SESSION 02 | AI FOR PUBLIC USE

## Presentation 02

# Prof. Ahreum Hong

Head of AI Technology Management Course,
Professor, Kyung Hee University

# Use of AI in Public R&D Evaluation

## Bridging the Maturity-Expectation Gap: Generative AI in Strategic Decision-Making for Public R&D

Professor. Ahreum Hong
Graduate School of Technology Management
Kyung Hee University

## Ah Reum Hong

**Professor, Graduate School of Techno Management, Kyung Hee University**
**Head of Global Management & AI Technology Management**

Technology management, technology policy, convergence industry policy

Ph.D. in Economics, Department of Technology Management and Economic Policy, Seoul National University

- Planning Committee Member, Future Generation Committee, Science and Technology Management Economics Association,
- Vice President, Future Generation Committee Member, Ministry of Science and ICT Preliminary Feasibility Assessment
- Expert Member, Korea Meteorological Administration Meteorological and Climate Human Resources Development Institute, 4th Industrial Revolution
- Advisory Member, Ministry of Trade, Industry and Energy Industrial Policy Research Institute
- Advisory Member, Seoul Metropolitan Police Agency Cyber Safety Advisory Member,
- Kyung Hee University Future Science Cluster Planning
- Member, Kyung Hee University Future Convergence R&D
- Expert Member, Seoul National University International IT Policy Program
- Research Professor Post-Doc, University of Southern California

# The State of Gen AI

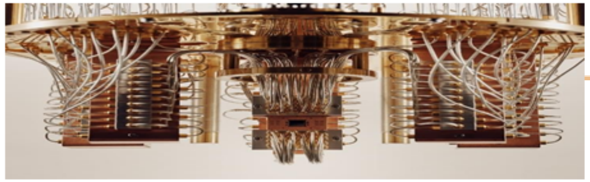
# New GPTs

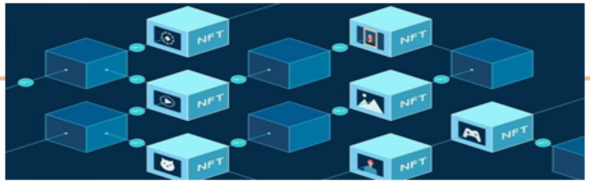
## General Purpose Technology

Quantum Computing

Block chain

AI and Machine learning







출처: https://time.com/6249784/quantum-computing-revolution/

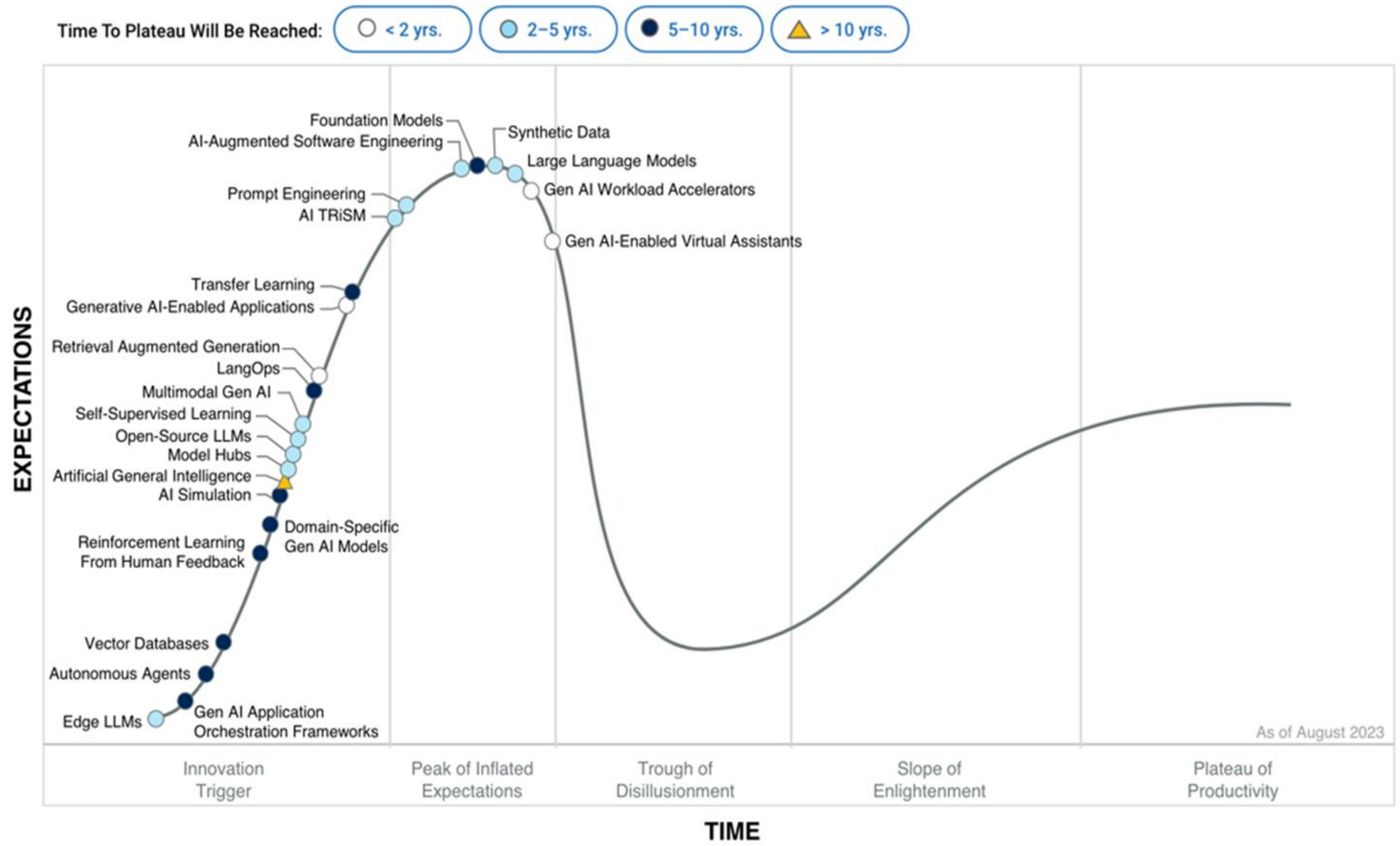
출처: https://online.stanford.edu/how-does-blockchain-work

출처: https://towardsdatascience.com/cousins-of-artificial-intelligence-dda4edc27b55

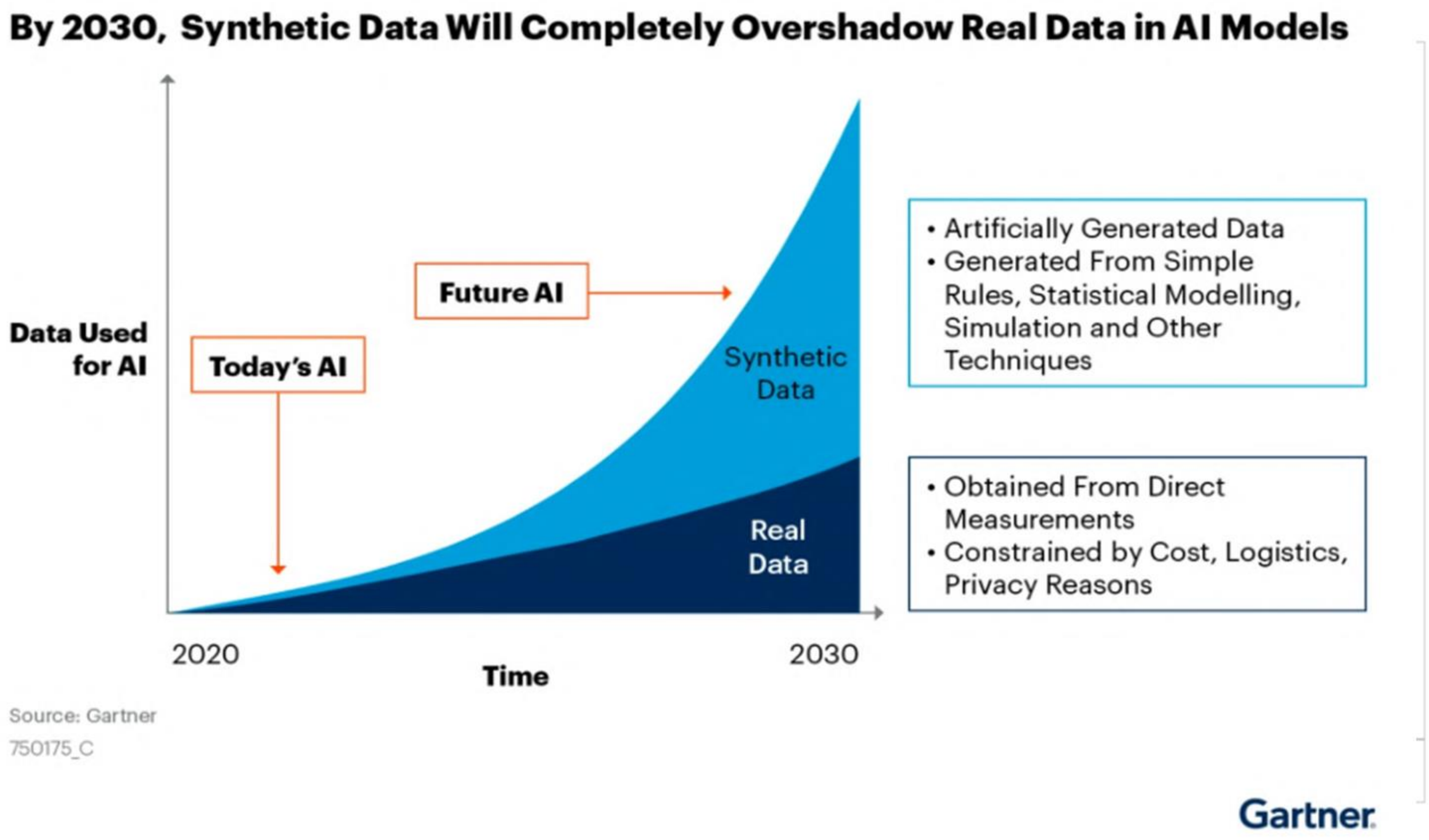
[August, 2024] California's Senate Bill (SB) 1047, which **mandates AI developers using significant computing resources to implement a Safety and Security Plan (SSP) to prevent critical harm, has garnered attention from the EU as it considers its own AI regulations.** Prominent AI experts, like Yoshua Bengio, suggest that the EU could draw inspiration from this bill. If passed, SB 1047 could align EU and US regulatory approaches, potentially reducing compliance costs for companies operating in both regions. However, major tech firms like Meta, OpenAI, and Google oppose the bill, fearing it could stifle innovation.
**Additionally, the appointment of chairs and vice-chairs for the Code of Practices for General-Purpose AI (GPAI) by the AI Office in the EU is seen as crucial for the effective implementation of the AI Act.** These leaders will play a key role in ensuring the credibility and balance of the process, with concerns that companies might only meet minimum requirements. The selection of independent and expert leaders is emphasized to maintain the integrity of the AI Act's co-regulatory approach.

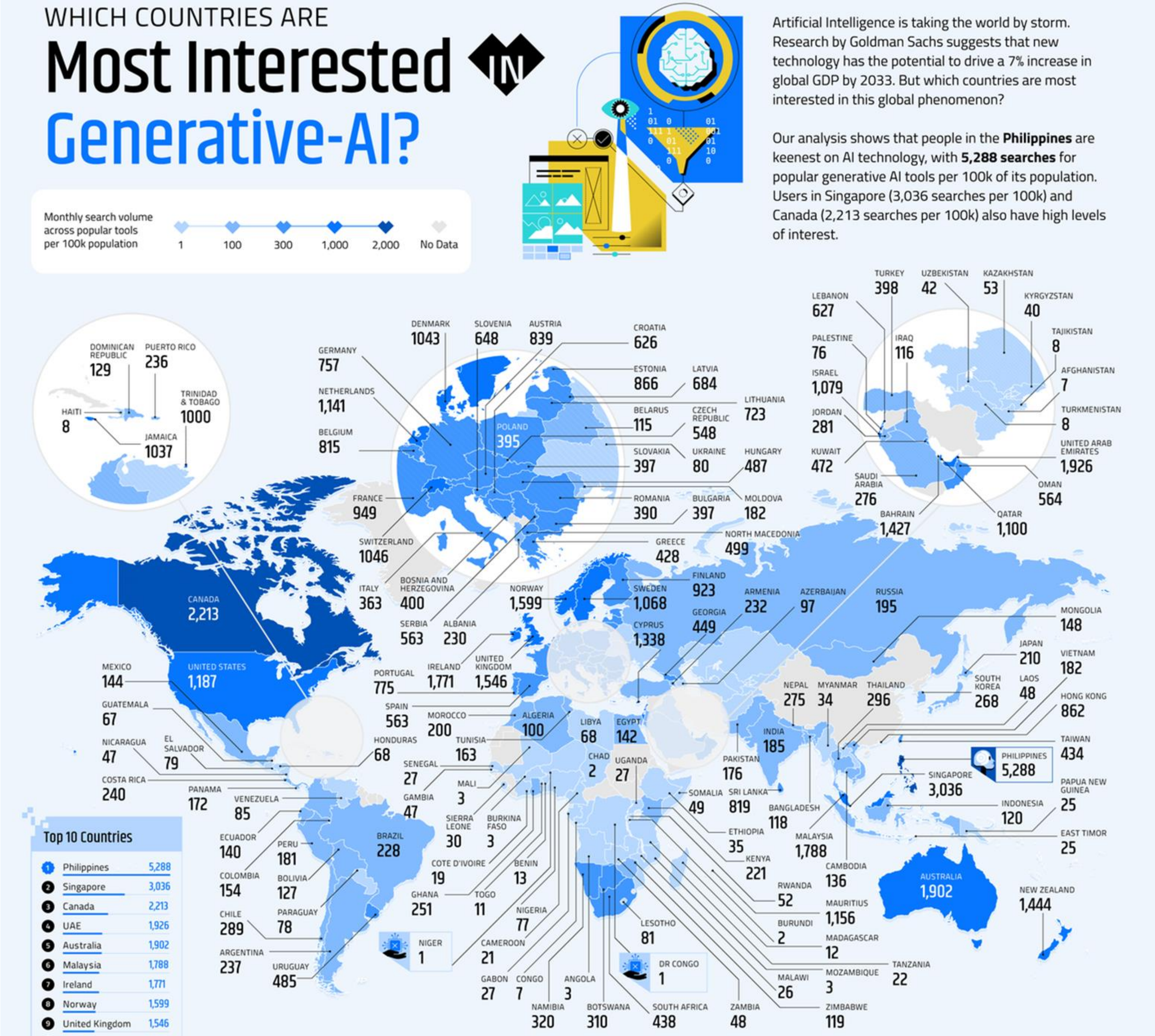
**출처:** Steen, J., Klein, G., & Potts, J. (2022). 21st-century general-purpose technologies and the future of project management. *Project Management Journal*, *53*(5), 435-437.

# Hype Cycle for Generative AI



# Hype Cycle for Generative AI




WHICH COUNTRIES ARE
Most Interested IN
Generative-AI?
Artificial Intelligence is taking the world by storm. Research by Goldman Sachs suggests that new technology has the potential to drive a 7% increase in global GDP by 2033. But which countries are most interested in this global phenomenon?
Our analysis shows that people in the Philippines are keenest on AI technology, with 5,288 searches for popular generative AI tools per 100k of its population. Users in Singapore (3,036 searches per 100k) and Canada (2,213 searches per 100k) also have high levels of interest.
Monthly search volume across popular tools per 100k population
1
100
300
1,000
2,000
No Data
DOMINICAN REPUBLIC 129
PUERTO RICO 236
TRINIDAD & TOBAGO 1000
HAITI 8
JAMAICA 1037
GERMANY 757
DENMARK 1043
SLOVENIA 648
AUSTRIA 839
CROATIA 626
NETHERLANDS 1,141
ESTONIA 866
LATVIA 684
LITHUANIA 723
BELGIUM 815
POLAND 395
BELARUS 115
CZECH REPUBLIC 548
SLOVAKIA 397
UKRAINE 80
HUNGARY 487
FRANCE 949
ROMANIA 390
BULGARIA 397
MOLDOVA 182
SWITZERLAND 1046
GREECE 428
NORTH MACEDONIA 499
ITALY 363
BOSNIA AND HERZEGOVINA 400
NORWAY 1,599
SWEDEN 1,068
FINLAND 923
ARMENIA 232
AZERBAIJAN 97
RUSSIA 195
SERBIA 563
ALBANIA 230
CYPRUS 1,338
GEORGIA 449
PORTUGAL 775
IRELAND 1,771
UNITED KINGDOM 1,546
CANADA 2,213
UNITED STATES 1,187
MEXICO 144
GUATEMALA 67
NICARAGUA 47
EL SALVADOR 79
COSTA RICA 240
PANAMA 172
HONDURAS 68
SPAIN 563
MOROCCO 200
ALGERIA 100
LIBYA 68
EGYPT 142
TUNISIA 163
SENEGAL 27
CHAD 2
UGANDA 27
MALI 3
GAMBIA 47
VENEZUELA 85
SIERRA LEONE 30
BURKINA FASO 3
ECUADOR 140
PERU 181
BRAZIL 228
COTE D'IVOIRE 19
BENIN 13
COLOMBIA 154
BOLIVIA 127
GHANA 251
TOGO 11
NIGERIA 77
CHILE 289
PARAGUAY 78
NIGER 1
CAMEROON 21
ARGENTINA 237
URUGUAY 485
GABON 27
CONGO 7
ANGOLA 3
DR CONGO 1
NAMIBIA 320
BOTSWANA 310
SOUTH AFRICA 438
ZAMBIA 48
LESOTHO 81
TURKEY 398
UZBEKISTAN 42
KAZAKHSTAN 53
LEBANON 627
KYRGYZSTAN 40
PALESTINE 76
IRAQ 116
TAJIKISTAN 8
ISRAEL 1,079
AFGHANISTAN 7
JORDAN 281
TURKMENISTAN 8
KUWAIT 472
UNITED ARAB EMIRATES 1,926
SAUDI ARABIA 276
OMAN 564
BAHRAIN 1,427
QATAR 1,100
MONGOLIA 148
JAPAN 210
VIETNAM 182
SOUTH KOREA 268
LAOS 48
NEPAL 275
MYANMAR 34
THAILAND 296
HONG KONG 862
INDIA 185
TAIWAN 434
PAKISTAN 176
PHILIPPINES 5,288
SINGAPORE 3,036
PAPUA NEW GUINEA 25
SOMALIA 49
SRI LANKA 819
BANGLADESH 118
INDONESIA 120
ETHIOPIA 35
MALAYSIA 1,788
EAST TIMOR 25
KENYA 221
CAMBODIA 136
RWANDA 52
MAURITIUS 1,156
AUSTRALIA 1,902
NEW ZEALAND 1,444
BURUNDI 2
MADAGASCAR 12
MALAWI 26
MOZAMBIQUE 3
TANZANIA 22
ZIMBABWE 119
Top 10 Countries
1 Philippines 5,288
2 Singapore 3,036
3 Canada 2,213
4 UAE 1,926
5 Australia 1,902
6 Malaysia 1,788
7 Ireland 1,771
8 Norway 1,599
9 United Kingdom 1,546

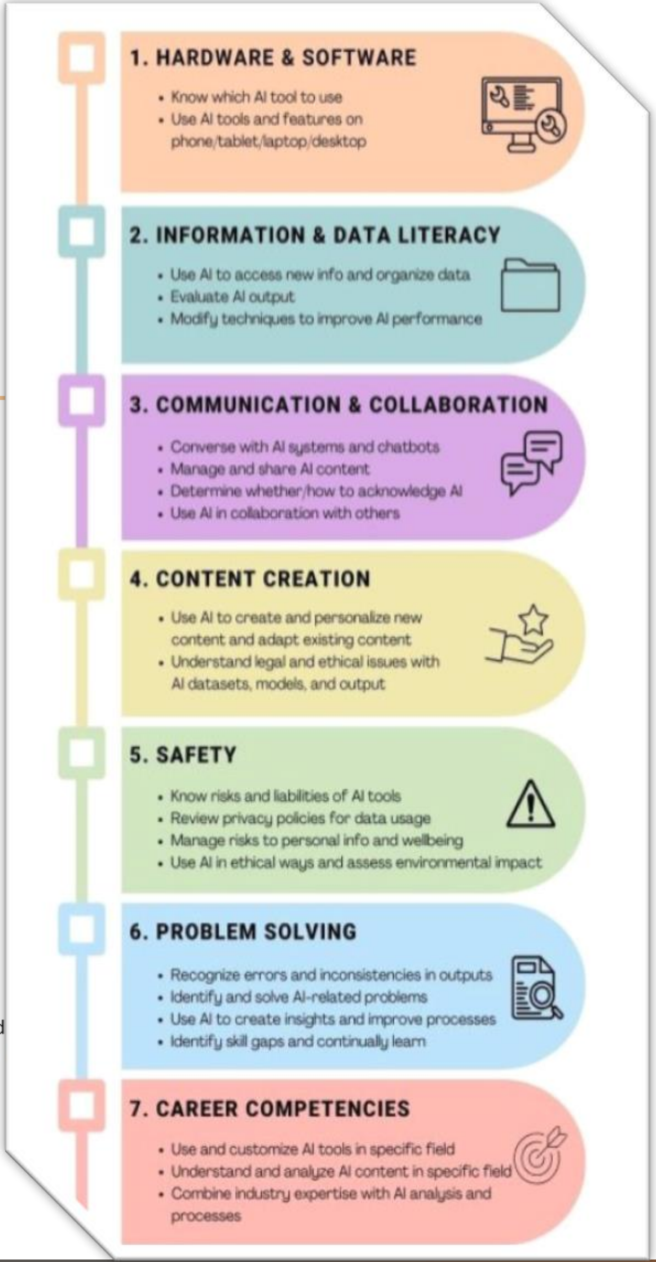
# Defining AI Literacy

- AI Literacy: Understanding basic AI concepts and applications in everyday life (UNESCO Framework)
- Importance of AI literacy:

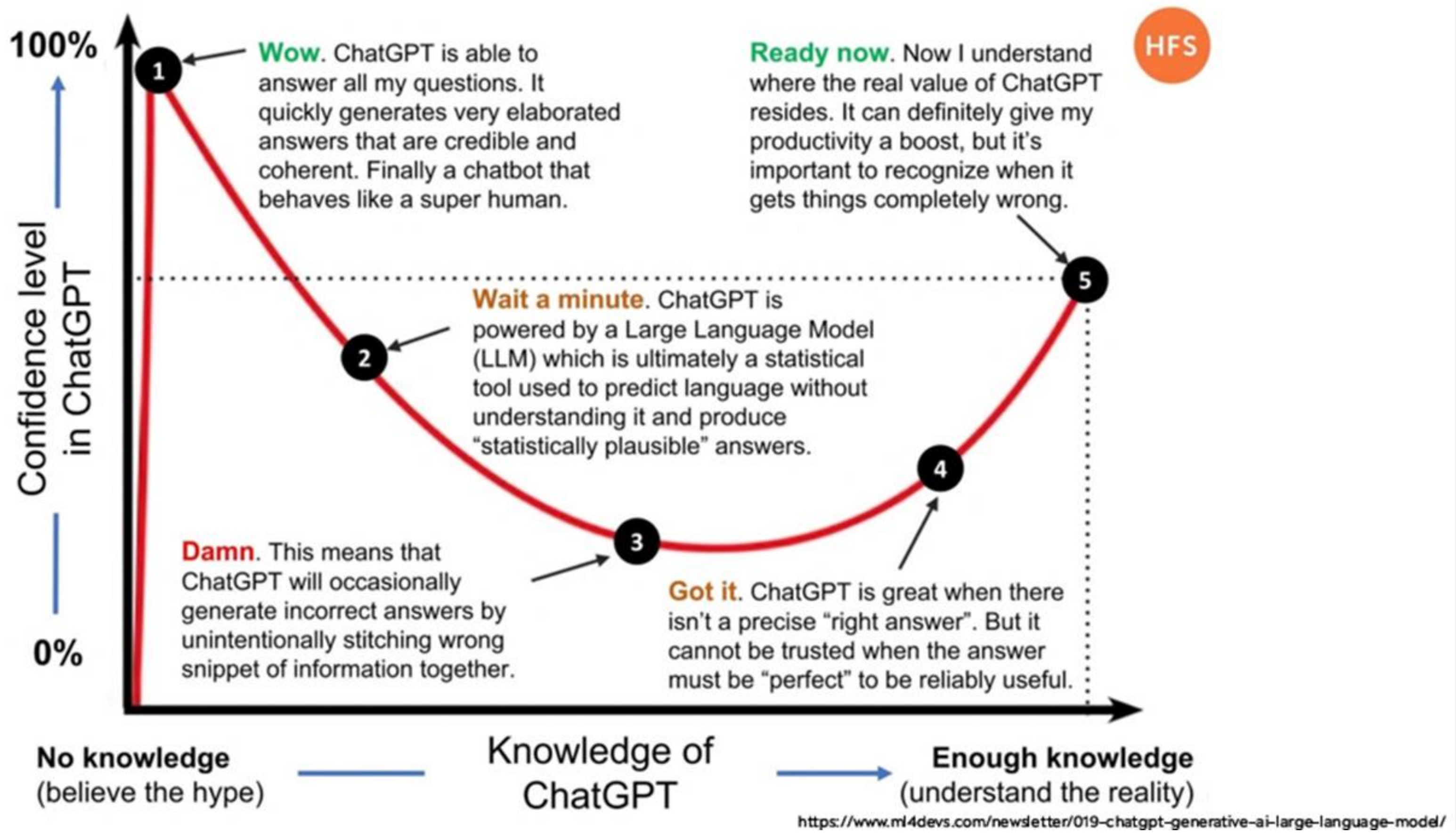
influence on job markets, & decision-making.

- Critical thinking, Data literacy, Ethical AI, Coding.
- Statistic: AI expected to contribute $15.7 trillion to the global economy by 2030 (PwC).

UNESCO. (n.d.). What you need to know about UNESCO's new AI competency frameworks for students and teachers. https://www.unesco.org/en/articles/what-you-need-know-about-unescos-new-ai-competency-frameworks-students-and-teachers



# Dunning Kruger Effect



# Survey among R&D decision-makers at leading companies shows that AI has arrived in many companies

## However there is still a long way to go

| | |
|---|---|
| Uncertain about AI capabilities & resources for transformation | 65% |
| Lack R&D AI strategy; 24% lack overall corporate AI strategy | 84% |
| Willing, but not yet ready for AI in R&D | 56% |
| See most benefits from AI in "product definition" & "prototyping" | 79% |
| Express dissatisfaction with AI implementation speed in R&D | 67% |
| Have not implemented any AI in R&D processes | 48% |
| AI in R&D is crucial now & growing in 1-2 years | All |

**Source** Roland Berger

Roland Berger

# Research Objectives

- Key Objectives:
  - Identify mature and adoptable areas for Gen AI in public R &D interim reviews.
  - Identify emerging technologies with high expectations for R&D evaluations.
  - Leverage MEG to enhance the impact of Gen AI applicati ons.

# Literature Review

**- Generative AI Integration in R&D Evaluation:** Represents a significant advancement with opportunities and challenges.

Dual nature: Enhances decision-making and efficiency but raises concerns about misuse due to AI literacy gaps (Chowdhury et al., 2024; Grimes et al., 2023).

**- Gap in Public Sector R&D Assessments:**

Existing research on AI's role in decision-making and efficiency is extensive. A gap remains in understanding AI's application in public sector R&D evaluations. .

Introduces the Metric for Evaluating Gen AI (MEG) for responsible technology adoption.

- **Advancements Peer Review with Gen AI:**

Gen AI and Large Language Models (LLMs) enhance peer review across various domains (Checco et al., 2021; Liang et al., 2024).

AI automates content analysis and reduces reviewers' workload (Chernikova et al., 2024).

Improves efficiency, objectivity, and quality of evaluations; facilitates global participation (Nauta et al., 2023; Sengar et al., 2024).

Successful AI integration in peer review requires assessing AI literacy and addressing perceptual differences.

- **Interim Reviews in R&D Assessment:**

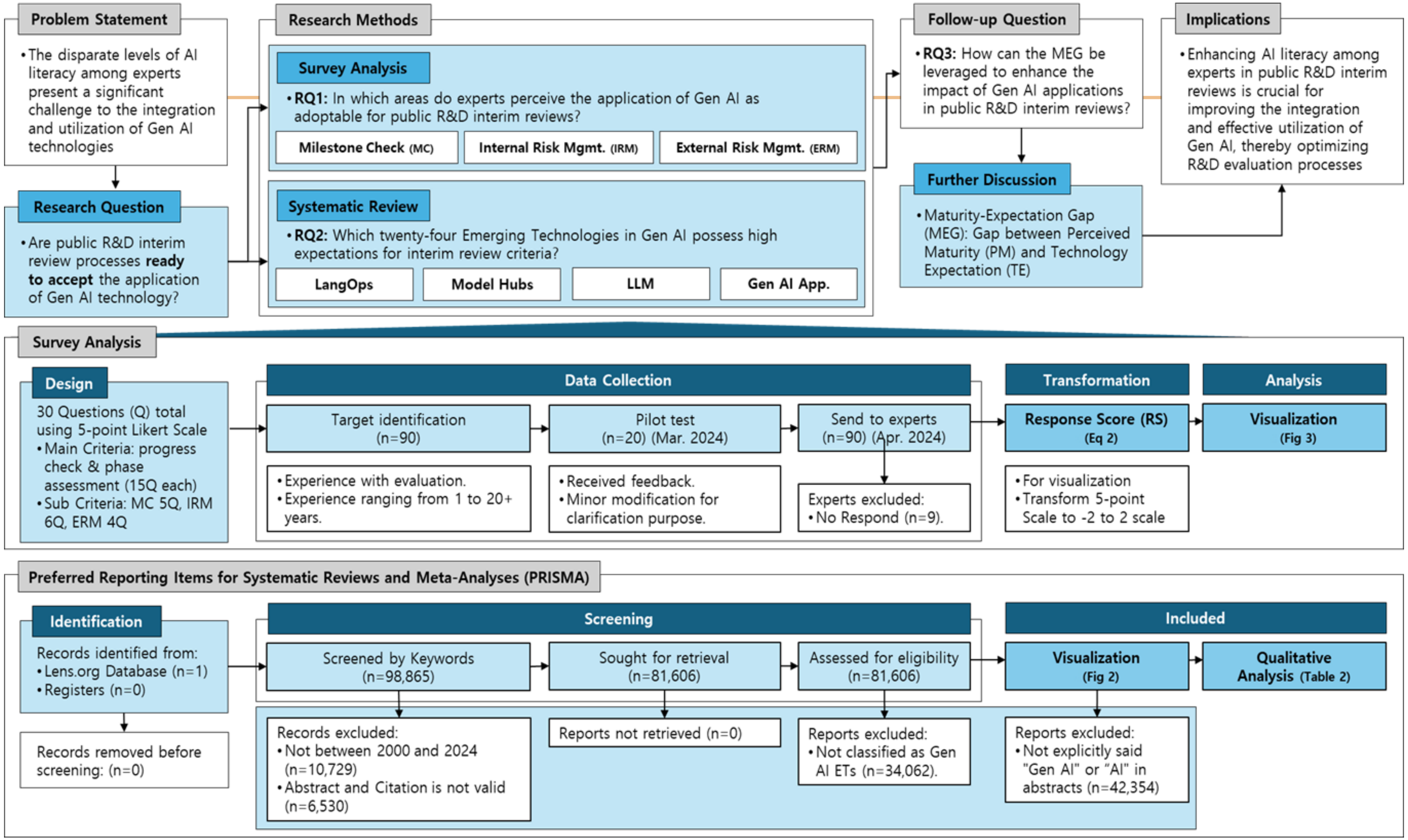
Interim reviews are critical for minimizing risks and deciding on project continuation.

Criteria typically involve milestone checks, internal risk management (IRM), and external risk management (ERM) (Aubert Bonn & Bouter, 2023; Vanino et al., 2019).

Gen AI can enhance interim review efficiency but requires improving AI literacy for successful adoption (Laureiro-Martínez & Brusoni, 2018; Luo et al., 2024).

KDI

# Research Model



# Key Concepts

- Generative AI Role and potential in public R&D assessments.
- Technology Acceptance Model Framework to analyze user perceptions and acceptance.
- Maturity Expectation Gap between perceived maturity & technological expectations.

Defined as the difference between PM and TE, MEG can contribute to responsible adoption of AI technologies. MEG exhibit low PM but high TE, then it indicates the potential advancement from the realization of future research outputs. By incorporating MEG, organizations can align AI adoption strategies with existing research on AI literacy (B. Wang et al., 2023), ensuring that users are equipped to understand AI's limitations and address explainability issues effectively. MEG can be expressed as Equation 1

$$MEG = TE - PM \qquad [Eq1]$$

# Methodology

1. Stakeholder perception analysis using surveys with R&D experts.
2. Systematic review of key emerging technologies within the Technology Acceptance Model framework.
3. Introduction and analysis of the Maturity Expectation Gap (MEG)

- PM, the RS is normalized and converted into a percentage, leading to the derivation of Equation 3. Here, $i$ represents each respondent:
- $PM_i = RS_i/(RS_{max} - RS_{min}) \times 100$ $[Eq3]$
- TE, the total paper counts from the systematic review are considered, totaling 47,544 papers. Each sub-criterion $t$ is evaluated by aggregating the paper counts and converting them into percentages, as shown in Equation 4. Here, $t$ represents each twenty-four ETs:
- $TE_t = Paper\ Counts_t/Total\ Paper\ Counts \times 100$ $[Eq4]$

KDI

## Tables

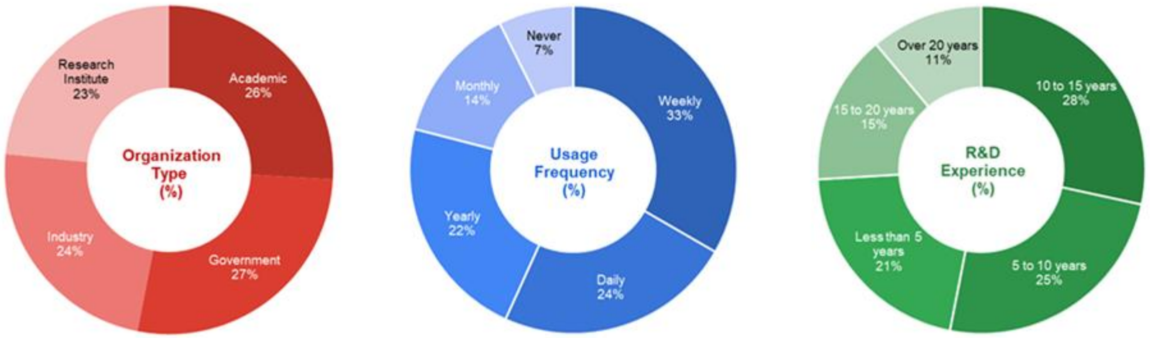
Table 1. Evaluation criteria table

| Criteria | Sub Criteria | Description | Examples of Metrics | Reference |
|---|---|---|---|---|
| Milestone Check | Effort Evaluation | Assessing researchers' effort and commitment in achieving project goals, including their adaptability when faced with setbacks. | • Number of repeated attempts<br>• Time spent on each task | (Aubert Bonn & Bouter, 2023; Vanino et al., 2019) |
| | Diligence Evaluation | Evaluating researchers' thoroughness in documenting work, managing data, and maintaining organized records. | • Completeness of research notes<br>• Adherence to document standards | (Moher et al., 2020; Vanino et al., 2019) |
| | Decision-Making | Determining if the AI system can support informed decisions on continuing, terminating, or redirecting projects based on reviews. | • Go/No-Go decision<br>• Impact of decisions on outcomes | (Balachandra, 1984; Carbonell-Foulquié et al., 2004) |
| | Progress Monitoring | Tracking project progress against planned timelines, milestones, and objectives, and identifying deviations or delays. | • Ratio of milestones achieved on time<br>• Deviation from planned timelines | (Bukoye et al., 2022; Meade & Presley, 2002) |
| | Plan Adjustment | Evaluating the need for adjustments to the research plan based on interim review findings. | • Number of plan revisions<br>• Impact of changes on outcomes | (Dvir & Lechler, 2004; Salimi & Rezaei, 2018) |
| Internal Risk Management | Resource Utilization | Assessing the efficient use of personnel, budgets, and equipment for effective resource management. | • Equipment usage rates<br>• Personnel allocation efficiency | (Jiang & Klein, 1999; Schniederjans & Santhanam, 1993) |
| | Equipment Safety | Evaluating the safe handling, maintenance, and management of research equipment and facilities. | • Number of maintenance checks<br>• Compliance with safety protocols | (Boccuni et al., 2020; Galasso et al., 2023) |
| | Financial Health | Analyzing the financial data and reports submitted by institutions to identify issues like deterioration, insolvency, or fraud | • Audit results<br>• Fraud detection metrics | (Carter & Edwards, 2001; Goldman & Peress, 2023) |
| | Compliance | Determining whether the researchers are adhering to relevant laws, regulations, standards, and ethical guidelines. | • Number of compliance violations<br>• Adherence to regulatory standards | (Cagnin et al., 2021; NIH, 1993) |
| | DE&I | Assessing the promotion of diversity, equity, and inclusion in personnel practices and the research environment. | • Diversity of research team<br>• Inclusion initiatives | (Hoisl et al., 2017; Kim & Hwang, 2022) |
| | Researcher's Satisfaction | Investigating research personnel's engagement, participation, and satisfaction. | • Satisfaction survey results<br>• Retention rates of research personnel | (M. Lee & Om, 1996; Mehrez et al., 1982) |
| External Risk Management | Market Risks | Evaluating the impact of market risks like financial market fluctuations, interest rates, and exchange rates on project viability. | • Analysis of market trends<br>• Impact of market changes on project | (Bromiley et al., 2017; Oriani & Sobrero, 2008) |
| | Security Risks | Assessing the project's vulnerability to security threats like technical issues and cyber-attacks. | • Number of security incidents<br>• Compliance with security protocols | (J. Wang et al., 2010; Y. Zhang & Yang, 2018) |
| | Environmental Risks | Evaluating the project's environmental impact, adherence to sustainability principles, and climate change-related risks. | • Environmental impact assessment<br>• Sustainability metrics | (Baker & Solak, 2011; Morrison-Saunders et al., 2024) |
| | Policy Risks | Assessing the project's alignment with government policies and the impact of policy changes on its success. | • Analysis of policy changes<br>• Alignment with government policies | (Becker, 2015; Dvir & Lechler, 2004) |

# Survey Analysis

- Participants 81 experts in public R&D.
  - Focus AreasPerceived Maturity (PM) of Gen AI in various evaluation criteria.
  - Data-driven insights on expert perceptions using Response Score (RS) methodology.
  - Visualization of results (heatmaps, bar graphs).

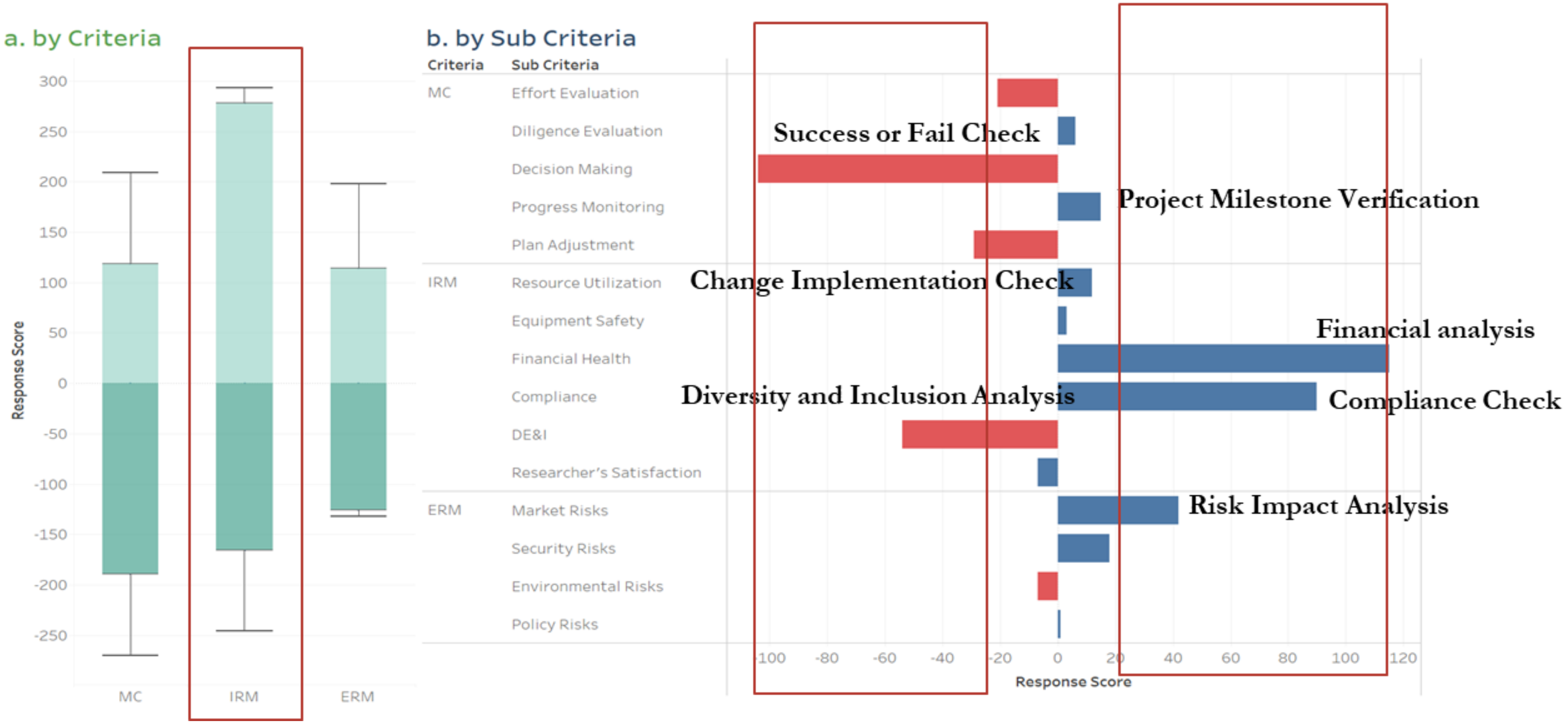
**Figure 4**. Respondents statistics



# Survey Results

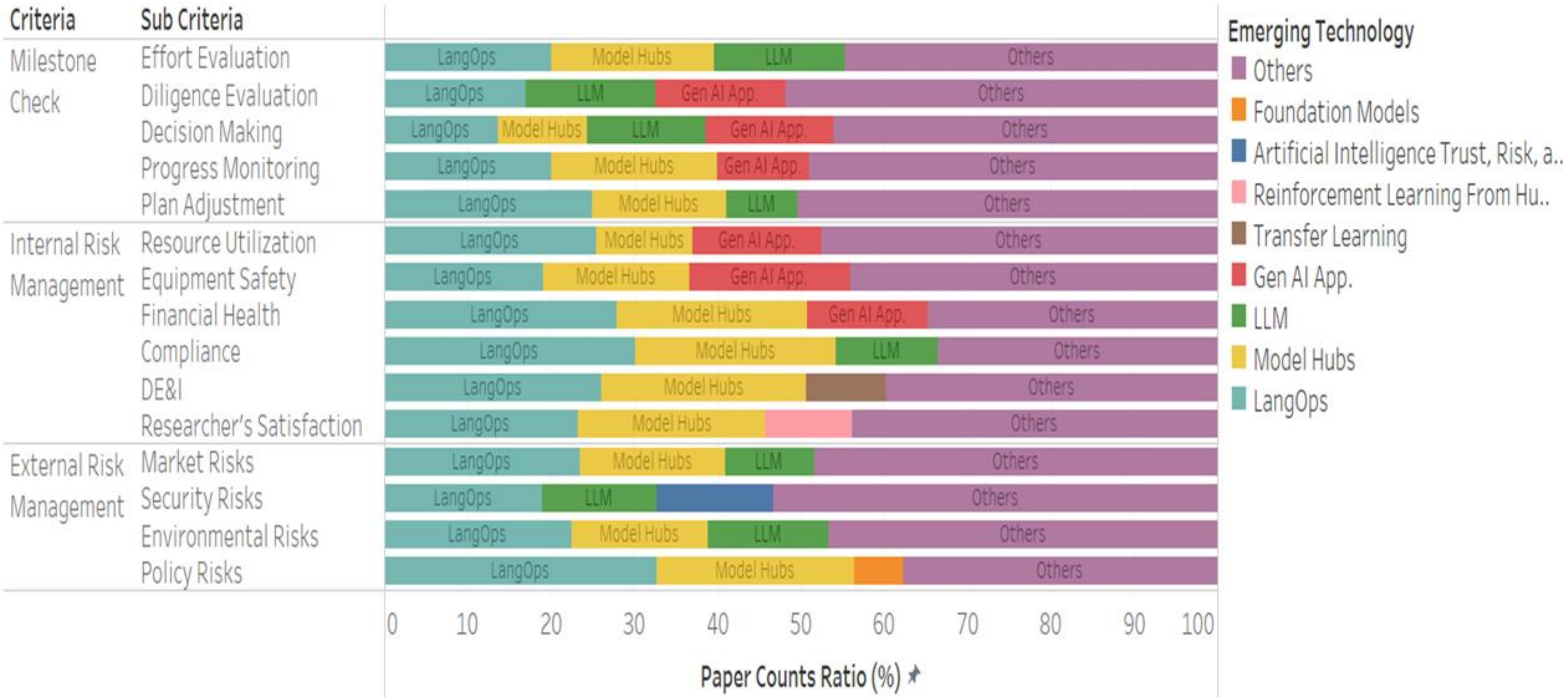
## Perceived Maturity (PM)

- Experts' awareness of maturity on Gen AI technologies for the application on public R&D interim review
- It was collected from eighty-one experts with R&D assessment experience, by using Response Score (RS) (Weighted Likert scale)
  - ✓ Experts saw potential in Financial Health and Compliance while expressed skepticism on Decision-Making and DE&I



# Survey results by category

# Systematic Review Results

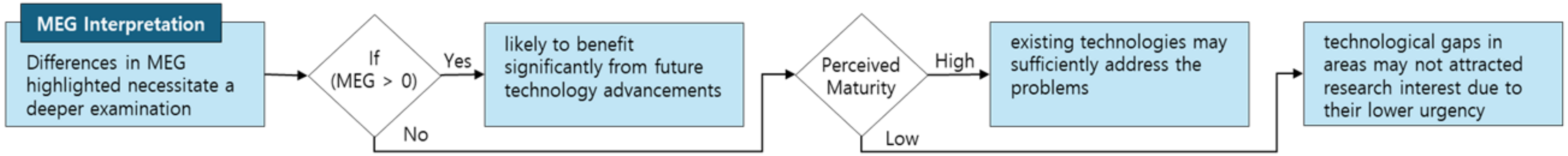



Emerging Technologies Identified: Language Operations (LangOps) Model Hubs, Large Language Models (LLMs), Gen AI-enabled Applications, Analysis of trends and potential integration areas in R&D assessments.

# PRISMA Flow and Result

| Technology | Description | Key Paper | Applications on Interim Review |
|---|---|---|---|
| **LangOps** | Technologies for automating and o ptimizing languag e operations. | Collective intelligence for deep learning: A survey of recent developments (Ha & Tang, 2022) | Enhancing the clarity of technical documents for public dissemination. Automating the translation of complex research findings into accessible language for policymakers and the public. Improving the accessibility of public information throug h advanced language processing. |
| **Model Hubs** | Platforms for stor ing and sharing v arious machine le arning models. | Detection and moderation of detrimental content on social media platforms: cu rrent status and future dire ctions (Gongane et al., 20 22) | Centralizing model repositories for public research instit utions to streamline R&D assessments. Facilitating collaboration between government agencies and academic researchers. Ensuring secure and efficient sharing of models for publi c projects, improving assessment and validation process es. |
| **LLM** | Pre-trained deep l earning models b ased on vast amo unts of data. | Machine learning techniq ues for IoT security: Curr ent research and future vis ion with generative AI an d large language models ( Alwahedi et al., 2024) | Multimodal LLM can enhance information exchange dur ing interim review meetings by addressing additional qu eries in diverse formats. RAG technology can improve the accuracy and relevanc e of information retrieval. Personalization and context-awareness capabilities can p rovide tailored evaluations and feedback. |
| **Gen AI-enabled Applications** | Applications that leverage generati ve AI to provide v arious functionalit ies. | Using AI-generated sugge stions from ChatGPT to o ptimize clinical decision s upport (Liu et al., 2023) | Ethical issues in generative AI, such as bias, privacy viol ations, and accountability for errors, are highly relevant t o public R&D evaluation processes. Gradual implementation of generative AI in public R&D reviews requires addressing copyright concerns, preventi ng discrimination, and ensuring transparency through dis cussion. |

# Findings
# Maturity Expectation Gap (MEG)


- While generative AI has demonstrated significant advancements, its perceived maturity among experts may not fully align with the heightened expectations surrounding its potential applications and transformative impact.
- MEG is an idea that measures the difference between perceived maturity (PM) and technology expectation (TE) of public R&D interim review criteria to assess whether a technology is undervalued or overvalued, and in gauging its future potential.
  - ✓ By analyzing this gap, stakeholders can make more informed decisions regarding technology adoption, resource allocation, and strategic planning in the rapidly evolving field of artificial intelligence.

# Theoretical and Practical Implications

**Theoretical:**

Expansion of TAM with the MEG concept.

**Practical:**

Recommendations for phased AI i ntegration in public R&D evaluatio ns.

Policy implications for enhancing AI adoption in complex decision-making environments.

To transform public R&D assessment using Explainable AI (XAI) and Generative AI:

1. **Set Transparent Metrics**: Define clear, measurable R&D outcomes and ensure AI processes are understandable.
2. **Build a Data Platform**: Create a secure platform for data collection and sharing, integrating XAI tools for transparency.
3. **Use XAI for Insights**: Apply XAI with Generative AI to analyze data, providing clear, actionable insights.
4. **Incorporate AI Insights**: Regularly integrate explainable AI insights to refine R&D strategies and improve decision-making.

# Conclusion and Further Study

- Highlights :
  - High MEG areas could benefit from future technological advancements.
  - Need for improved AI literacy and tailored educational programs.
  - Potential for Gen AI Significant in specific R&D domains with clear expectations but lower perceived maturity.
  - The study provides valuable insights into the potential and challenges of integrating Generative AI into public R&D assessments.
- Future Directions:
  - Navigating AI Ethics: Integrating Explainability and Transparency through Network Analysis and Topic Modeling

# Implementation Concerns

### 1. Automating Data Analysis for R&D Evaluation

- Enhancing the efficiency and accuracy of public R&D evaluations. This can significantly reduce manual work and provide real-time in sights.
- **Technological Gap:** AI lacks the ability to fully understand complex and unstructured data, especially in niche scientific fields.

### 2. Predictive Modeling for Strategic Investment

- Utilize **predictive modeling** powered by Generative AI to forecast the potential impacts of R&D projects, enabling strategic investm ent decisions based on anticipated returns and innovation potential.
- Existing predictive models still struggle with uncertainty and dynamic changes in R&D projects, especially in emerging fields where hi storical data is limited.

### 3. Enhanced Decision-Making Through Scenario Generation

- Enable **Generative AI** to create various decision-making scenarios by analyzing potential outcomes of strategic choices. To assess ris ks and benefits before making critical R&D funding decisions.
- Scenario generation requires AI to simulate complex socio-economic environments, which currently remains limited in scope. Q

KDI

# After Assessment?

*Who Invent AI ?*

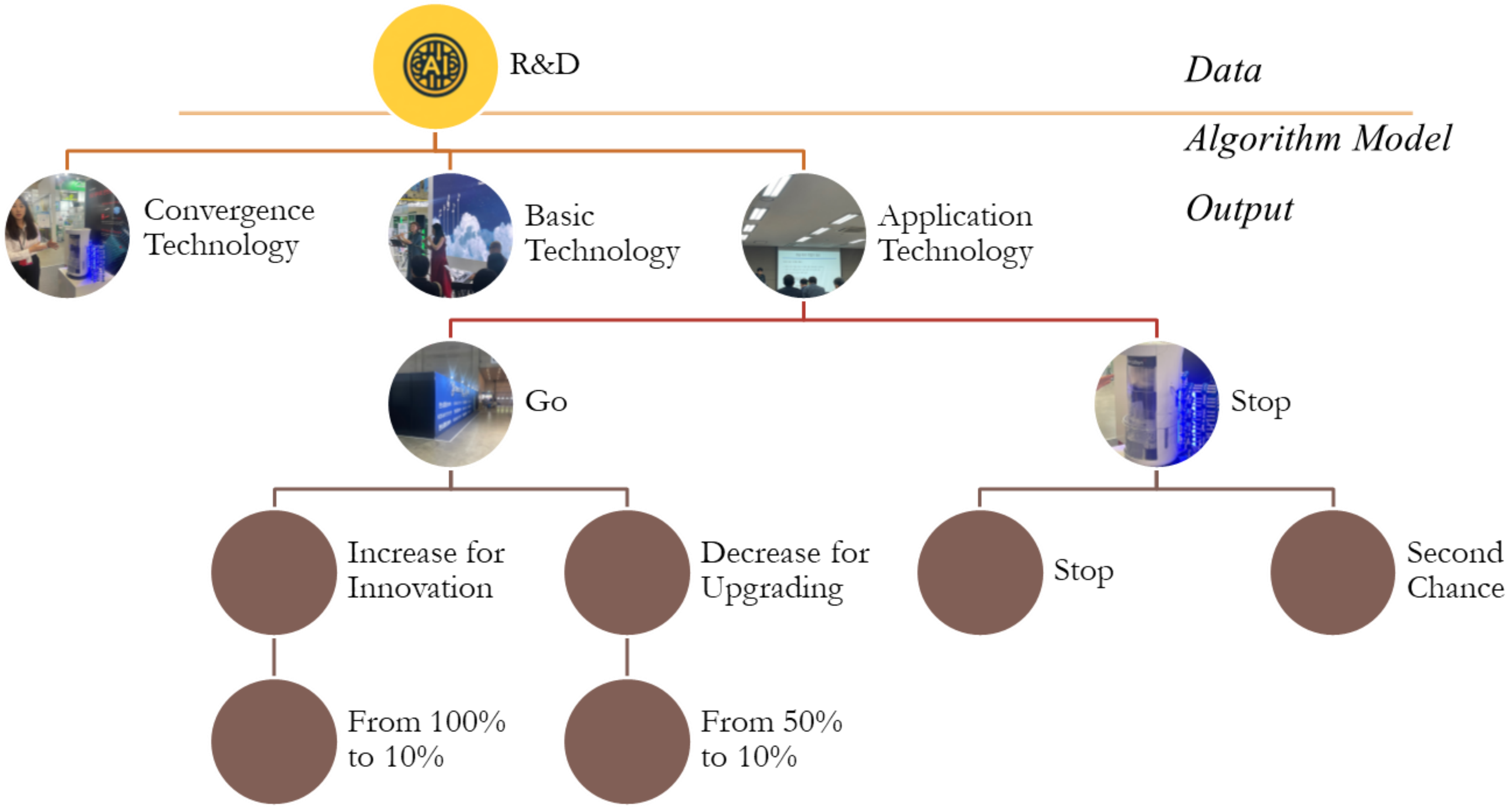
*Who use AI ?*

*How use AI ?*

*Data*

*Algorithm Model*

*Output*



# Navigating AI Ethics in Public R&D:

## Integrating Explainability and Transparency through Network Analysis and Topic Modeling

- **Identifying Research Gaps:** Through the analysis of existing literature, this research emphasizes the need to consider cultural factors and areas requiring further exploration. Understanding how AI ethics can integrate diverse cultural backgrounds and technological developments is crucial.
- **Guiding Future Research:** By synthesizing current research, this study proposes new research directions, helping researchers explore important and innovative topics in AI ethics.
- **Topic Modeling:** The study extracts key topics from AI ethics-related papers, focusing on the core issues of explainability and transparency. By applying BERTopic, BERT embeddings are used to convert text data into vectors, followed by dimensionality reduction using UMAP. Document clustering is then performed using HDBSCAN, and c-TF-IDF is employed to extract representative terms from each cluster to define topics.
- **Keyword Network Analysis:** The relationships between major keywords are visualized to concretely illustrate the interaction between explainability and transparency. Based on the representative keywords of each cluster, the topics are defined and analyzed.
- **Conclusion:** By deeply analyzing the interaction between explainability and transparency in AI ethics research, this study presents new directions to enhance the reliability and accountability of AI systems. The approach using network analysis and topic modeling provides new insights into AI ethics research.

Kim, K., Kogler, D. F., & Maliphol, S. (2024). Identifying interdisciplinary emergence in the science of science: combination of network analysis and BERTopic. Humanities and Social Sciences Communications, 11(1), 1-15.

# What's Next?

Optimization & Equilibrium using Gen AI

# Appendix