



Anna Abramova
Anastasia Ryzhkova
Iuliia Tserekh

GLOBAL AI ETHICS ASSESSMENT KEY SUBINDEXES: GOVERNMENT AND STATE POLICY

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. KEY SUBINDEXES: GOVERNMENT AND STATE POLICY”**
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. 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 includes several subindexes in accordance with AI lifecycle and covering all the stakeholders involved. In this regard government and state policy is one of the key subindexes for further estimations of AI ethics. The groups of indicators it is based in includes AI employees, AI cases, state policy development tracks. The levels of estimation from federal to cities could vary depending on the country and its administrative structure. It is evident that some of the indicators could be assessed only with reliance on the survey data.

“AI ETHICS ASSESSMENT AT NATIONAL AND INTERNATIONAL LEVELS. KEY SUBINDEXES: GOVERNMENT AND STATE POLICY” is the part of AI FOR DEVELOPMENT research paper collection from MGIMO Centre for AI

JEL F01, F20, F42, F53, F55, F60, F63, F68

Key words: artificial intelligence, ethics, index, state policy, government

© 2022 MGIMO. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Cover photo: canva.com

Moscow, 2022

Contents

List of abbreviations.....	4
Introduction.....	5
National priorities for AI: main tracks and challenges.....	6
Methodology for Global AI ethics subindex Government	8
Discussion.....	12
References.....	12
Appendix 1.....	14

List of abbreviations

AI	Artificial intelligence
ICT	Information-communication technology
IP	Intellectual property
OECD	Organisation for Economic Co-operation and Development
UN	United Nations
UNCTAD	United Nations Conference on Trade and Development
UNESCO	United Nations Educational, Scientific and Cultural Organization

Introduction

This working paper is the next step in the ongoing research of MGIMO Centre for AI on elaboration of Global AI ethics index. This research is focused on detailed coverage of state policy and government activities and priorities that influence AI ethics perception and its understanding at state level.

The role of AI in public administration and the importance of a state policy and regulation in AI life cycle is increasing steadily for more than last 5 years. On one hand, AI regulation has become the next evolutionary step in the long way of digital economy strategic developments decisions started in 90s with the first pillars in e-commerce. AI national policies and strategies become the general frame for further developments of AI at the national level.

On the other hand, state at all the levels of governance (federal, regional, cities and towns) is one of the consumers of AI. AI in governance can automate routine tasks and perform critical ones and provide national cybersecurity. Moreover, many countries have already adopted national AI strategies shaping their priorities in AI development, main goals and policy frameworks.

Current level of digital technologies advancement and AI implementation both in government & business sectors raise a number of questions and challenges in the field of regulation. Thus, it is important to outline key indicators that allow to assess the level of AI use in a country with regard to ethical issues.

We suppose that the research will have practical and theoretical value for policy makers, AI developers and business and suggest “GOVERNMENT AND STATE POLICY” index to address the above challenges.

General framework for proposed Global AI ethics of index rely on the concept of life cycle (in this case AI life cycle) and includes the following subindexes interconnected with each – key groups of actors (government, business, civil society, research centers) and enabling factors relevant for all the groups of actors (AI literacy, R&D investments and ICT infrastructure development).

The structure of the research is follows – the first section is introduction, the second one is covers national priorities for AI midterm development, the third part is devoted to methodology of subindex government and state policy. The final one is dedicated to the discussion covering the most challenging issues in practical implementation of the subindex within different groups of countries depending on the degree of digital divide.

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

National priorities for AI development: main tracks and challenges

The international application of the assessment frameworks requires our understanding of the level of digitalization of the countries. In the case of the subindex Government elaboration, we focus on national policies and priorities. The general development track in digital economy with focus on AI has several main elements – national policies in digital economy development (general one or devoted to concrete sectors), e-government programmes, national AI policies, AI adoption for state agencies and AI ethics initiatives.

For the purpose of the research, we use the approach introduced in regular UN DESA report World economic situation and prospects. UN experts divide all the countries in three main groups depending on the “basic economic country conditions” (UN 2022), developed, transition economies and developing countries (Appendix 1).

Developed countries are the leaders of digital economy developments. The majority of these countries passed a long way in digitalization since the second half of the 20th century. In different combinations they enjoy all the key strategic elements in regulations of digital economy. It made possible to consistently elaborate and apply strategic documents for AI. More attention is now attracted to AI ethics with the revision of the initial AI strategies or with the adoption the documents having the ethical issues as one of the main points. So, AI ethics became one of the essential elements in Blueprint for an AI Bill of Rights introduced in USA in 2022 (The White House 2022). Some developed of the countries start the revision of the first AI strategies. For, example, United Kingdom adopted the new edition of the midterm national AI strategy in 2021 for the next 10 years (UK Office for Artificial intelligence, 2021). And AI ethics in the case of United Kingdom is indicated as one of the national priorities.

Transition economies

According to the UN classification¹, such countries as Albania, Bosnia and Herzegovina, Montenegro, Serbia, The former Yugoslav Republic of Macedonia, Armenia, Azerbaijan, Belarus, Georgia, Kazakhstan, Kyrgyzstan, Republic of Moldova, Russian Federation Tajikistan, Turkmenistan, Ukraine and Uzbekistan have a state “economies in transition”. IMF also includes² in this category other **CEE** countries (Croatia, Czech Republic, Hungary, Bulgaria, Poland, Romania, Slovak Republic, Slovenia, **Baltics** (Estonia, Latvia, Lithuania) and Cambodia, China, Laos, Vietnam.

¹ https://www.un.org/en/development/desa/policy/wesp/wesp_current/2014wesp_country_classification.pdf

² <https://www.imf.org/external/np/exr/ib/2000/110300.htm>

Despite UNCTAD claims that “the category “transition economies” is no longer used”³ we suppose that it is important to study their national priorities for AI placing them separately from developed and developing countries.

These group of countries in terms of digital transformation and AI development is not homogeneous. It includes recognized leaders in digital transformation – China, Russia, Estonia. These countries have adopted national policies focused on extensive digitalization and AI adoption – their strategies have a potential to become a framework pattern for other countries in the list.

The other subgroup of countries – Belarus, Kazakhstan, Vietnam, Ukraine, Baltic states, Czech Republic, Hungary, Bulgaria, Poland, Slovak Republic, Slovenia are relatively fast growing digital markets with increasing potential for computer services production and export. This subgroup adopts a wide range of legislative initiatives and stimulates the active work in the field, but a lack of funding and the matters corresponding with the current problems in economy developments affects the process of AI development and implementation on government level. But it worth mentioning, that AI ethics aspects are introduced in some national strategies. Thus, AI ethics is in the list of the objectives of Serbian national Strategy for the Development of AI for mid-term 2020-2025 with indication of responsible state agencies and key indicators.

Albania, Bosnia and Herzegovina, Montenegro, Serbia, The Republic of North Macedonia, Armenia, Azerbaijan, Georgia, Kyrgyzstan, Republic of Moldova, Tajikistan, Turkmenistan, and Croatia, Cambodia, Laos are going through a process of digitalization, but the number of legislative initiatives on AI regulation is smaller than in developed economies. AI ethics discussions are enlarging.

Developing countries

Developing countries is also not homogenous group. The first subgroup is represented by the regional leaders in ICT development and application (ex. Singapore, India, some of the Gulf states). These countries are elaborating the national strategies and implement AI for public services. Besides, some of them are also made the first steps in the field of AI ethics regulation. The second group involves the countries with lower level of implementation of ICT and AI in particular, being technologically dependent from the leaders in digitalization. And the last group is so called fragile states with the lowest level of ICT implementation.

³ <https://hbs.unctad.org/classifications/>

Methodology for Global AI ethics subindex Government and state policy

Subindex about AI ethics assessment at government and state policy level requires involving the broad number of professionals or a creation the special statistical unit with experts in digital transformation. It should be taken to consideration that the list of national agencies involved can't be static enlarging from year to year depending on the number of the sectors involved. In addition, each country should decide on the levels of segmentation from federal to cities.

Firstly, the authors reviewed and compared the five most common and proven methods for evaluating complex and dynamic systems, such as:

- brainstorming;
- analysis of weaknesses and strengths;
- method of charting;
- Delphi method;
- expert evaluation.

Each of the overviewed methods has its own characteristics and limitations in application.

For government and national policy subindex framework, the authors recommend using Expert valuation method, because it is the valuable compromise between high cost, long-time calculations, and investigation depth.

1) Brainstorming

For the government and national policy level this method is not realistic in view on the fact that it's highly hard to gather all the decision makers at one place and one time. In this case only creation of special units focused on AI ethics detailed analysis could be helpful.

As a rule, brainstorming is carried out within the project team with the possibility of involving a third-party expert in the work. An expert may have broad, or vice versa, highly specialized knowledge, which, in the opinion of the project team leader, is important for the implementation of the project.

The algorithm of method is rather simple and contains of several steps:

1. The participants make the most detailed list of parameters, that are relevant for the project. In this case it could be structured for concrete state agencies and their challenges in AI ethics.
2. The paraments with least realization probability are deleted from long-list by the majority of participants. In case of AI ethics some of the problems could be common for some state agencies.

Advantages of the method: the speed of obtaining the result, the ease of implementation of the method.

Disadvantages of the method: the quality of the analysis directly depends on the experience and outlook of the persons participating in the brainstorming session.

The possibility of applying the method for evaluating the ethical aspects of AI technologies implementation:

- requires the experience project team involving for implementing similar products,
- a high cost
- the complexity of involving relevant professionals.

2) Analysis of weaknesses and strengths

For the government and national policy level this method is not realistic in view on the fact that it's highly hard to push the annual research based on the analysis of weaknesses and strengths on AI ethics with large list of the officials involved.

The method is similar to the assumption analysis method, however, the project team compiles a list of potential parameters, identifying and subsequently analyzing their weaknesses / strengths.

Advantages: detailed consideration of the Index parameters.

Disadvantages:

- the long-time realization of the method;
- excessive detail of the method;
- the quality of the analysis directly depends on the experience and outlook of the professionals involved.

The possibility of applying the method for evaluating the ethical aspects of AI technologies implementation: the project team, with insufficient experience, may miss significant parameters and aspects.

3) Charting analysis

For the government and national policy level this method is not realistic in view on the fact that it's highly hard in implementation for limited number of experts who can describe the complex situation at the field of AI ethics in state agencies

The method is carried out within the project team with the possibility of inviting an external expert. The analysis takes place in three stages:

- drawing up cause-and-effect relationships,
- creating a flowchart of the processes being implemented,
- drawing up impact diagrams.

Advantages: qualitative consideration of potential risks of projects in context of AI ethics.

Disadvantages: the implementation of the charting method requires the skills of the project team to work with this method and significant time costs.

The possibility of applying the method for evaluating the ethical aspects of the use of AI technologies: the application of this skill requires specialized competencies and experience.

4) Delphi method

For the government and national policy level this method is not realistic, because it has too long-time for realization.

The Delphi method involves conducting a large anonymous survey of external and internal experts, summarizing the collected data, issuing completed questionnaires to another expert group, followed by a face-to-face results discussion, and then re-conducting an anonymous survey with summing up the final results and compiling a list of potential risks.

Advantages: high-quality study of AI ethics challenges.

Disadvantages: the method requires the long-time realization and financial resources for implementation.

The possibility of applying the method for evaluating the ethical aspects of AI technologies implementation: the method requires a lot of time and money.

5) Method of expert assessments

The method of expert assessments is similar to the Delphi method, however, involves an open survey of experts. In case of the first attempts for calculations of AI ethics index we consider it as the most applicable being less cost affective in terms of collecting the data from the agencies involved .

Advantages: a qualitative study of the identification of potential risks.

Disadvantages: it is required creation of a base of experts who are ready to participate in a large survey.

The possibility of applying the method for evaluating the ethical aspects of AI technologies implementation: the method requires a lot of time.

Calculation formula

The authors based the assessment of groups of indicators on the significance index, which is calculated by the formula:

$$r_{ij}^k = \alpha_{ij} \beta_{ij}^k, \quad (1)$$

where

r_{ij}^k - the significance of the i-indicator , assessed by the j-th respondent, in terms of the impact on the k- factor,

$i = (1 \dots N)$, where N is the number of parameters considered in the study,

$j = (1 \dots n)$, where n is the number of responses received,

$k = (1 \dots 5)$, where 1...5 are the numbers of influence groups, respectively (respectively, cost, execution time of IT project, product quality, environment, security),

α_{ij} - the weight of the significance of the indicator i, estimated by the j-th respondent,

β_{ij}^k - the value of the "effect" of the influence of the indicator on the considered stakeholder and/or the goals pursued by him.

To assess the average value of indicators, the Index of Significance of the indicator is calculated by the formula:

$$R_i^k = \frac{\sum_{j=1}^n r_{ij}^k}{n} = \frac{1}{n} \sum_{j=1}^n \alpha_{ij} \beta_{ij}^k \quad (2)$$

Proposed parameters for calculation

AI ETHICS CASES

- number of AI ethic cases
- dynamic of AI ethic cases (current/previous)
- project scope
- project cost
- number of involved participants
- level: local, national, transnational
- project compliance with the ethical requirements of UNESCO

AI ETHICS EMPLOYEERS:

- Total number of employees
- Number of developers
- Kind of activity
- Regional binding
- association membership for AI ethical aspects development
- collaboration with research centers and think tanks

FINANCIAL SUPPORT PROGRAMS FOR AI ETHICS:

- Number of programs
- level: local, national, international
- compliance with the ethical requirements of UNESCO

STATE POLICY:

- Number of programs
- Amount of budget appropriations
- Collaboration between government and business
- Number of regulatory initiatives in ethical aspects
- Compliance with the ethical requirements of UNESCO

Discussion

Global AI ethics assessment framework use multidisciplinary approach having in the heart key stakeholders along AI lifecycle. Information and data for the subindex Government could be reached from official country publications on of national policies and reports on AI developments, incl. ethical aspects. AI ethics national policies is one of the issues that was discussed and implemented in the states that are considered as one of the leaders in AI creation and implementation, USA, EU, China or Saudi Arabia. In case of OECD members, the OECD AI policy observatory project could be very helpful for further estimations and calculation as we could get detailed information on AI national policies and priorities of the reporting countries.

Expert evaluation method is considered as the most applicable at the first iterations of Global AI ethics index calculations. Later on the improvements in national statistics with focus on assessment of AI development and its challenges could support lite changes in the methodology. Advances in introduction of national AI ethics codes could support the institutional framework within the subindex “Government and National policies”. For, example the Russian universal AI ethics code in terms of institutional structure introduce AI ethics representatives, national commission and working groups. So, that could be also the ground for further elaboration of the part of national statistics dedicated to digital economy developments and AI contribution in it.

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>

AI Ethic Code. https://ethics.a-ai.ru/assets/2022/12/30/AI_Ethics_Code_en_1.pdf

Blueprint for AI Bill of Rights. The White House 2022. <https://www.whitehouse.gov/wp-content/uploads/2022/10/Blueprint-for-an-AI-Bill-of-Rights.pdf>

National AI Strategy. UK Office for Artificial intelligence, 2021. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1020402/National_AI_Strategy_-_PDF_version.pdf

Shapiro, D.L. Supplemental Joint Brainstorming: Navigating Past the Perils of Traditional Bargaining. *Negotiation Journal* 16, 409–419 (2000). <https://doi.org/10.1023/A:1026645106984>

The OECD Artificial Intelligence Policy Observatory - OECD.AI <https://oecd.ai>

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>

Strategy for the Development of Artificial Intelligence in the Republic of Serbia 2020-2025.

https://www.media.srbija.gov.rs/medsrp/dokumenti/strategy_artificial_intelligence.pdf

World Economic situation and prospects 2022. UN 2022. Available at: URL <https://www.un.org/development/desa/dpad/publication/world-economic-situation-and-prospects-2022/>

Appendix 1

Table A
Developed economies

Northern America	Europe		Major developed economies (G7)
	European Union	Other Europe	
Canada United States	EU-15 Austria ^a Belgium ^a Denmark ^a Finland ^a France ^a Germany ^a Greece ^a Ireland ^a Italy ^a Luxembourg ^a Netherlands ^a Portugal ^a Spain ^a Sweden	Iceland Norway Switzerland United Kingdom ^c	Canada France Germany Italy Japan United Kingdom United States
Developed Asia and Pacific			
Australia Japan New Zealand	EU-13 ^b Bulgaria Croatia Cyprus ^a Czechia Estonia ^a Hungary Latvia ^a Lithuania ^a Malta ^a Poland Romania Slovakia ^a Slovenia ^a		

^a Member of euro area.

^b Used in reference to the 13 countries that joined the EU since 2004.

^c The United Kingdom withdrew from the EU on 31 January 2020 and is therefore excluded from all EU aggregations.

Table B
Economies in transition

South-Eastern Europe	Commonwealth of Independent States and Georgia ^a	
Ibania Bosnia and Herzegovina Montenegro North Macedonia Serbia	Armenia Azerbaijan Belarus Georgia Kazakhstan Kyrgyzstan	Republic of Moldova Russian Federation Tajikistan Turkmenistan Ukraine Uzbekistan

Table C
Developing economies by region^a

Africa		Asia	Latin America and the Caribbean
North Africa	Southern Africa	East Asia ^b	Caribbean
Algeria	Angola	Brunei Darussalam	Bahamas
Egypt	Botswana	Cambodia	Barbados
Libya	Eswatini	China	Belize
Mauritania	Lesotho	Democratic People's	Guyana
Morocco	Malawi	Republic of Korea	Jamaica
Sudan	Mauritius	Fiji	Suriname
Tunisia	Mozambique	Hong Kong SAR ^c	Trinidad and Tobago
Central Africa	Namibia	Indonesia	Mexico and Central America
Cameroon	South Africa	Kiribati	Costa Rica
Central African Republic	Zambia	Lao People's	Cuba
Chad	Zimbabwe	Democratic Republic	Dominican Republic
Congo	West Africa	Malaysia	El Salvador
Equatorial Guinea	Benin	Mongolia	Guatemala
Gabon	Burkina Faso	Myanmar	Haiti
Sao Tome and Principe	Cabo Verde	Papua New Guinea	Honduras
East Africa	Côte d'Ivoire	Philippines	Mexico
Burundi	Gambia	Republic of Korea	Nicaragua
Comoros	Ghana	Samoa	Panama
Democratic	Guinea	Singapore	South America
	Guinea-Bissau	Solomon Islands	Argentina
	Liberia	Taiwan Province of	Bolivia (Plurinational State of)
	Mali	China Thailand	Brazil
	Niger	Timor-Leste	Chile
	Nigeria	Vanuatu	

Republic of the Congo Djibouti Eritrea Ethiopia Kenya Madagascar Rwanda Somalia South Sudan Uganda United Republic of Tanzania	Senegal Sierra Leone Togo	Viet Nam South Asia Afghanistan Bangladesh Bhutan India Iran (Islamic Republic of) Maldives Nepal Pakistan Sri Lanka Western Asia Bahrain Iraq Israel Jordan Kuwait Lebanon Oman Qatar Saudi Arabia State of Palestine Syrian Arab Republic Turkey United Arab Emirates Yemen	Colombia Ecuador Paraguay Peru Uruguay Venezuela (Bolivarian Republic of)
--	---------------------------------	--	--

^a Economies systematically monitored for the World Economic Situation and Prospects report. These analytical groupings differ from the geographical aggregations defined according to M49.

^b Throughout the report the term ‘East Asia’ is used in reference to this set of developing countries, and excludes Japan.

^c Special Administrative Region of China.

Table D
Fuel-exporting countries

Developed countries	Economies in transition	Economies in transition			
		Latin America and the Caribbean	Africa	East Asia	Western Asia
Norway	Azerbaijan Kazakhstan Russian Federation Turkmenistan	Bolivia (Plurinational State of) Colombia Ecuador Trinidad and Tobago Venezuela (Bolivarian Republic of)	Algeria Angola Cameroon Chad Congo Equatorial Guinea Gabon Ghana Libya Mozambique Nigeria	Brunei Darussalam Indonesia Mongolia Papua New Guinea South Asia Iran (Islamic Republic of)	Bahrain Iraq Kuwait Oman Qatar Saudi Arabia United Arab Emirates Yemen

Source: World Economic situation and prospects 2022. UN 2022. Available at: URL <https://www.un.org/development/desa/dpad/publication/world-economic-situation-and-prospects-2022/>



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-reviewed articles and research papers. Our enlarging network of strategic partnerships makes it possible to provide AI consulting and policy solutions both for business and government agencies.

The Centre was founded in October, 2021

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>



приоритет2030[^]
Лидерами становятся