



Ethical Issues Surrounding the Future of AI in Education

LG AI Research's Approach and Beyond

Myoung-Shin KIM

Principal Policy Officer

22 October 2024

LG AI Research in LG Group

In order to improve our AI capabilities on the group level, LG AI Research is spreading unique elements to our affiliates. This includes improving AI technology capabilities, overcoming business challenges, and conducting joint research.



EXAONE (EXpert Ai for everyONE)

EXAONE 3.0 excels in real-world AI applications, with top-tier performance benchmarks, transparent global validation, and an open-sourced 7.8B model to contribute to the AI research ecosystem

Improving Customer Service Efficiency by Real-Time Speech-to-Text Analysis

Real-time analysis of consultation content (summarizing the conversation, classifying the type and tone of the consultation, extracting data on malfunctions mentioned in the conversation).

Mar. 2023 LG Household & Health Care
Jul. 2023 LG Electronics



AI Partnership to Revolutionize Biomedical Science

LG AI Research and Jackson Laboratory join forces to use AI for breakthroughs in Alzheimer's and cancer research, aiming to revolutionize personalized medicine.

Mar. 2024 LG AI Research & Jackson Lab

LG-Qraft AI ETF Listing on the New York Stock Exchange in the United States

AI-powered Exchange-Traded Fund (ETF) selecting the top 100 large-cap stocks in the NYSE, based on an AI model that focuses on capital market predictions and the construction of investment portfolios.

Nov. 2023 LG AI Research



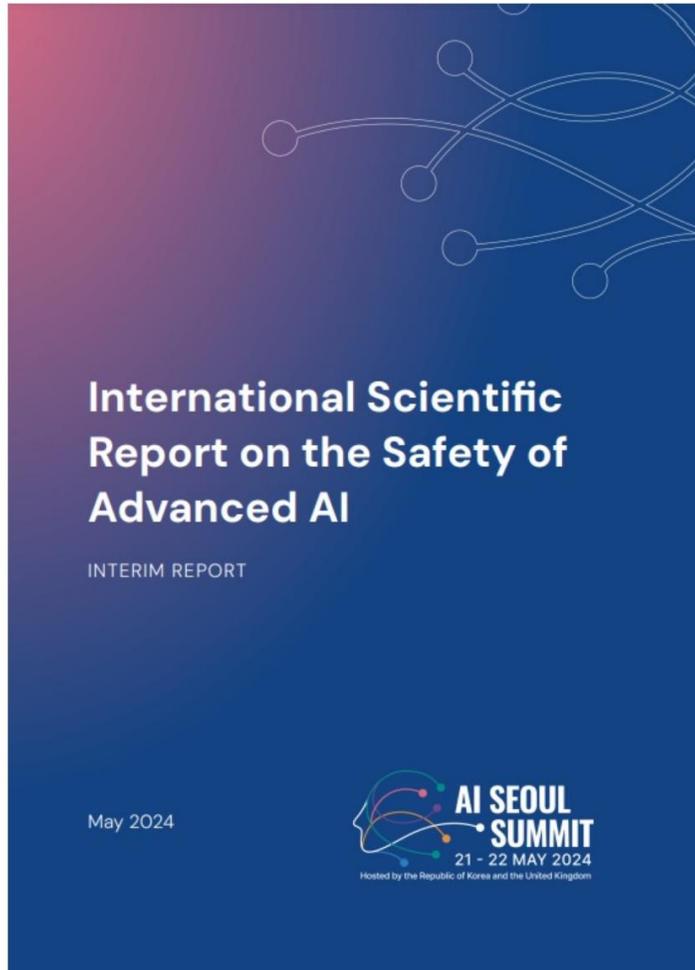
Chapter

1

**Global Perspectives
on AI Risks**

Three Key AI Risks Highlighted by International Experts

The report, developed by a global panel of 75 AI experts from 30 countries, identifies three major risks associated with AI systems. The final version will be published at the France AI Action Summit next February.



General-Purpose AI Risks

1. Malicious Use: AI can be exploited for scams, disinformation, and cyberattacks.
2. Malfunctions: AI may produce biased results and pose risks of losing control.
3. Systemic Risks: AI may disrupt jobs, widen inequalities, strain energy use, and raise privacy issues.

Cross-Cutting Risks

- Difficulty ensuring AI safety, with rapid development outpacing regulation.

Beyond AI Safety: The Rising Importance of Inclusive AI

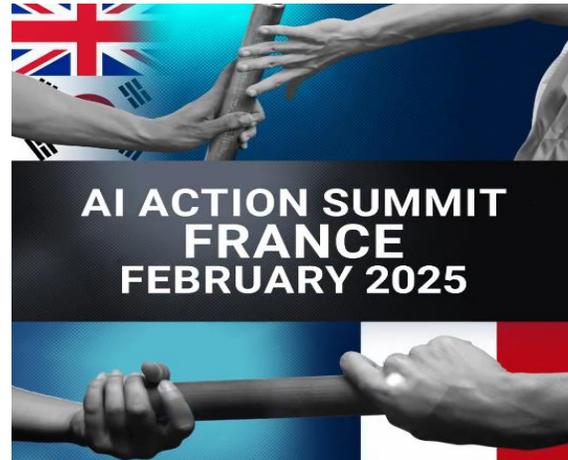
Concerns are rising over the growing gap between those benefiting from AI and those being left behind.

United Nations



- UNESCO Recommendation on the Ethics of AI (**Principles**, 2021)
- UN High-Level Advisory Body on Artificial Intelligence (2023–2024)
- UN Global Digital Compact (2024)
 - **Bridging the digital divide**, accelerating sustainable development, protecting human rights, and strengthening global governance

AI Safety Summit



- 1st Summit (UK, 2023): **Safety**
- 2nd Summit (Korea, 2024)
 - Innovation, **Safety**, **Inclusion**
- 3rd Summit (France, 2025)
 - **Public Interest**, Future of Work, Innovation & Culture, **Trust**, Global AI Governance

Big Tech



- Announcement of AI Ethics Principles (since 2018)
- Frontier Model Forum (Enhancing **Safety**, 2023)
- Partnership for Global **Inclusivity** on AI (2024)
 - Meta: Open-source AI accessibility, "No Language Left Behind"
 - AWS/Google/IBM/Microsoft/Nvidia/OpenAI: Expanding education
 - AWS/Anthropic/Google/Microsoft: Investment and support for developing countries

Advancing AI *for a* better Life



Responsible AI

- Securing reliability and expertise for industrial applications
- Moving beyond the declaration of AI Ethics Principles to implementation and reporting



Inclusive AI

- Open access to models and industry data
- Promoting AI literacy through education

Chapter

2

Responsible AI

Embedded Ethics

LG AI Ethics Principles

Going beyond technology, LG believes that AI will facilitate better lives for people. We respect the AI Ethics Principles to make our society healthier and more sustainable.



Humanity

- LG AI provides benefits to humanity and society.
- LG AI does not violate human rights.



Fairness

- LG AI treats all people fairly while respecting diversity.
- LG AI avoids unfair discrimination based on individual characteristics.



Safety

- LG AI operates in a safe and robust manner.
- LG AI predicts and responds to potential risks.



Accountability

- LG AI clearly defines roles and responsibilities for those involved in development and use.
- LG AI strives to fulfil the responsibility to operate appropriately in line with the intention of the human.



Transparency

- LG AI is transparent with our customers to help them better understand and trust how AI works.
- LG AI follows principles and standards to ensure transparency in our algorithms and data.

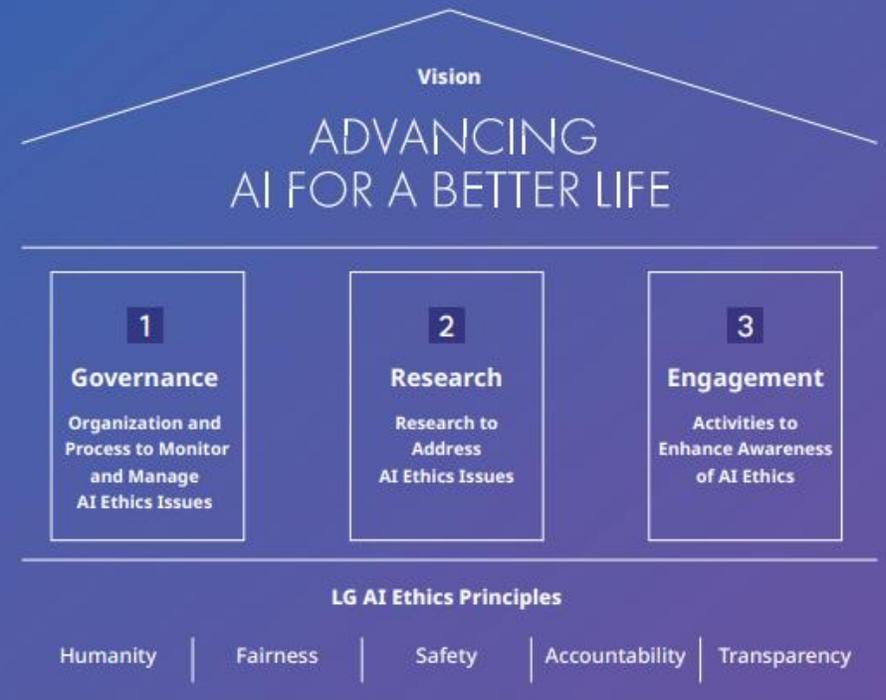
Three Strategies for Implementing LG AI Ethics Principles

:Governance, Research, and Engagement

LG AI Research is actively pursuing the following initiatives based on the 5 core values of the AI ethics principles to create responsible and trustworthy AI:

- ① Establishing and operating AI ethics governance
- ② Conducting research to address AI ethics issues
- ③ Engaging in activities to enhance awareness of AI ethics

Considering the impact of AI technology on various industrial and societal sectors, we are committed to not only improving AI performance but also to developing AI that contributes to a better life for all.



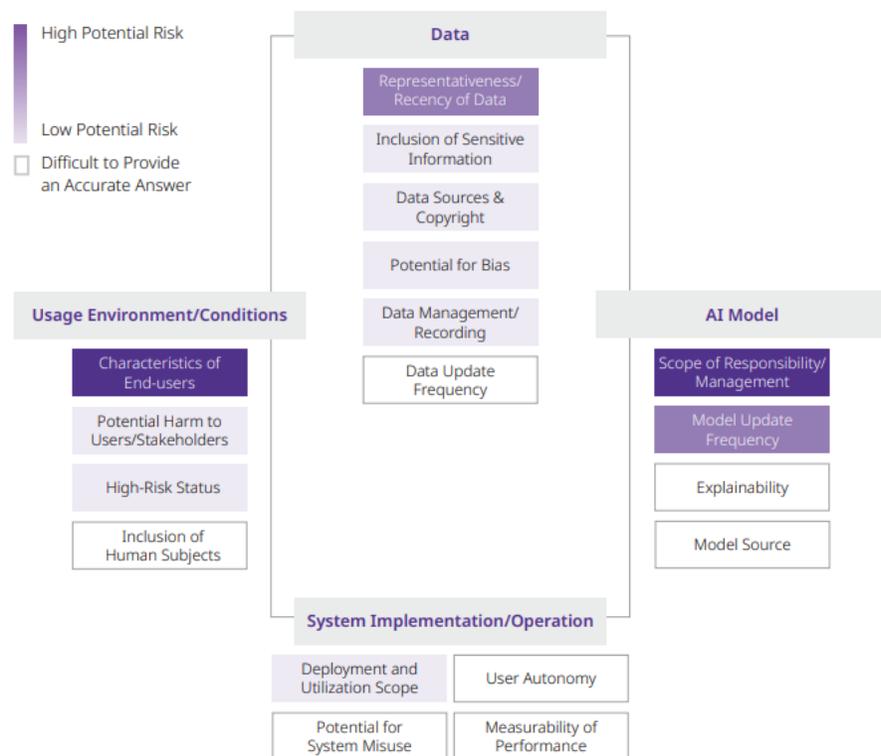
Governance_AI Ethical Impact Assessment(EIA)

Procedure designed to identify and address potential risks that may arise in the AI lifecycle

[Step 1] Analyzing Project Characteristics

The process begins with forming a Task Force (TF) comprising members specialized in AI technology, business, and ethics for each project. These members, with their varied expertise, assess the potential risks and solutions for the AI project. Using an online survey with approximately 30 questions (rated on a 1-5 scale), the TF evaluates the project's overall features.

Project Characteristic Analysis Results (Example)



[Step 2] Setting Problem-Solving Priority

In the second phase of setting problem-solving priorities, the survey results are used to discuss specific risks that could arise based on the identified characteristics of the project and potential solutions.

Problem-Solving Priority Categories (Example)



[Step 3] Verifying Results and Documentation

In the final stage, the resolution of identified risks is verified, and the entire AI Risk Management Process, including its outcomes, is documented. This documentation encompasses the project's objectives, end-users, stakeholders, training data, and model performance, as well as limitations and vulnerabilities. Such thorough documentation efforts are undertaken to ensure the transparency and accountability of AI systems.

"Before the AI Risk Management Process, I believed that the research project I was conducting did not have any particularly ethical issues. However, during the AI Risk Management process, discussions with TF members from various perspectives made me consider the impact on users, something I hadn't thought of before, and it influenced the way I conducted my research."

Yongchul Park
Language Lab

Governance_Human Rights as Compass for AI EIA

We are committed to respecting, protecting, and promoting human rights throughout the entire lifecycle of our AI systems

AI Ethical Impact Assessment

Could the completed project negatively impact users or stakeholders (e.g., physical harm, mental distress, economic damage, human rights issues)?

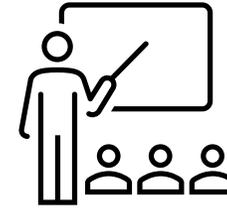
Are vulnerable or marginalized groups among the stakeholders affected by the project (e.g., children, elderly, people with disabilities, low-educated, irregular workers)?

Could this project be applied in high-risk areas?

Does the dataset include legally sensitive information, such as personal data, children's data, health and medical information?

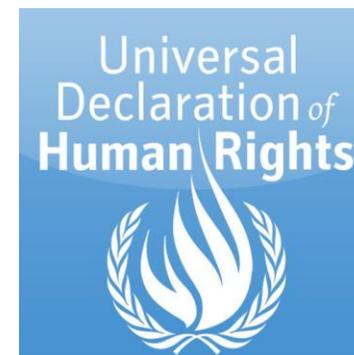
Is the dataset representative (e.g., diverse population groups, sufficient collection period, various regions)?

Can users, whether intentionally or unintentionally, use the AI system in ways that could harm people?



◀ Red teaming based on the Universal Declaration of Human Rights ▶

- Educate data workers on the Universal Declaration of Human Rights and the Sustainable Development Goals
- Provide guidelines to ensure that the data they produce upholds these values.



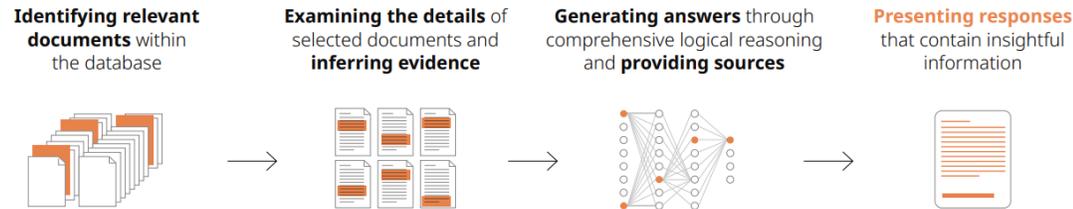
Research_AI Delivering Trustworthy Answers

Efforts to Overcome Structural Limitations of AI Models and Prevent and Address Unintended Human Rights Violations

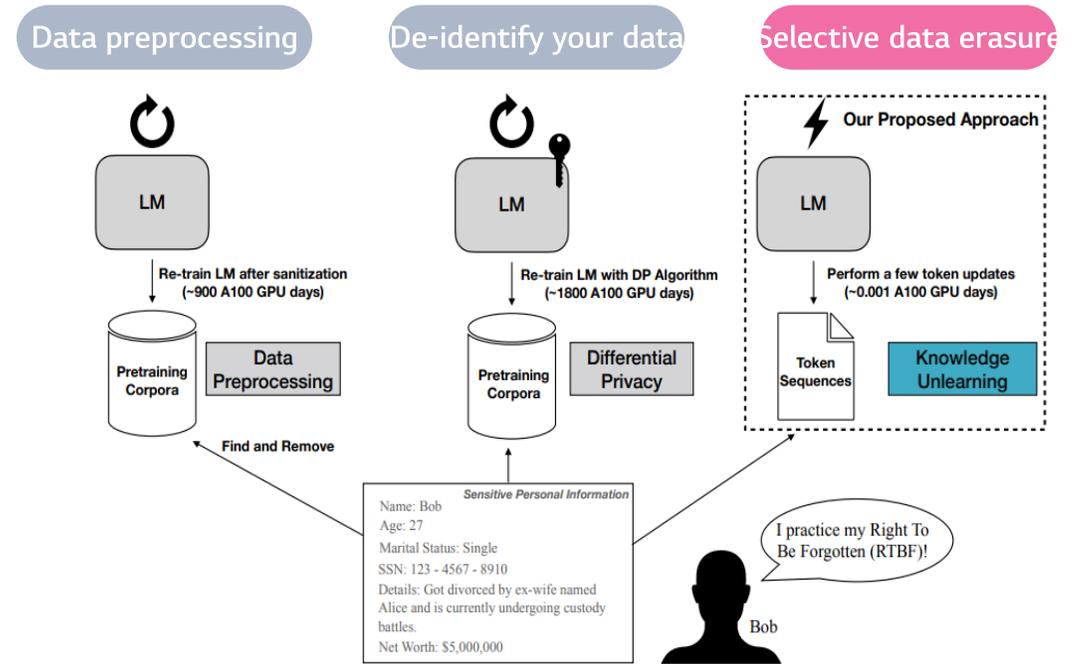
AI model generating answers based on factual information and evidence

The screenshot shows a user interface for a Language Model. At the top, there is a search bar with the text "Ask me a scientific question". Below it, a "User's question" is displayed: "How does human feedback affect language models?". The model's response is an "Evidence-based answer containing insights" that discusses the impact of human feedback on language models. To the right of the answer, there is a "Reference" section listing several research papers, with one highlighted: "The text of the specific research paper that forms the basis of the answer". On the left side, there is a "Recommended Tags" section and a "Sources" section listing various research papers, with one highlighted: "A list of research papers utilized in formulating the answer".

Answer generation process



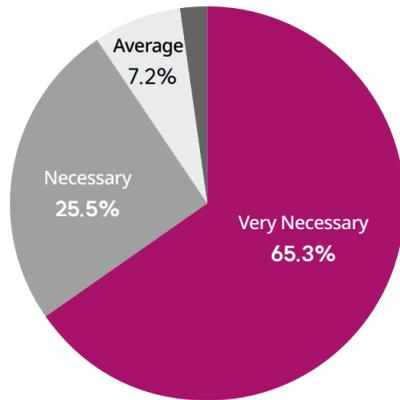
Unlearning Research on selectively removing pre-learned bias/sensitivity information from models to increase safety



Engagement_AI Ethics Awareness Survey & Seminar

Annual 'AI Ethics Awareness Survey' to assess the awareness and practice of AI ethics and to improve adherence to the institute's AI ethics principles
 Hold a Bi-weekly seminar where everyone learns from everyone

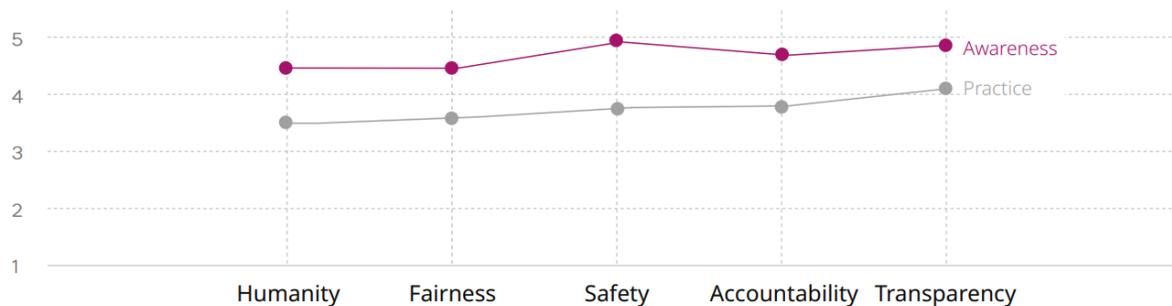
Importance of AI Ethics



Reasons for Upholding AI Ethics



Awareness-Practice Gap by Core Value of LG AI Ethics Principles



AI Ethics Seminar Season #1

Advancing AI for a better Life



'45 Minutes that Change LG AI Research'

01 Seminar Curriculum (Bi-weekly Tuesdays 12:30~13:15)

7/11(Tue)	AI Ethics in AI Companies: Focusing on Implementation	So-young Ahn (AI X Unit)
7/25(Tue)	Ontology/Epistemology of AI: Why Ask Ethics of Machines?	Woo-young Kim (Language Lab)
8/8(Tue)	The Future of Work in the AI Era: Change and Adaptation	Jun-ha Park (AI Biz Unit)
8/22(Tue)	AI Ethical Impact Assessment: A Compass for Implementing AI Ethics Principles	Myoungshin Kim (AI X Unit)
9/5(Tue)	Explainable AI: Reality, Limitations, and Possibilities	Jong-seong Jang (Vision Lab)
9/19(Tue)	Data Governance: An Essential Tool for AI Ethics	Yong-min Park (AI Biz Unit)
10/10(Tue)	(Special Lecture by External Expert) Copyright and Intellectual Property in the AI Era	Prof. Yoon-myeong Kim (Kyung Hee University)
10/24(Tue)	Core Components of AI: The Role and Ethical Considerations of Training Data	Hyeon-gu Yun (EXAONE Lab)
11/7(Tue)	Latest Research Trends in AI Ethics Benchmarks and Evaluation Indicators	Soo-hyun Cho (People Unit)
11/21(Tue)	The Ripple Effect of AI: Future and Challenges of Social Impact	Jee-soo Bang (Language Lab)
12/5(Tue)	Global AI Ethics Policy: Diverse Approaches and Limits Between Countries	Ji-min Park (People Unit)

Advancing AI for a better Life: 45 Minutes of Ethical Exploration

LG AI Accountability Report on AI Ethics

Align with the UNESCO's Recommendation on the Ethics of AI, and South Korea's National Guidelines for AI Ethics and the Digital Bill of Rights

LG AI Research

ADVANCING AI FOR A BETTER LIFE

LG ACCOUNTABILITY REPORT ON AI ETHICS 2023

UNESCO's Recommendation on the Ethics of AI (Nov. 2021)

The Recommendation on the Ethics of AI provides ethical principles for the development and use of AI. The Recommendation emphasizes that AI technology should not infringe upon human rights and fundamental freedoms. They encompass not only values and principles essential for the sound development and use of AI but also contain details on specific policy actions.

	Article	Description	Location of relevant information
Values	13-16	Respect, protection and promotion of human rights and fundamental freedoms and human dignity	8-9p
	17-18	Environment and ecosystem flourishing	20-21p
	19-21	Ensuring diversity and inclusiveness	8-9p
	22-24	Living in peaceful, just and interconnected societies	8-9p
Principles	25-26	Proportionality and Do No Harm	9p, 13-14p
	27	Safety and security	15p
	28-30	Fairness and non-discrimination	9p, 13-14p, 17-19p, 25p
	31	Sustainability	20-21p
	32-34	Right to Privacy, and Data Protection	15, 17p
	35-36	Human oversight and determination	12-14p
	37-41	Transparency and explainability	14p, 16-17p
	42-43	Responsibility and accountability	12-15p
	44-45	Awareness and literacy	22-25p
	46-47	Multi-stakeholder and adaptive governance and collaboration	26-27p
Policy Action	50-53	Ethical impact assessment	13-14p
	54-70	Ethical governance and stewardship	12-14p
	71-77	Data policy	15p, 17-18p
	78-83	Development and international cooperation	27p
	84-86	Environment and ecosystems	20-21p
	87-93	Gender	13-14p, 18-19p
	94-100	Culture	6p, 15p, 18p
	101-111	Education and research	24-25p
	112-115	Communication and information	25p
	116-120	Economy and labour	25p
121-130	Health and social well-being	20-21p	

South Korea's National Guidelines for AI Ethics (Dec. 2020)

The National Guidelines for AI Ethics are standards that all members of society, including governments, public institutions, companies, and users, should follow in all stages of development and utilization to realize ethical AI. The Guidelines are composed of 3 basic principles and 10 key requirements.

	Article	Description	Location of relevant information
Basic Principles	1	Respect for Human Dignity	9p, 12-14p
	2	Common Good of Society	9p, 25p, 27p
	3	Fitness for Purpose	9p, 20-25p
Key Requirements	1	Human Rights	8-9p, 12-15p, 18-19p
	2	Protection of Privacy	13-14p, 17-19p
	3	Respect for Diversity	9p, 13-14p, 18-19p, 25p
	4	Prevention of Harm	9p, 12-15p, 18-19p
Key Requirements	5	Public Good	16-17p, 24-25p, 27p
	6	Solidarity	25p, 27p
	7	Data Management	15p, 18-19p
	8	Accountability	9p, 12-17p
	9	Safety	9p, 13-14p
	10	Transparency	9p, 13-14p, 16-17p

South Korea's Digital Bill of Rights (Sep. 2023)

The Digital Bill of Rights sets out standards and principles for ensuring individual freedoms and rights in the digital age, and for achieving a digitally prosperous society in which digital innovation is pursued and its benefits are enjoyed justly and fairly by all. The Digital Bill of Rights consists of 6 chapters and 28 articles.

	Article	Description	Location of relevant information
Fundamental Principles	1	Guarantee of Freedom and Rights	9p
	2	Fair Access and Equal Opportunities	25p
	3	Security and Trust Assurance	9p, 12-14p
	4	Fostering of Digital Innovation	5-8p
	5	Advancement of Human Well-being	20-21p, 27p
Guarantee of Freedom and Rights in the Digital Environment	6	Guarantee of Digital Access	-
	7	Freedom of Digital Expression	-
	8	Respect for Digital Diversity	9p, 18-19p
	9	Access and Control of Personal Information	-
Fair Access and Equal Opportunities to Digital	10	Demand for Non-Digital Alternatives	-
	11	Guarantee of Digital Work and Rest	-
	12	Promotion of Fair Competition	-
	13	Protection of Digital Assets	-
Fair Access and Equal Opportunities to Digital	14	Enhancement of Digital Literacy	25p
	15	Guarantee of Data Access	26p
	16	Enhancement of Social Safety Nets to Digital	25p
	17	Ethical Development and Use of Digital Technology	2-30p
	18	Response to Digital Risks	12-14p
	19	Protection of Digital Privacy	17p
	20	Fostering of Sound Digital Environment	18-19p
	21	Protection of Children and Youths	13-14p
	22	Freedom of Digital Innovation Activities	8p
	23	Improvement of Digital Regulation	26p
	24	Support for Digital Innovation	25-26p
	25	Conflict Resolution in Digital Transition	26p
	26	Sustainable Digital Society	20-21p
27	Global Reduction of Digital Divide	27p	
28	Cooperation for Global Digital Norms	27p	

Chapter

3

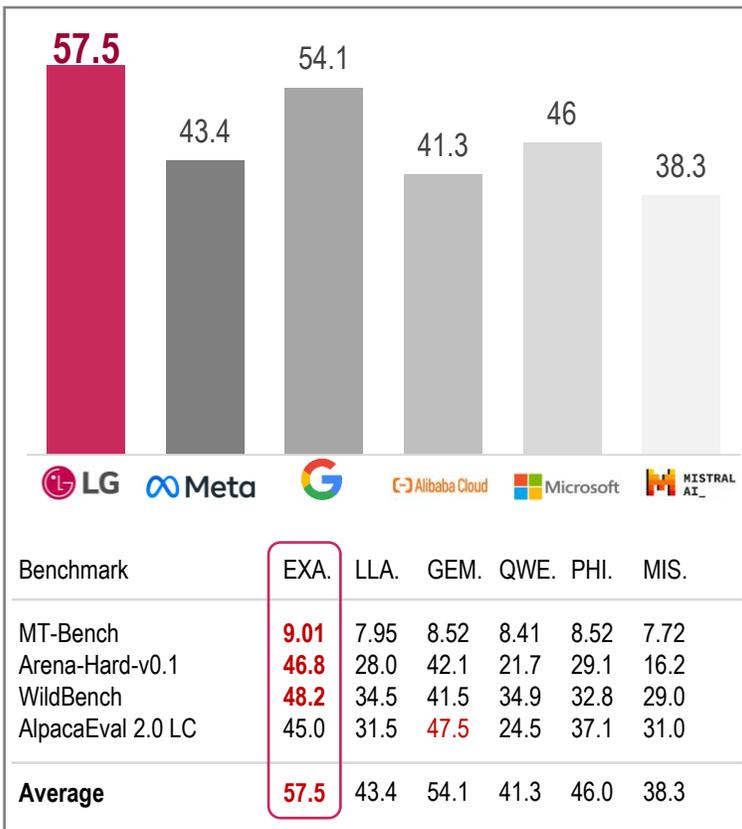
Inclusive AI

No one left behind

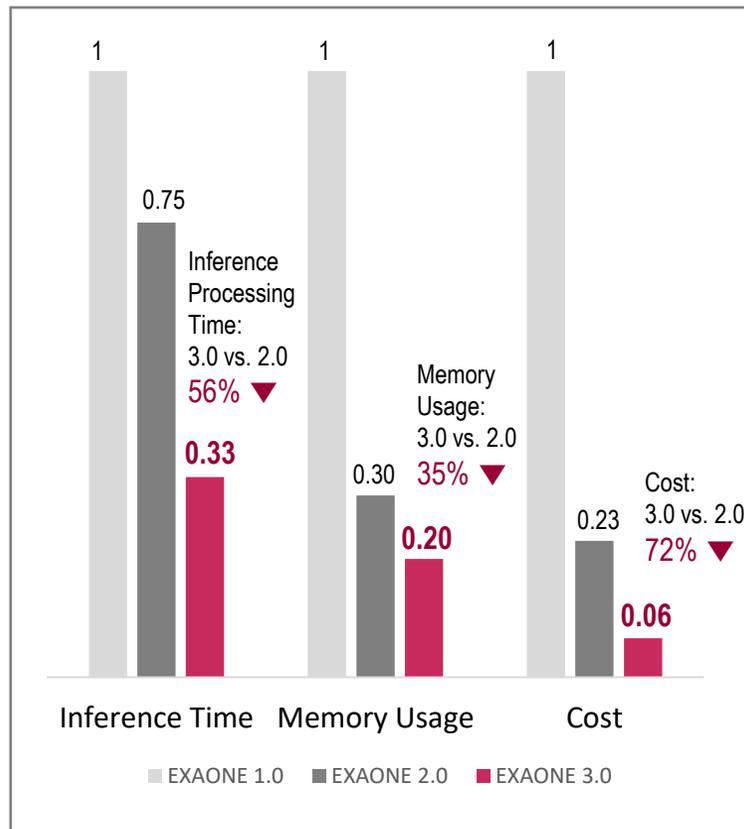
Releasing the EXAONE 3.0 language model as open source

We released the EXAONE 3.0 language model as open source to advance global AI research and accessibility.

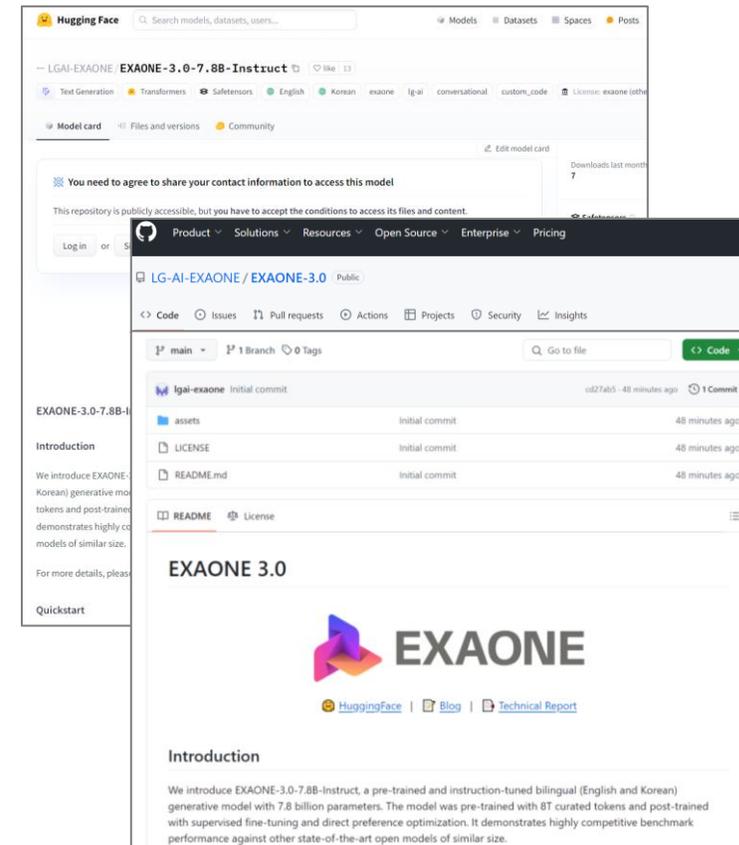
High Performance in Real-world



Cost Competitiveness



HuggingFace / Github



AI Literacy and Education Activities

Tailored AI education programs for students of all ages (over 30,000 individuals annually) for free to enhance public awareness of AI and address educational inequality

LG Discovery Lab
(AI Education for Teenagers)



LG Aimers
(AI Expert Training for Youth, 19-29)



LG AI Academy
(Specialized AI Training for LG Employees)

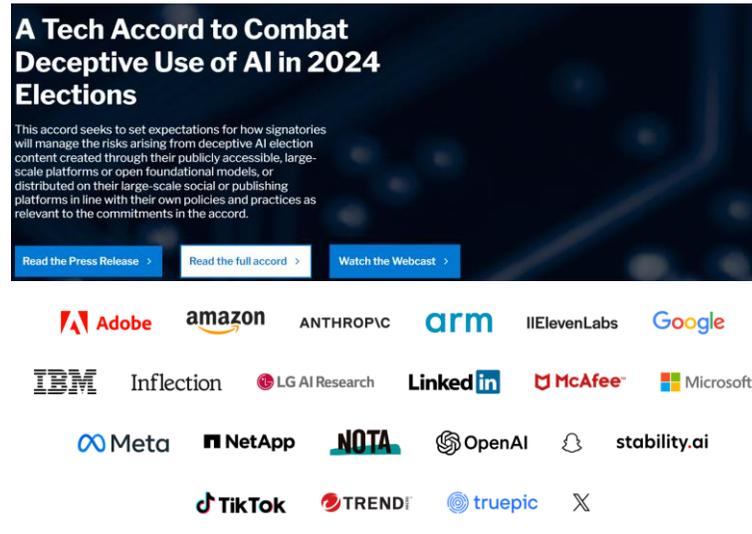


Global Partnership for building a more Ethical AI

LG AI Research is the only South Korean enterprise that has formed an AI ethics partnership with UNESCO and signed a Tech Accord to Combat Deceptive Use of AI in 2024 Elections



Partnership with UNESCO(23.11)



Tech Accord to Combat Deceptive Use of AI(24.2)
(Munich Security Conference)



Seoul AI Business Pledge(24.2)
(AI Seoul Summit)

LG-UNESCO Global MOOC on the Ethics of AI

Empower experts to become champions and advocates on the ethical practices of AI

Those that harness AI's potential while adhering unwaveringly to ethical principles like fairness, transparency, and human rights

Mind the AI Divide

Shaping a Global Perspective on the Future of Work



(Source) <https://youtu.be/AiK0iYZuNS0?si=yPhjfl4T-5twOQ3g>

- **Target:** AI practitioners, researchers, policymakers, and advocates
- **Contents:** Provides knowledge and tools to navigate the ethical landscape of AI, covering crucial topics such as bias mitigation, privacy safeguards, transparency, and human oversight
- **Schedule**
 - First half of 2025: Collect and solicit outstanding AI ethics cases globally
 - Throughout 2025: Develop MOOC content
 - Early 2026: Release globally via UNESCO platform

Chapter

4

**AI Governance
& Education**

Adaptive Governance_What to change and what to preserve?

Rather than trying to create a fully detailed AI governance framework from the beginning, we should adopt an adaptive approach —making decisions first in areas with social consensus and then updating governance as technology advances and societal perceptions change.



Source: <https://quotefancy.com/quote/1467944/Alfred-North-Whitehead-The-art-of-progress-is-to-preserve-order-amid-change-and-to>

AI's Impact Across the Life Cycle: Tailored Educational Approaches

We need to analyze the impact of AI on individuals at each stage of the life cycle and design and implement tailored educational policies that align with these insights



Source: https://stock.adobe.com/kr/search?k=human+life+cycles&asset_id=771431249

How can we build an AI Governance that gains universal trust?

Multi-sectoral Cooperation

- **Industries:** develop reliable AI models that can overcome technical challenges like hallucinations and establish robust AI risk management governance
- **Governments:** create new governance through social consensus among diverse stakeholders in the face of new conflicts arising from the proliferation of AI and ensure no one is left behind by AI's benefit (Education/Training, Reduced Working Hours, Income Redistribution)
- **International Organizations:** identify and standardize best practices in AI ethics globally and disseminate them, while also strengthening the capacities of developing countries
- **Individuals:** realize that they are not only consumers, but also data producers, content creators, and influencers, and rethink their role in AI ethics (AI ethics must be at the core of AI education)

Personal Identity in the Age of AI

Consumers



Data Producers



Content Creators



Influencers

The background features several decorative wavy lines. At the top, there are cyan lines. At the bottom, there are purple lines. On the right side, there are pink lines that form a large, abstract shape. The central text is white and stands out against the dark blue background.

Advancing AI *for a* better Life