aicentre.mgimo.ru
E: aicentre@inno.mgimo.ru
P: +7 903 623-95-15
https://t.me/aicentremgimo