

Does AI increase or decrease transparency?

Tools and our relations with them

Yung-Hsuan Wu
Researcher, DiploFoundation

DIPLO

Geneva Internet Platform

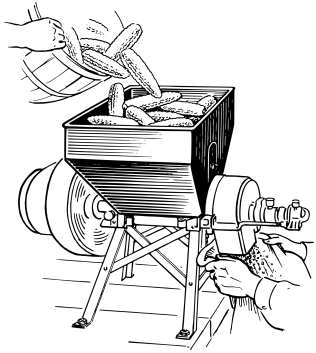


Know the talk before you walk the talk.

At Diplo, we insist on knowing what AI is before delving deep into AI governance issues.

It is crucial to our research, teaching, discussions, and work.

Workflow of creating a survey



Writing questions

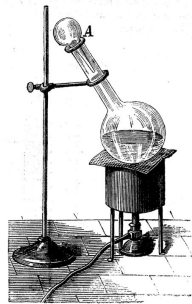
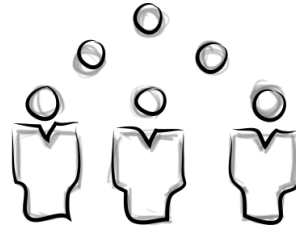
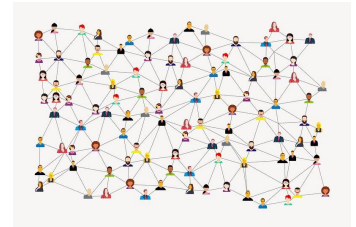


Fig. 29.

Testing survey

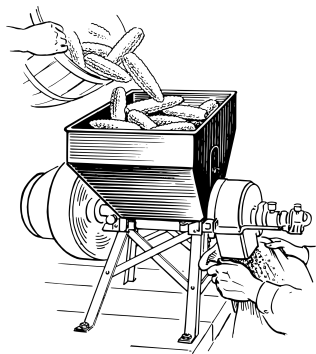


Distribute survey/
collect responses



Analyse responses

How does AI fit in this?

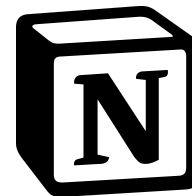


Writing questions

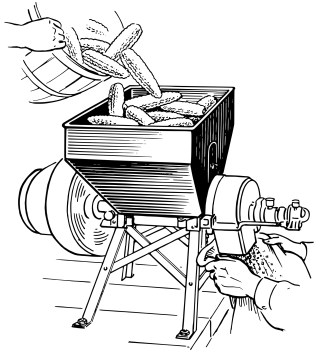
- AI-based writing tools
 - ChatGPT
 - Copy AI
 - Nichesss
 - NotionAI
 - Google Docs
 - Microsoft Teams Premium
 - ...



copy.ai



How does AI fit in this?

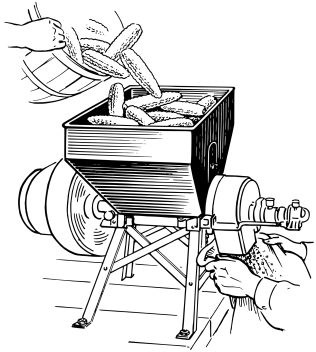


Writing questions

- AI-based writing tools
 - ChatGPT
 - Copy AI
 - Nichesss
 - NotionAI
 - Google Docs
 - Microsoft Teams Premium
- Online survey creators
 - SurveyMonkey
 - Survey Sparrow
 - Startquestion
 - ...



How does AI fit in this?

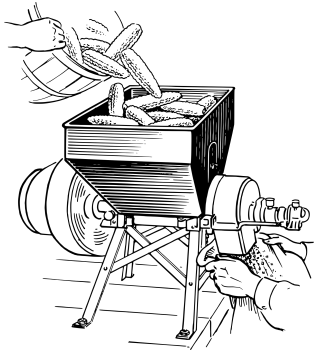


Writing questions

- AI-based writing tools
 - ChatGPT
 - Copy AI
 - Nichesss
 - NotionAI
 - Google Docs
 - Microsoft Teams Premium
- Online survey creators
 - SurveyMonkey
 - Survey Sparrow
 - Startquestion
 - ...

Use: Take prompts to generate questionnaire templates, detect and recommend question types based on question wordings, etc.

How does AI fit in this?



Writing questions

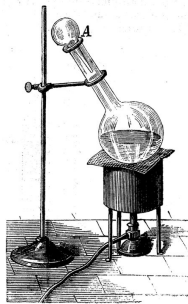


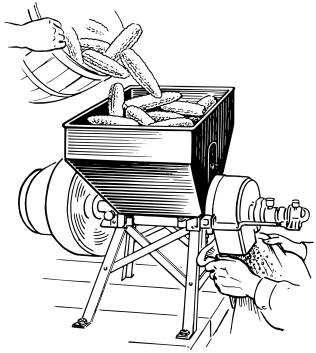
Fig. 29.

Testing survey

- Online survey creators
 - Issue detecting
 - Estimated response rate
 - ...

Big data-enabled survey experience analyses

How does AI fit in this?



Writing questions

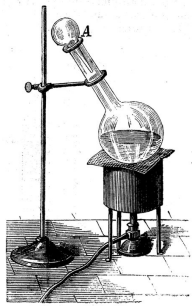
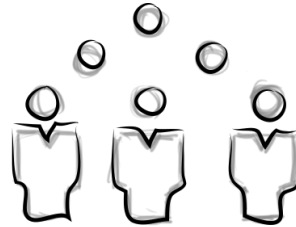


Fig. 29.

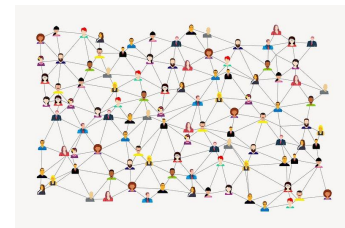
Testing survey



Distribute survey/
collect responses

How does AI fit in this?

- Response summarisation:
 - Works well with open-ended responses
 - Tone detection
 - Key word finding
- Metadata analysis:
 - How long did participants spend on each question
 - Patterns in questionnaire answering
 - Trend analyses



Analyse responses

How does AI fit in this?

Global Digital Compact

Submissions | Details of Individual Submission

Type of entity / organization

- Selection/merit based
- Academia/research institution/think tank
- Civil society or non-governmental organization
- Government
- Individual
- International, regional or multilateral organization
- Other
- Private sector/Corporation
- Technical community

Geographic scope of activity

Tout

Region

Tout

Country

Tout

Areas for which inputs submitted

- Selection/merit based
- Accountability for Discrimination/Misleading Content
- Connect all People

Welcome to the United Nations

Entity/Org type	Geographic scope	Entity Name	Contact Name	# of Core Principals	Nombre de Commentaires, pledges or actions
Academia/research institution/think tank	Global	New York University School of Professional Studies	Ananya Balasubramanian	8	8
Academia/research institution/think tank	Global	The Glenlead Centre	Ann Kristin Gjenster	8	8
Academia/research institution/think tank	Global	United Nations University, Institute in Macau	Dr. Franc Gatzweiler	8	8
Academia/research institution/think tank	Global	Minderoo Centre for Technology and Democracy, University of Cambridge	Jeremy Hughes	8	8
Academia/research institution/think tank	Global	Minderoo Centre for Tech University of Cambridge	Welcome to the United Nations		
Academia/research institution/think tank	Global	Collaboratory for Indigene WAMPUM LAB			
Academia/research institution/think tank	Global	South School on Internet			
Academia/research institution/think tank	Global	GEODE			
Academia/research institution/think tank	Global	Center for Global Affairs			
Academia/research institution/think tank	Regional (Africa)	Research ICT Africa			
Academia/research	Regional	E4E			

Major Groups: **Civil Society and other Stakeholders**

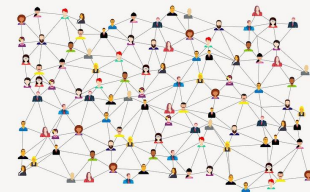
Civil Society and other Stakeholders

*all inputs are in PDF format.

- 1 of 8 Billion
- 100 Million

- Global Call to Action Against Poverty (GCAAP)
- Global Campaign for Education/US
- Global Futurist Initiative
- Global Cities Hub
- Global Coalition for Social Protection Floors
- The Global Crop Diversity Trust
- Global Enviro-Action
- Global Forum for Media Development
- Global Initiative Against Transnational Organized Crime
- Global Initiative for Digital Rights
- Global Initiative for Inclusive Information and Communication Technologies (G3ict)
- Global Network of Women Peacebuilders (GNWP)
- Global Pact Coalition
- Global Partners Digital, Access Now, European Center for Not-for-Profit Law, Derechos Digitales - América Latina,
- Alliance for Universal Digital Rights

- Postnety International
- Private Sector Mechanism of UN Committee on Food Security
- PROGETTO ALFA ASSOCIATION & CAMPUS SALUTE ASSOCIATION
- Project Liberty's Institute
- Project Ploughshares
- Qatar Youth Hostels
- R.K Atikpo Foundation LBG
- Red Agenda 2030 MX
- Reporters sans frontières (RSF)
- The Republican Public Association for



Analyse responses

Submission Papers

- Submission by Swiss Internet Governance Forum, June 2022
- Submission by New York University's Center for Global Affairs, May 2022
- Submission by Action Coalition on Innovation and Technology for Gender Equality, September 2022
- Submission by Chinese Academy of Sciences, September 2022
- Submission by International Federation of Library Associations and Institutions, September 2022
- Submission by Global Partners Digital, October 2022
- Submission by International Chamber of Commerce, October 2022
- Submission by Stimson Center (specifically pages 26-36), October 2022
- Submission by the fellows of the 2022 European Summer School on Internet Governance (EuroSSIG), November 2022
- Submission by UNFPA (part I), January 2023
- Submission by UNFPA (part II), April 2023
- Submission from multistakeholder Africa consultations facilitated by the Government of Germany and the Government of Kenya, Jan 2023
- Submission by the Alliance for Universal Digital Rights, January 2023
- Submission by the Jacob Blaustein Institute for the Advancement of Human Rights, February 2023
- Submission by the United States Council for International Business, February 2023
- Submission by the Global Security Alliance Greater China Region (GSA GCR), February 2023
- Submission by the Youth IGF Germany, February 2023
- Submission by Ahreum Lee, February 2023
- Submission by Fabian Luetz, February 2023
- Submission by the International Electrotechnical Commission (IEC), March 2023
- Submission by Concordia Consulting, March 2023
- Submission by European Union (EU) on behalf of the EU and its 27 Member States, March 2023
- Submission by Global Initiative for Inclusive Information and Communication Technologies (G3ict), March 2023
- Submission by the South School on Internet Governance, March 2023
- Submission from the Southern Alliance, March 2023
- Submission from the Open Data Charter, March 2023

How does AI fit in this?

Transparency

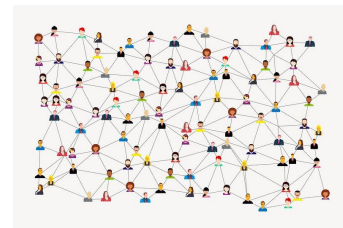
Attribution

Traceability

Agency

Contestability

Accountability



Analyse responses

Does AI decrease or
increase transparency?

Is transparency
technically possible?

GDC Assistant to the UN Tech Envoy Office

What are civil society organisations' opinions on governing AI?

Mar 13, 16:40

Civil society organizations generally support the regulation of AI to prioritize the protection and promotion of human rights, fairness, equity, and public interest. They advocate for clear legal frameworks, public debate, oversight mechanisms, ethical research, international collaborations, education, industry self-regulation, and collaboration between civil society and industry. They also emphasize the importance of proportionality, flexibility, and enforcement capacity in AI regulation.



Mar 13, 17:40

Sources ?

Send

Application powered by **DIPLO**

Is transparency technically possible?

1 File: Unknown File >> (Page: N/A)

53 Question 4: Should regulation of AI prioritize the protection and promotion of human rights, including freedom of expression, privacy, and due process? Respondent 01 02 Answer Yes Yes
Question 5: Should AI systems be designed and deployed in ways that promote fairness and equity, including through the elimination of bias and the inclusion of diverse perspectives in the development process? Respondent 01 02 Answer Yes Yes
accessible to all individuals, including those vulnerable to digital surveillance or digital technology or digital infrastructure? Respondent 01 02 Answer Yes Yes
regulation of AI prioritize the public interest? Respondent 01 02 Answer Yes Yes
must be designed to benefit society as a whole? Respondent 01 02 Answer Yes Yes
Should development and regulation of AI involve the participation of governments, and other stakeholders, and the public? Respondent 01 02 Answer Yes Yes
participation? Respondent 01 02 Answer Yes Yes

4 File: Unknown File >> (Page: N/A)

Virtual Consultation 24 of the conversation at the LAWS GGEs focuses narrowly on the risk to civilians of autonomous drones. There should also be scope for the discussion of escalation, non-proliferation, and strategic stability. The GDC could recommend a general ban on non-human control over nuclear weapons platform launches, building on commitments from nuclear weapons states, such as the United Kingdom and the US. • High-Level Panel of AI Experts: Society is facing a deluge of challenges, both short and long-term, emanating from AI. We find ourselves in a state of affairs where parties involved in the AI lifecycle or its governance cope with issues in a decentralized manner. In other words, many of the lessons learned about mitigating AI risks are either selectively shared or remain secret to protect a competitive advantage. Missing is a credible body that generates consensus-based assessments on fundamental AI challenges affecting society. Much like the Intergovernmental Panel on Climate Change, the UN would benefit its sustainability agenda by establishing a trustworthy body of AI experts. Its objective would be to

Is transparency technically possible?

Foundation models

AI models that are trained on broad data and can be adapted to a wide range of tasks across different domains. They serve as the basis for more specialised models dedicated to specific tasks or domains.

Is transparency technically possible?

Foundation models

Generative Models

A type of deep learning models that could generate high-quality content (texts, audios, images, videos, etc.) that are not seen in the data

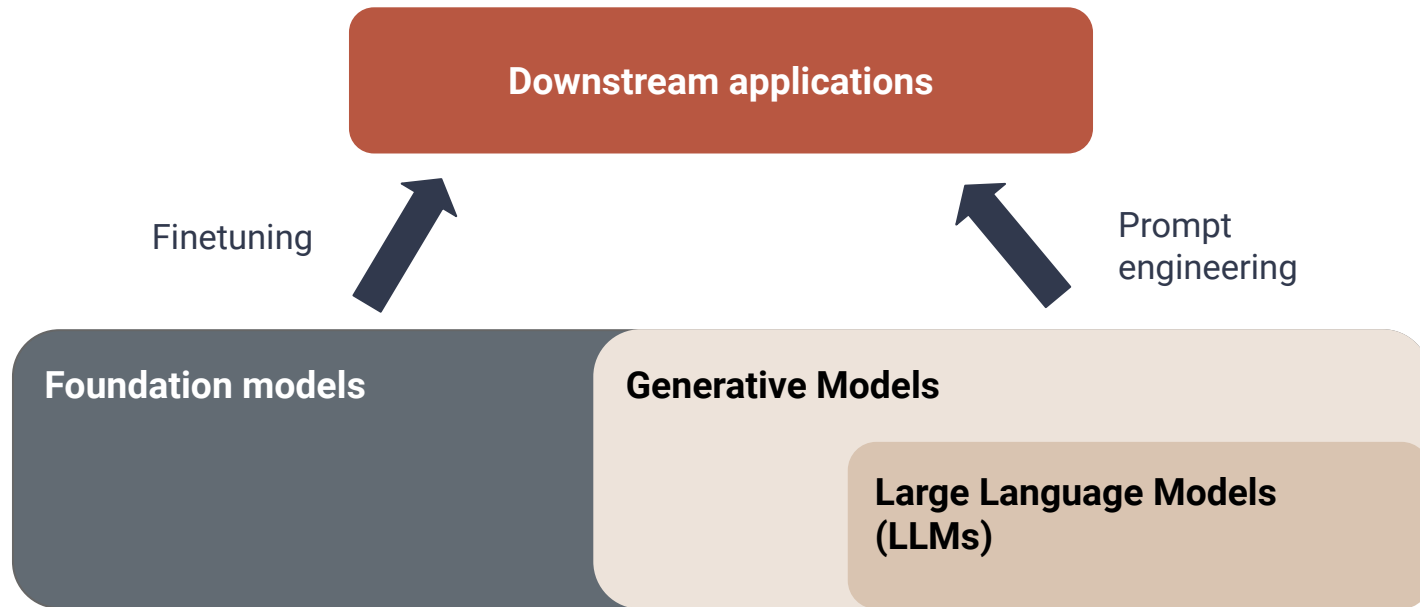
Is transparency technically possible?

Foundation models

Generative Models

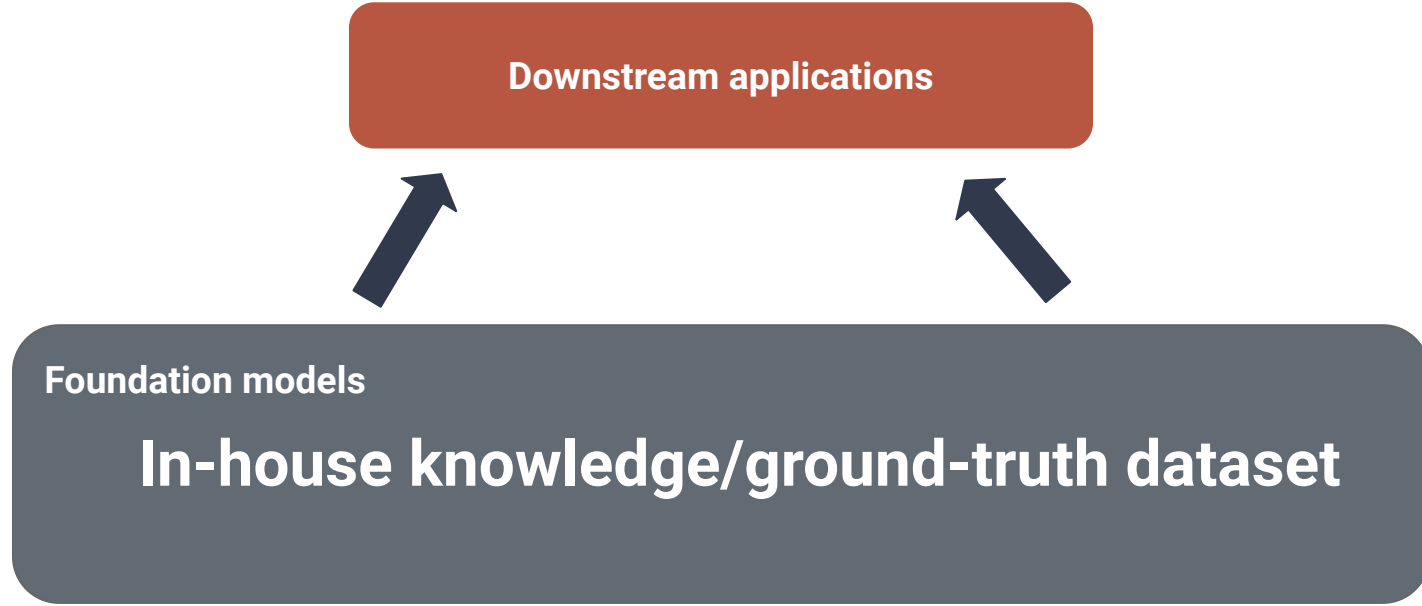
**Large Language Models
(LLMs)**

Is transparency technically possible?



Is transparency technically possible?

Q: What are civil society organisations' opinions on governing AI?

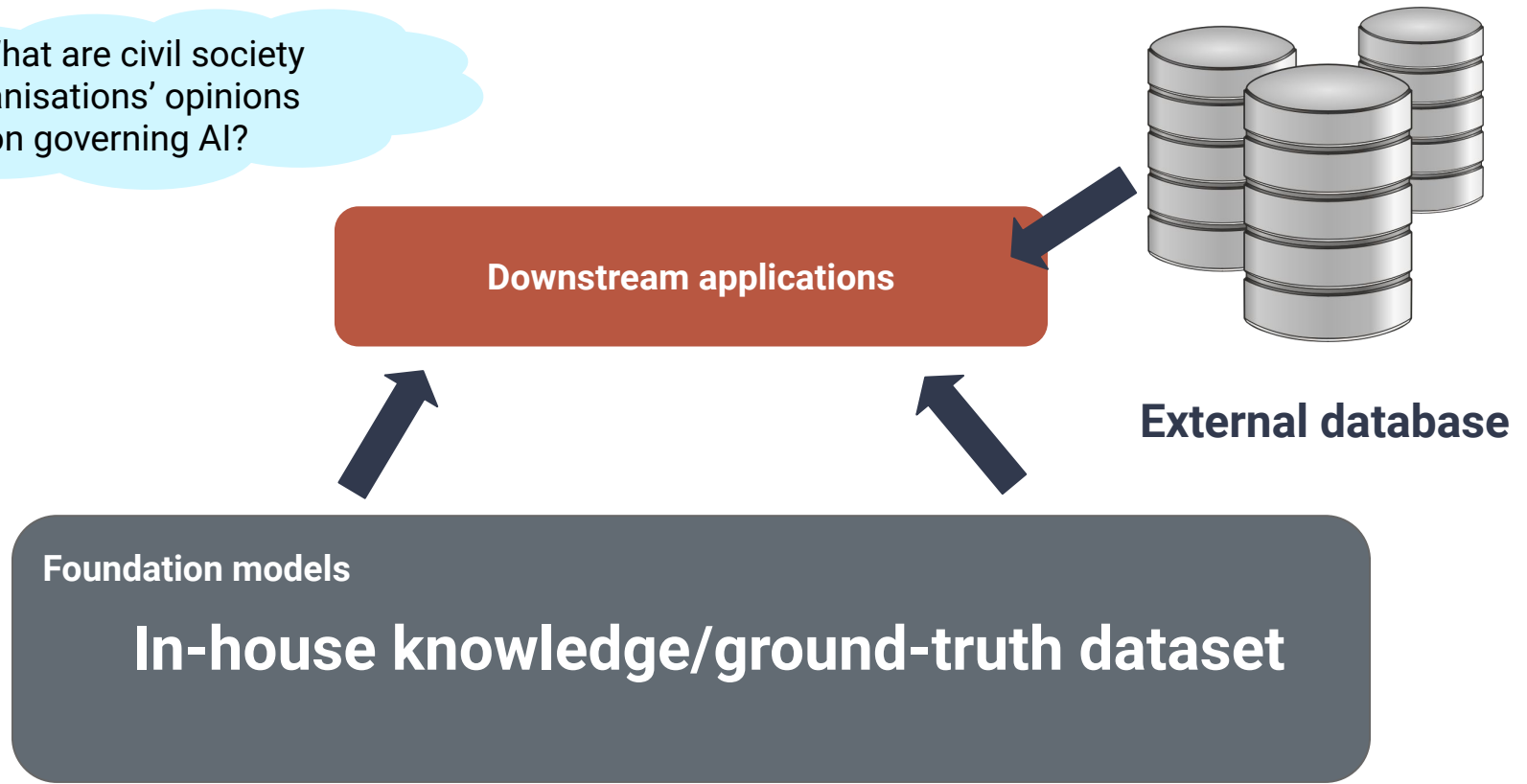


Is transparency technically possible?

- Retrieval-augmented generation (RAG)
 - Provide curated databases of documents for the LLM to reference before relying on in-house knowledge
 - More context, less hallucination

Is transparency technically possible?

Q: What are civil society organisations' opinions on governing AI?



Is transparency technically possible?

- Retrieval-augmented generation (RAG)
 - Provide curated databases of documents for the LLM to reference before relying on in-house knowledge
 - More context, less hallucination
- Source attribution
 - Provide the source on which the AI model generates the answer improves transparency and traceability

Is transparency technically possible?

GDC Assistant to the UN Tech Envoy Office

What are civil society organisations' opinions on governing AI?

Mar 13, 16:40

Civil society organizations generally support the regulation of AI to prioritize the protection and promotion of human rights, fairness, equity, and public interest. They advocate for clear legal frameworks, public debate, oversight mechanisms, ethical research, international collaborations, education, industry self-regulation, and collaboration between civil society and industry. They also emphasize the importance of proportionality, flexibility, and enforcement capacity in AI regulation.



Mar 13, 17:40

Sources ?

Send

Application powered by **DIPLO**

Is transparency technically possible?

1 File: Unknown File (Page: N/A)

53 Question 4: Should regulation of AI prioritize the protection and promotion of human rights, including freedom of expression, privacy, and due process? Respondent 01 02 Answer Yes Yes
Question 5: Should AI systems be designed and deployed in ways that promote fairness and equity, including through the elimination of bias and the inclusion of diverse perspectives in the development process? Respondent 01 02 An:
accessible to all individuals, including those v
technology or digital infrastructure? Responc
regulation of AI prioritize the public interest c
must be designed to benefit society as a whc
Should development and regulation of AI inv
governments, and other stakeholders, and m
participation? Respondent 01 02 Answer Yes >

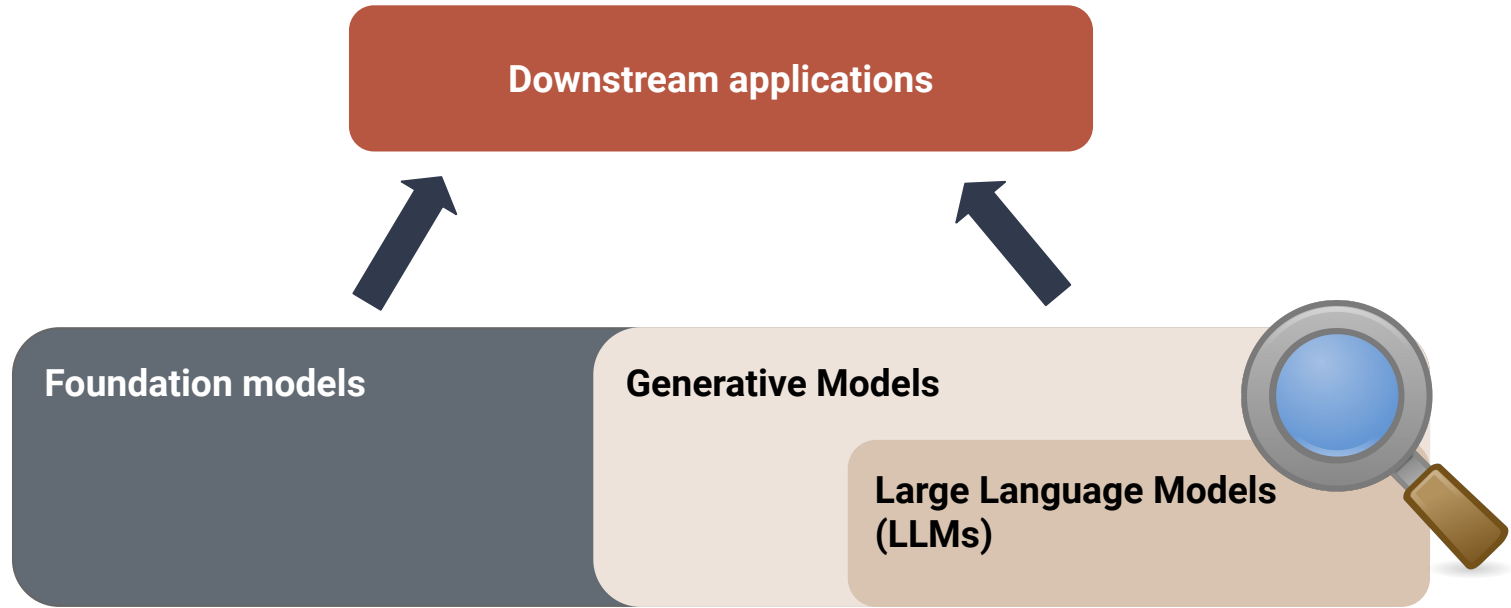
4 File: Unknown File (Page: N/A)

Virtual Consultation 24 of the conversation at the LAWS GGEs focuses narrowly on the risk to civilians of auton- omous drones. There should also be scope for the discussion of escalation, non-prolifer- ation, and strategic stability. The GDC could recommend a general ban on non-human control over nuclear weapons platform launches, building on commitments from nuclear weapons states, such as the United Kingdom and the US. • High-Level Panel of AI Experts: Society is facing a deluge of challenges, both short and long-term, emanating from AI. We find ourselves in a state of affairs where parties in- volved in the AI lifecycle or its governance cope with issues in a decentralized manner. In other words, many of the lessons learned about mitigating AI risks are either selectively shared or remain secret to protect a competitive advantage. Missing is a credible body that generates consensus-based assessments on fundamental AI challenges affecting society. Much like the Intergovernmental Panel on Climate Change, the UN would benefit its sustainability agenda by establishing a trustworthy body of AI experts. Its objective would be to

Is transparency technically possible?

- Retrieval-augmented generation (RAG)
 - Provide curated databases of documents for the LLM to reference before relying on in-house knowledge
- Source attribution
 - Provide the source on which the AI model generates the answer improves transparency and traceability
- Open-source models
 - Allow better insights into the architecture of the AI models

Is transparency technically possible?



Is transparency technically possible?

Is it black boxed by technology...

Or by humans?

Thank you!

diplomacy.edu

dig.watch

humanism.ai

wu@diplomacy.edu

D*i*PLO

Geneva Internet Platform

