

ACI'25 MUN



ANTALYA PRIVATE ACI COLLEGE MODEL UNITED NATIONS CONFERENCE 2025

UNESS AGENDA ITEM:

PROCESS EVALUATION AND IMMEDIATE ACTIONS CAN BE
TAKEN REGARDING THE CONSEQUENCES OF FAILURE OR
MISUSE OF ARTIFICIAL INTELLIGENCE CONTROL SYSTEMS

UNDERSECRETARY GENERAL:
HUSEYİN DEMİREL

ACADEMIC ASSISTANT
MUSTAFA ASLAN

Table of Contents,

1. *Letter from Secretary General*
2. *Letters from the Academy of F-UNESS*
 - a. *Letter from the Under Secretary General*
 - b. *Letter from Academic Assistant*
3. *Introduction of the Committee*
 - a. *History of the Committee*
 - b. *Main Goal of the Committee*
4. *Agenda Item: Process evaluation and immediate actions to be taken regarding the consequences of failure or misuse of artificial intelligence control systems*
 - a. *Introduction of the “Artificial Intelligence Systems”*
 - b. *Introduction of the frame for understanding the control system of AI Systems*
 - c. *Introduction of the usage, popularity and authority of the AI Systems in the future*
 - d. *Terms Related to Agenda*
5. *Countries' policies in the future related to the agenda item*
 - a. *Introduction of the country's resources and the budget on Artificial Intelligence*
 - b. *General information about the trustworthiness and transparency of the countries' AI control systems*
 - c. *General informations about the detrimental or affirmative consequences if any kind of errors happen on the AI control systems*
6. *Organisations that can assist UNESS*
 - a. *UNCSTD (The United Nations Commission on Science and Technology for Development)*
 - b. *TEC (Technology Executive Committee)*
7. *Questions to be Addressed*
8. *Further Reading*
9. *Bibliography*

1. Letter from Secretary General

First of all, let me extend a huge and warm welcome to all of the participants to the very first official session of Private Açı High School Model United Nations Conference of 2025, which will be held in Antalya from September 2nd to 4th!

I, Yasemin Raithel, as the Secretary-General of this well-planned and coordinated conference, it is my immense pleasure to be able to present this organization with this position. I sincerely wish you a fulfilling and thriving one. Model United Nations Conferences are not just a conference, you have a wonderful opportunity to level-up your knowledge of international relations and today's issues, gaining confidence whilst being a representative in the committees. Especially, members of the Academic Team; our Under-Secretary Generals, Academic Assistants and Board Members are the best you could ever ask for in Antalya society. Each of our committees are carefully selected for you to have the best experience ever. From advanced committees to beginner ones.

My utmost gratitude to the Principals and Teachers of our school, their effort and guidance made this conference possible. Moreover, I want to thank every single attendee of AÇIMUN'25, particularly our Academic and Organization Teams, who worked hard and served their best. Also a special appreciation goes to the only other person of the Executive Team, his organization skills are beyond any measure.

Last but not least, I want to leave a quote from Founder Father of the Republic of Turkey, Mustafa Kemal Atatürk:

Turkish Youth, your first duty is to preserve and to defend Turkish Independence and the Turkish Republic forever. This is the very foundation of your existence and your future. This foundation is your most precious treasure.

Lastly, whether you are a first timer delegate or an experienced member of ours, I wish everyone a marvelous experience and success during the conference. Hope to see you soon!

Sincerely,

Yasemin RAITHEL

Secretary-General of AÇIMUN'25

2. Letter from the Academy of F-UNESS

a. Letter from the Under Secretary General

Dear Participants,

It gives me the utmost pleasure to serve as the Under-Secretary-General of UNESS(United Nations Emergency Special Session) at this prestigious conference.

One of the largest problems in the usage of artificial intelligence is the detection and the limits of them; the transparency and trustability of these processes. That's why, as the UNESS committee, we must take action and find various solutions.

Every single delegate is encouraged to read this study guide carefully, conduct further and extra research specifically on their country, and come prepared for the committee. I sincerely hope that this committee will be an unforgettable experience for all of you.

For my closing remarks, I would like to express my thanks to the hardworking Executive Team . I also want to welcome each and every one of my delegates.

Last but not least, I would like to thank my handsome academic assistant for helping to create such an enjoyable committee.

King Regards

Under-Secretary General of the UNESS
Hüseyin Demirel

+90 551 026 99 33

demirelhuseyin0411@gmail.com

b. Letter from the Academic Assistant

Dear Participants,

First of all, I would like to welcome you all to the AÇIMUN'25 F-UNESS committee. I am Mustafa Aslan, a 12th grade student at Bahçeşehir Aspendos Campus, and I am honored to serve as the Academic Assistant of the committee.

I have been attending Model United Nations conferences in Antalya and many other cities for the last two years and this is my 24th conference.

I would like to thank Yasemin Raithel for giving me an opportunity to be here as a member of the academic team.

Besides these, if you have any questions about the committee, please contact me from my contact information below, even for the tiniest thing you want to ask.

We have added all the necessary information in the committee to the study guide as an academic team.

I wish you all success in advance.

Academic Assistant, Mustafa ASLAN

+05324246907 whatsapp
aslanmustafa0770@gmail.com

3. Introduction of the Committee

The Emergency Special Session (ESS) of the United Nations General Assembly is a rare but powerful diplomatic forum convened under Resolution 377A(V), famously titled 'Uniting for Peace.'

Adopted in 1950, this resolution empowers the General Assembly to consider urgent threats to peace and security when the Security Council is deadlocked, and is unable to fulfill its duty of maintaining international peace and security.

This mechanism has been invoked only 11 times in UN history.

UNGA ESS challenges delegates to respond rapidly, diplomatically, and strategically in the face of crisis. If you're a delegate who enjoys fast-paced negotiations, handling high-pressure situations, and crafting impactful resolutions while working around global power dynamics.

What makes these sessions ‘special’?

Only 11 emergency special sessions have been called in the last 73 years. The eleventh was convened, following a Security Council veto, six days after Russia’s full-scale invasion of Ukraine in 2022, resulting in a broadly supported resolution on the crisis.

What makes it ‘special’ is that despite Council disunity, the countries of the world can address urgent global peace and security matters together. At the request of Member States or the Council itself, the Assembly President shall convene an emergency special session within 24 hours.

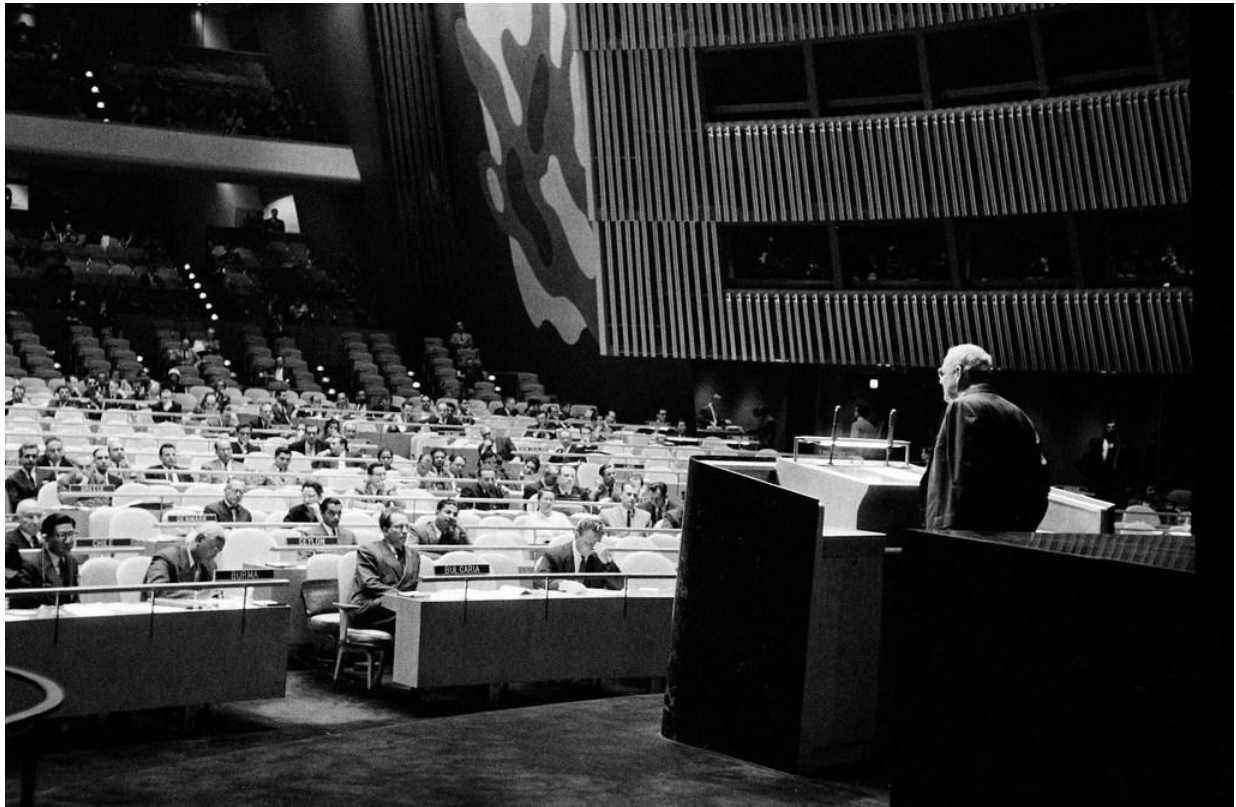
The establishment and guiding rules of these sessions came after the outbreak of the Korean Peninsula conflict in 1950, when the Assembly adopted its landmark “Uniting for peace” resolution.

Another unique element is that unlike the 15-member Security Council, a veto cannot be used in the 193-member General Assembly. If a two-thirds majority of the Assembly votes in favour, a proposed draft resolution is adopted.

What’s another difference between the Council and the Assembly? While the former’s resolutions are legally binding, the latter’s are not.

a. History of the Committee

The first emergency special session of the United Nations General Assembly was convened on 1 November and ended on 10 November 1956 resolving the Suez Crisis by creating the United Nations Emergency Force (UNEF) to provide an international presence between the belligerents in the canal zone. The [First Emergency Special Session](#) was convened due to the failure of the [Security Council](#) to resolve the instability at the Suez Canal. This forced an invocation of the "Uniting for Peace" resolution, which transferred the issue from the Security Council to the General Assembly in its [Emergency Special Session](#) (ESS) guise. On the fourth day of the ESS the Canadian representative, Lester B. Pearson, introduced the concept of a UN police force. The creation of the United Nations Emergency Force (the first peacekeeping force) was approved by the General Assembly with 57 supports and zero opposes. The vote had 19 countries abstaining, including the United Kingdom, France, Egypt, the Soviet Union and several Eastern European countries.



UN Photo The first ever emergency special session of the UN General Assembly on the Middle East takes place in 1956.

The issue developed after years of attacks by Egypt on Israel. Egypt abrogated the Anglo-Egyptian Treaty of 1936, began restricting Israeli shipping, and in 1955 turned to Czechoslovakia to purchase weapons. In July 1956 the United States withdrew financial assistance for Egypt's Aswan Dam project, leading Egypt to nationalize the Suez Canal Company. In September, the Security Council convened to consider "the situation created by the unilateral action of the Egyptian government in bringing to an end the system of international operation of the Suez Canal" and "actions against Egypt by some Powers, particularly France and the United Kingdom, which constitute a danger to international peace and security and are serious violations of the Charter of the United Nations". In October the Security Council passed Resolution 118 calling on the sovereignty of Egypt to be respected and the operation of the Suez Canal to be insulated from the politics of any country. However, Israel invaded Egypt shortly after. An American draft resolution calling on Israel to withdraw from Egypt to behind armistice lines was vetoed by France and the United Kingdom who joined Israeli forces in the invasion. Security Council Resolution 119 passed on October 31 admitted its failure to maintain international peace and security invoking the 1950 "Uniting for Peace" resolution triggering an emergency special session of the General Assembly.

b. Main Goal of the Committee

When the UN Security Council is unable to shoulder its primary responsibility to maintain international peace and security, as has thus far been the case following the 7 October outbreak of the latest Israel-Palestine crisis, the General Assembly can consider the matter immediately in an emergency special session.

UN Member States can request the Assembly President to convene such a session, with a view to making appropriate recommendations for collective measures, including calling for ceasefires, and in the case of a breach of the peace or act of aggression, as a last resort, the use of armed force when necessary.

At the first plenary meeting of an emergency special session or upon its resumption, as is the case with the current tenth session, the Assembly typically hears from the country or countries of concern. Member States then debate the issue and vote on a draft resolution.

An emergency special session swiftly happens when the UN Security Council is deadlocked. That occurs when one of the five permanent Council members (China, France, Russia, United Kingdom, United States) uses their UN Charter-mandated veto power to quash a related draft resolution.

While veto use can also be quickly scrutinized at a regular General Assembly meeting, emergency special sessions are one option the UN membership has to debate urgent peace and security matters.

The nature of the Uniting for Peace mechanism has been extensively analyzed, such that only a brief outline of its key features is necessary here. In response to Security Council deadlock on continued UN military action in Korea, the General Assembly adopted [Resolution 377A](#) on 3 November 1950. The resolution stipulates that either the General Assembly or Security Council can initiate an Emergency Special Session where, due to a “lack of unanimity of the permanent members,” the Council “fails to exercise its primary responsibility for the maintenance of international peace and security.” In this case, the Assembly “shall consider the matter immediately with a view to making appropriate recommendations to Members for collective measures, including in the case of a breach of the peace or act of aggression the use of armed force when necessary, to maintain or restore international peace and security.” In the ten prior Emergency Special Sessions, the Assembly took and recommended a variety of measures. It has condemned violations of international law and called for cessation of these breaches. It has recommended the imposition of sanctions against offending states. It has established peacekeeping forces with host state consent. Perhaps most famously associated with the Uniting for Peace mechanism, in 1951 the Assembly called upon states to support continued UN military action in Korea, including to repel Chinese aggression, a feat that it has not repeated since.

4. Agenda Item: Process evaluation and immediate actions to be taken regarding the consequences of failure or misuse of artificial intelligence control system

a. Introduction of the “ Artificial Intelligence Systems”

Artificial intelligence is a branch of research concerned with developing computers and machines that can reason, learn, and behave in ways that would ordinarily need human intelligence, or that involve data on a scale that humans cannot handle.

AI is a multidisciplinary field that includes computer science, data analytics and statistics, hardware and software engineering, languages, neurology, philosophy, and psychology.

On an operational level, AI is a group of technologies based mostly on machine learning and deep learning that are used for data analytics, predictions and forecasts, object classification, natural language processing, recommendations, intelligent data retrieval, and other applications.

b. Introduction of the frame for understanding the control system of AI Systems

According to the current results, Artificial Intelligence Systems are improving day by day but on the other hand these systems need to be under control and be detected. That's why the AI Control Systems has been developed.

AI Control Systems are programs which fully or semi-controlled by the humans that are professionals in current technology. These systems basically has responsible for management of the improving process of the AI. And they are here to control any kind of misuse or failure situations possible.

As it mentioned, these control systems are working under the human professionals. Because of that these control systems has limited power and usage frame. Despite control systems, AI systems has unlimited developing capacity.

Note from Under Secretary General: My dear delegates, the academy of this committee is highly suggested to do further research about the AI control systems and their capacity. We are waiting for your new ideas related to improving process of AI control systems

c. Introduction of the usage, popularity and authority of the AI Systems in the future

In the future, approximately all of the scientist expected detrimental and affirmative things such as,

- Usage of AI will be increased all over the world including undeveloped countries
- The AI systems will be the main power of the countries
- Possible wars will be happen in the cyber world not on the war field

- Humans will not be necessary for so many sectors which is currently popular such as textile, agriculture, industry, education, health.
- In the future, humans will have approximately full access to any kind of human needs
- Class discrimination among the people according to their economic situations will be concluded
- Welfare levels will be increased because of the sustainable productions and trade

d. Terms Related to Agenda Item

Cyber Security: Cybersecurity is the activity of defending systems, networks, and programs against digital threats. These intrusions are typically intended to access, change, or delete sensitive information; extract money from users via ransomware; or disrupt normal corporate activities.

White Hat Hackers: White hat hackers, often known as "ethical hackers" or "good hackers," are the antithesis of black hats. They exploit computer systems or networks to identify security holes and give recommendations for enhancements.

Black Hat Hackers: Black hat hackers are criminals who infiltrate computer networks with harmful intent. They may also distribute malware that destroys files, holds computers hostage, or steals passwords, credit card numbers, and other personally identifiable information.

Black hats are driven by self-interest, such as financial gain, revenge, or simply wreaking devastation. Sometimes their aim is ideological, such as targeting persons with whom they strongly disagree.

AI (Artificial Intelligence) Experts: AI professionals design, build, and optimize artificial intelligence systems to address challenging issues. These individuals combine extensive academic knowledge with practical applications in machine learning, natural language processing, and computer vision.

Beyond coding and model creation, AI specialists manage the ethical implications of their work, ensuring that AI technologies are used properly. They bridge the gap between theoretical research and real-world applications by frequently cooperating with cross-functional teams to integrate AI solutions into existing systems and visiting AI conferences to demonstrate their knowledge.

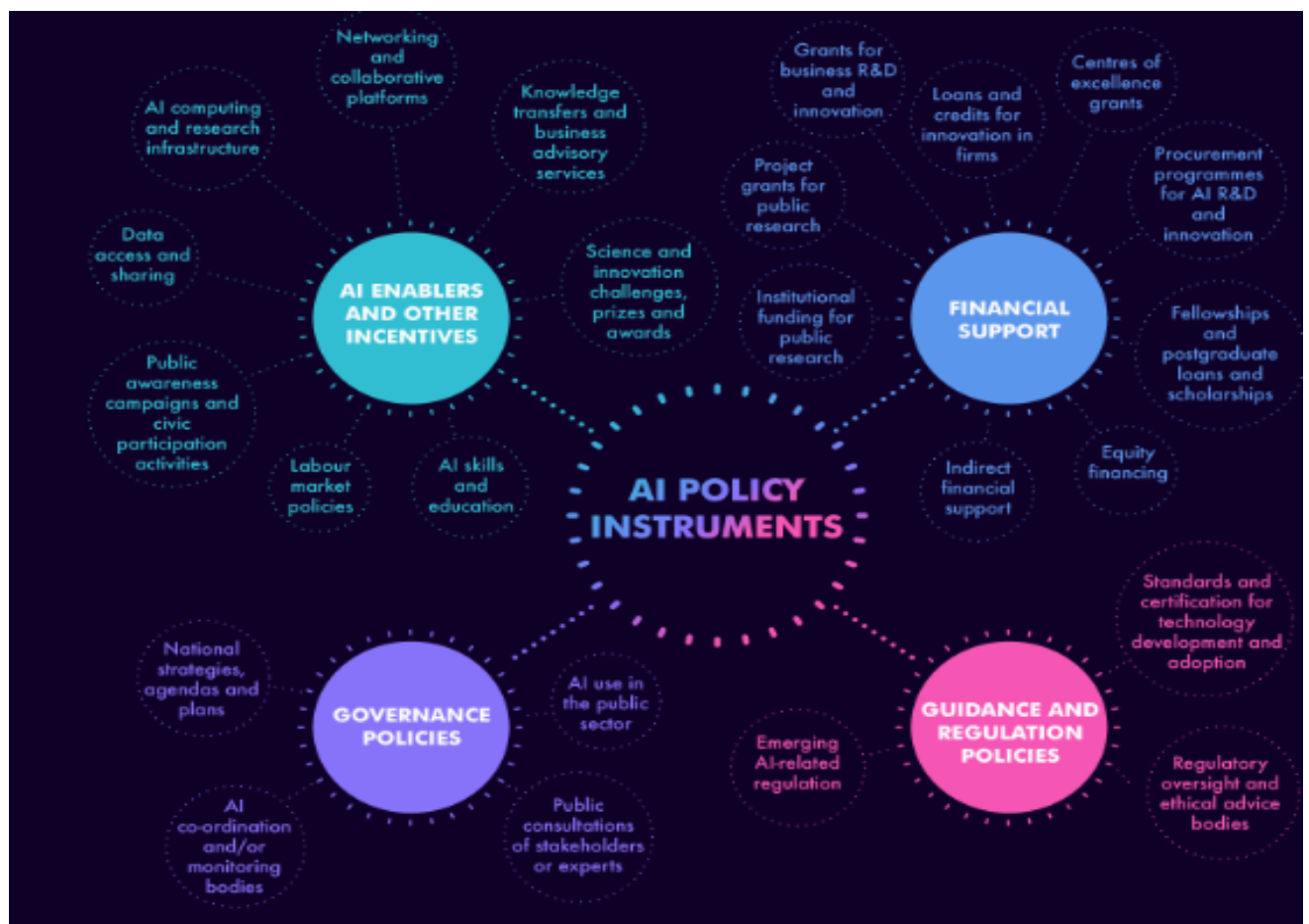
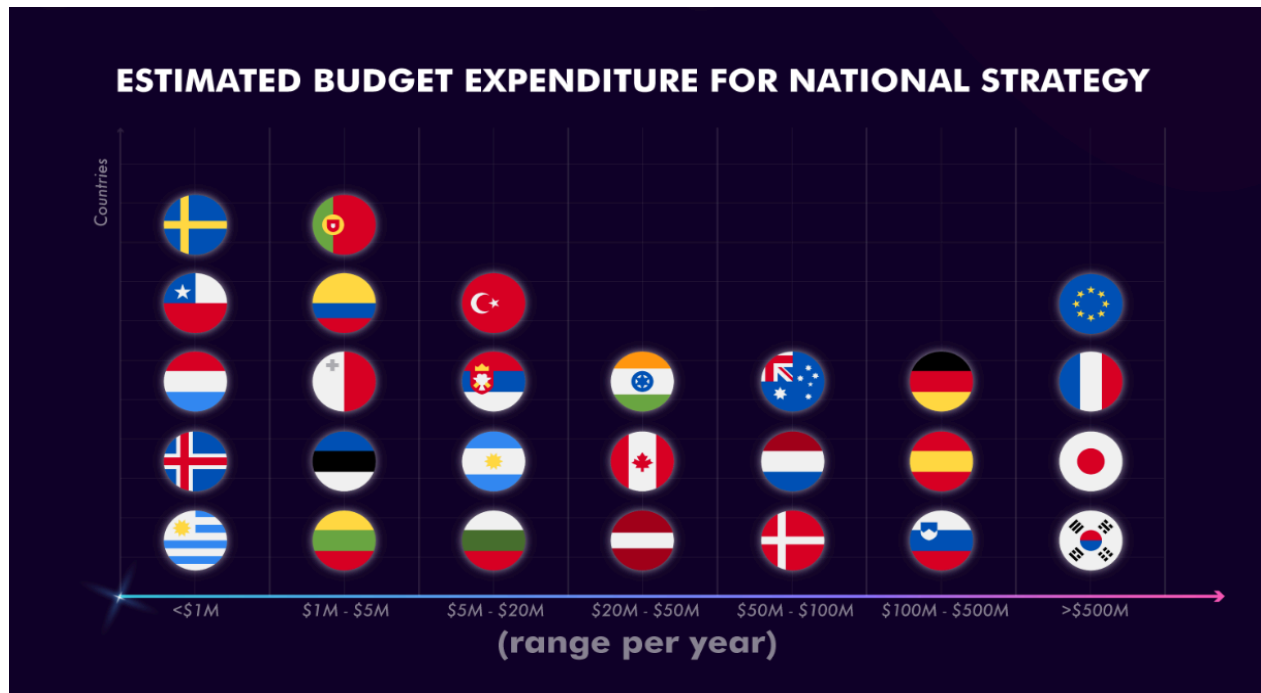
5. Countries' policies in the future related to the agenda item

a. Introduction of the country's resources and the budget on Artificial Intelligence



1	United States	82	YES	Executive Order on the Safe, Secure Development and Use of Artificial Intelligence	2023	2023
2	European Union	63	YES	Commission Regulation on Artificial Intelligence	2023	2023
3	United Kingdom	61	YES	Artificial Intelligence Strategy	2023	2023
4	Australia	44	YES	Australian AI Action Plan	2023	2023
5	Germany	38	YES	National AI Strategy	2023	2023
6	Portugal	36	YES	AI Strategy 2024 - National Strategy for AI	2024	2024
7	Turkey	36	YES	National AI Strategy	2023	2023
8	France	34	YES	National Strategy on AI	2023	2023
9	India	32	YES	National Strategy on Artificial Intelligence	2023	2023
10	Colombia	30	YES	AI National Strategy	2023	2023
11	Spain	28	YES	National AI Strategy	2023	2023
12	Japan	27	YES	AI Strategy	2023	2023
13	Singapore	25	YES	National AI Strategy	2023	2023
14	Belgium	25	YES	National Governance Plan for the Development of AI	2023	2023
15	China	22	YES	National AI Strategy	2023	2023
16	Norway	21	YES	National Strategy for AI	2023	2023
17	Serbia	21	YES	Strategy for the Development of AI in the Republic of Serbia	2023	2023
18	Malta	19	YES	National AI Strategy	2023	2023
19	Peru	19	YES	National AI Strategy	2023	2023
20	Israel	18	NO	-	-	-
21	Estonia	17	YES	National AI Strategy	2023	2023
22	Canada	16	YES	First Canadian AI Strategy	2023	2023
23	Mexico	15	YES	National Strategy on Artificial Intelligence	2023	2023
24	Slovenia	15	YES	National AI Strategy	2023	2023
25	Hungary	14	YES	National AI Strategy	2023	2023
26	Italy	14	YES	AI Strategy - Program	2023	2023
27	Korea	14	YES	National Strategy for AI	2023	2023
28	Netherlands	13	YES	Strategy for the Development of AI	2023	2023
29	Sweden	13	YES	National Strategy for AI	2023	2023
30	Argentina	12	YES	AI National Plan	2023	2023
31	Brazil	12	YES	Artificial Intelligence Strategy	2023	2023
32	Denmark	12	YES	National Strategy for AI	2023	2023
33	Finland	12	YES	AI Strategy	2023	2023
34	Chile	11	YES	AI National Strategy	2023	2023
35	New Zealand	11	NO	-	-	-
36	Austria	10	YES	AI Strategy 2024	2024	2024
37	Czech Republic	10	YES	National Strategy of the Czech Republic	2023	2023
38	Egypt	10	YES	Egypt National AI Strategy	2023	2023
39	Ireland	10	YES	National AI Strategy	2023	2023
40	U.A.E.	10	YES	National Strategy for AI	2023	2023
41	Costa Rica	7	NO	-	-	-
42	Kazakhstan	7	NO	-	-	-
43	Tunisia	7	NO	-	-	-
44	Kenya	6	NO	-	-	-
45	Luxembourg	6	YES	AI Strategy: Vision for Luxembourg	2023	2023
46	Rwanda	6	NO	-	-	-
47	Saudi Arabia	6	YES	National Digital and AI Strategy	2023	2023
48	Switzerland	6	YES	Digital Switzerland Strategy	2023	2023
49	Thailand	6	YES	Thailand National AI Strategy	2023	2023
50	Viet Nam	6	YES	National Strategy on AI	2023	2023
51	Latvia	6	YES	National AI Strategy	2023	2023
52	Greece	4	NO	-	-	-
53	Iceland	4	YES	National AI Strategy	2023	2023
54	Lithuania	4	YES	AI Strategy	2023	2023
55	Morocco	4	YES	Plan for the Development of Artificial Intelligence in Morocco	2023	2023
56	Poland	4	YES	National AI Strategy	2023	2023
57	Romania	4	YES	National AI Strategy	2023	2023
58	South Africa	4	NO	-	-	-
59	Uruguay	4	YES	AI Strategy for the Digital Development	2023	2023
60	Armenia	4	NO	-	-	-
61	Uganda	4	NO	-	-	-
62	Uzbekistan	4	NO	-	-	-
63	African Union	4	NO	-	-	-
64	Bulgaria	4	YES	Strategy for the Development of Artificial Intelligence in Bulgaria	2023	2023
65	Cyprus	2	YES	National AI Strategy	2023	2023
66	Mauritius	2	YES	National AI Strategy	2023	2023
67	Nigeria	2	NO	-	-	-
68	Slovakia	2	NO	-	-	-
69	Croatia	2	YES	National Plan for the Development of AI	2023	2023
70	Indonesia	1	YES	National AI Strategy	2023	2023
71	Ukraine	1	NO	-	-	-

Note from Under Secretary General: My dear delegates, you can reach full table from this link ; <https://www.aiprm.com/ai-laws-around-the-world/ai-laws-around-world.png>



b. General information about the trustworthiness and transparency of the countries' AI control systems

- United States of America: According to the results, The USA has very improved AI systems. They are using these systems in so many sectors such as textile, education, production, industry etc. Because of that America is focusing on the education of the future experts of AI to successfully manage the process of the development of AI control systems. Despite the insufficient information about the America's budget on AI control systems they are known as a powerful country in the technology sector.
- United Kingdom: UK is known as a powerful country for their education, scholarship, and the capacity for technological systems such as AI. With the help of this education system UK is a trustworthy country to raise a new generation of AI experts.

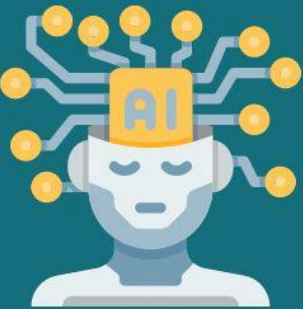
Note from Under Secretary General: My dear delegates, you can always get a more information to find a new ideas about the structure of the education systems in UK via this link;

<https://www.findaphd.com/phds/departments/1367/school-of-science-engineering-and-environment>

c. General informations about the detrimental or affirmative consequences if any kind of errors happen on the AI control systems

This article focuses on the risks and consequences of irresponsible AI implementation within organizations, offering guidance and resources on building responsible AI systems and adhering to relevant regulations.

To align compliance, security, and business goals, integrate security measures into business processes, prioritize risk management, and ensure regulatory compliance.



www.trustcloud.ai

RISKS AND CONSEQUENCES OF irresponsible AI in organizations

TrustCloud

- 1 Introduction to AI in organizations
- 2 Understanding the risks of irresponsible AI
- 3 The consequences of irresponsible AI
- 4 Real-life examples of AI gone wrong
- 5 Ethical considerations in AI implementation
- 6 Regulatory frameworks
- 7 Building responsible AI systems
- 8 The role of stakeholders in ensuring responsible AI
- 9 The future of responsible AI in organizations

It also emphasizes the importance of ethical AI development and compliance with frameworks such as GDPR, [HIPAA](#), and ISO standards. It provides an overview of several compliance standards and offers assistance in achieving them through TrustCloud's services.

Understanding the risks of irresponsible AI

The risks associated with irresponsible AI implementation are multifaceted and can manifest in various ways. One of the primary concerns is the potential for AI systems to perpetuate and amplify existing biases, leading to discriminatory outcomes. This can occur when the training data used to develop AI models is skewed or when the algorithms themselves are designed with inherent biases. Another significant risk lies in the lack of transparency and accountability in AI decision-making processes. Opaque "black box" algorithms can make it challenging to understand how decisions are reached, raising concerns about fairness, privacy, and due process. Additionally, the potential for AI systems to be manipulated or misused for malicious purposes, such as cyberattacks or the spread of disinformation, poses a severe threat to organizational security and integrity.

The consequences of irresponsible AI

The implementation of Artificial Intelligence (AI) in organizations has been a game-changer in terms of efficiency, productivity, and decision-making. However, the consequences of irresponsible AI implementation can be severe and far-reaching.

One of the major consequences is a loss of trust among customers and stakeholders. If AI algorithms are not properly designed or if they are biased, it can lead to unfair treatment or discrimination. This can damage the reputation of the organization and result in a loss of customers. Moreover, if AI systems are not properly secured, they can also be vulnerable to hacking or data breaches, which can have serious consequences for both the organization and its customers.

Another consequence of irresponsible AI implementation is the potential loss of jobs. While AI has the potential to automate repetitive tasks and improve efficiency, if not implemented responsibly, it can also lead to job displacement. This can have a significant impact on employees and their livelihoods. Organizations need to ensure that they have a clear plan in place to reskill or retrain employees whose jobs may be affected by AI implementation.

Furthermore, irresponsible AI implementation can also lead to legal and ethical implications. If AI systems are used for purposes that violate privacy laws or ethical guidelines, organizations can face legal repercussions and damage their reputation. For example, if AI is used to collect and analyze personal data without proper consent or transparency, it can be seen as a breach of privacy rights. Organizations must ensure that they adhere to relevant laws and regulations and implement AI in an ethical manner.

Here are six potential consequences of irresponsible AI use within an organization:

- Data Privacy Violations:** Misuse of sensitive data in AI algorithms can lead to privacy breaches, violating regulations like GDPR or HIPAA and leading to legal penalties and loss of trust.

- Bias and Discrimination:** AI models trained on biased data may reinforce stereotypes or make discriminatory decisions, harming reputation and exposing the organization to lawsuits.

- Security Vulnerabilities:** Insufficiently secured AI systems can be susceptible to hacking and data manipulation, increasing the risk of breaches and information leaks.

- Financial Loss:** Misguided AI-driven decisions in areas like finance or operations can lead to costly errors, impacting revenue and operational stability.

Reputational Damage: Negative public perception from irresponsible AI use can damage customer trust, brand reputation, and long-term profitability.

Regulatory Scrutiny: Non-compliance with emerging AI regulations and ethical standards may result in increased scrutiny, penalties, and restrictions on AI usage within the organization.

Responsible AI governance is essential to avoid these risks and build reliable, fair, and secure systems.

The importance of risk mitigation in addressing risks and consequences of irresponsible AI

Artificial Intelligence (AI) has become a powerful force for innovation, efficiency, and growth in organizations. But as with any transformative technology, it comes with risks—especially when implemented irresponsibly. From biased algorithms to data privacy violations, the consequences of unchecked AI can be severe. That's where risk mitigation comes into play, helping organizations harness AI's potential while avoiding pitfalls.

The risks of irresponsible AI

Bias and Discrimination:

AI systems trained on biased data can unintentionally amplify inequalities, leading to unfair decisions in hiring, lending, or law enforcement.

Data Privacy Breaches:

Mishandling sensitive data in AI models can expose organizations to legal penalties and erode customer trust.

Operational Failures:

Relying on poorly designed AI can result in errors, misinterpretations, or system crashes that disrupt business processes.

Reputation Damage:

Public backlash from unethical AI use can harm an organization's brand and stakeholder confidence.

The role of risk mitigation in AI governance

To address these challenges, organizations need a structured approach to risk mitigation. This involves identifying potential risks early, putting safeguards in place, and continuously monitoring AI systems to ensure responsible use.

Key strategies for risk mitigation in AI

Transparent Development:

Build AI models with explainability in mind. Stakeholders should understand how decisions are made and be able to audit processes when needed.

Ethical Data Practices:

Use diverse, high-quality datasets to reduce biases. Implement strict data governance policies to ensure privacy and compliance with regulations.

Human Oversight:

Risk mitigation in AI means keeping humans in the loop for critical decisions. This reduces the likelihood of automated systems making harmful or irreversible mistakes.

Scenario Testing:

Test AI systems for potential edge cases and unintended consequences. This helps identify weaknesses before deployment.

Continuous Monitoring and Updates:

AI evolves, and so do risks. Regular reviews and updates ensure the system remains ethical, accurate, and aligned with organizational goals.

Benefits of risk mitigation in AI

Protects Stakeholders:

Mitigating risks ensures fair treatment of customers, employees, and other stakeholders impacted by AI decisions.

Strengthens Compliance:

Adhering to legal and ethical standards helps organizations avoid costly penalties.

Enhances Trust:

A well-managed AI strategy builds confidence among users and partners.

Drives Sustainable Innovation:

Responsible AI practices encourage innovation without sacrificing ethics or accountability.

AI offers immense opportunities, but it's not without its challenges. Irresponsible use can lead to significant risks and consequences, potentially overshadowing its benefits. By focusing on risk mitigation, organizations can strike a balance—leveraging AI's power responsibly while safeguarding against potential harm.

Risk mitigation isn't just about avoiding problems; it's about creating AI systems that are fair, transparent, and built to support long-term success. With the right strategies in place, organizations can ensure their AI efforts lead to meaningful, positive outcomes for everyone involved.

6. Organisations that can assist UNESS

a. UNCSTD (The United Nations Commission on Science and Technology for Development)

The United Nations Commission on Science and Technology for Development (CSTD) is a subsidiary body of the Economic and Social Council (ECOSOC). It holds an annual intergovernmental forum for discussion on timely and pertinent issues affecting science,

technology and development. Since 2005, the Commission has been mandated by ECOSOC to serve as the focal point in the system-wide follow-up to the outcomes of the World Summit on the Information Society (WSIS).

Its members are composed of national Governments, however civil society contributes to discussions that take place. Strong links exist with other UN bodies (The Commission on Status of Women, Regional Commissions, ITU, UNESCO).

Outcomes of the CSTD include providing the United Nations General Assembly and ECOSOC with high-level advice on relevant science and technology issues. UNCTAD provides secretariat service and substantive support to the Commission.

What is it?

The United Nations Commission on Science and Technology for Development (CSTD) is the United Nations' home for discussions on science and technology – what is new, what matters, what is changing, what the impact is – and how this affects development and a sustainable future for all.

It is the forum that helps ask and frame the critical issues influencing the fields of science and technology today.

Some of the important normative issues raised include the technology and life interface, as well as governance of the use and development of frontier technologies – namely, big data analytics, biotech and genome editing, the Internet of things (IoT) and artificial intelligence. The CSTD is also an open platform where proposals, ideas, experiences, cases, and intellectual thought can be channeled toward making a policy impact. It facilitates concrete collaborations between member states, NGOs and actors in the science, technology and development space.

Why was it established?

The distant origins of the CSTD are at the UN Conference on Science and Technology for Development held in Vienna in 1979, where an Intergovernmental Committee on Science and Technology for Development was created.

In 1992 the General Assembly decided to transform the Committee into a functional commission of the Economic and Social Council (ECOSOC) and set up the CSTD.

The CSTD was created to provide the UN with high-level advice through analysis and policy recommendations to:

- Guide the future work of the United Nations
- Develop common policies
- Agree on appropriate actions

It provides a forum where countries can raise critical challenges and explore opportunities presented by rapid technological development – to ensure developing countries and people do not get left behind.

The world has changed radically since 1992. Even then, the seeds of scientific and technological discovery sown throughout the twentieth century were demonstrating exponential growth. Now it is unparalleled.

Today the CSTD matters more than ever, as the opportunities and obstacles at the juncture of science, technology and innovation and the need for sustainable development become more complex, global and far-reaching.

What does it do?

The CSTD provides the annual intergovernmental platform and forum for discussion of timely and pertinent issues affecting science, technology and development.

It has a long-established process for facilitating dialogue that can lead to policy change.

Why does it matter?

The transfer of knowledge, skills and solutions in the science and technology fields can have a profound and lasting impact on the development trajectories of nations and peoples.

The world cannot develop or advance – nor can we achieve the Sustainable Development Goals – without modern technology and new solutions to old problems.

The CSTD provides a space for the crucial discussions of how we do this practically, ethically, timeously and to the benefit of all.

It also helps map out the challenges -- some long-standing and many new -- so that we can keep pace with the rapid advances that alter the landscape of science and technology fields.

The CSTD offers a space for collaborations between member States, academia, civil society and the business community engaged in science and technology for development.

The CSTD also serves to map new ways that developing countries contribute to science, technology and innovation. In some cases, innovation is also redefined by those in developing countries while trying to reach the targets set under the Sustainable Development Goals.

How does it fit into the UN ecosystem?

The CSTD is a subsidiary body of ECOSOC. The Commission met for the first time in April 1993 in New York, USA. Since July 1993, UNCTAD has been hosting the secretariat of the Commission which meets annually in Geneva, Switzerland.

Who are the major stakeholders?

The members of the CSTD are national Governments, but discussions also involve civil society representatives. Strong links exist with other UN bodies (i.e. Commission on the Status of Women, ITU, Regional Commissions, UNESCO).

So what?

Today, more than ever before, people risk being left behind as technology and innovation outpaces both their and Government's ability to keep up.

There is a great risk of the digital divide widening further and with it the ability of developing countries to harness science, technology and innovation for their own development.

Time is running out – mainly because of the extreme pace at which innovation is happening.

The goal is to help developing countries – and all people who lack access and opportunity – to benefit from and use science, technology and innovation to address their own development challenges – and innovate within these.

b. TEC (Technology Executive Committee)

Created in 2010, the Technology Executive Committee (TEC) is the policy arm of the Technology Mechanism. It focuses on identifying policies that can accelerate the development and transfer of low-emission and climate resilient technologies.

The TEC and the Climate Technology Centre and Network (CTCN) form the Technology Mechanism. With the Technology Mechanism serving the Paris Agreement, the TEC will play a key role in supporting countries to identify climate technology policies that support them to achieve the Agreement's objectives.

The TEC consists of 22 technology experts representing developed and developing countries. It meets at least twice a year and holds climate technology events to support efforts to address technology-related policy issues. Each year the TEC reports to the Conference of the Parties (COP) on its performance and activities. Specifically, the TEC analyses climate technology issues and develops balanced policy recommendations, supporting countries to accelerate action on climate change.

Policies

The TEC's key outputs are its annual technology-related recommendations to the COP. Through these, the TEC highlights proven measures that countries may take to speed up climate technology action. The TEC also produces policy briefs, called TEC Briefs, and other technical documents to enhance information sharing on climate technology efforts. Read the TEC key messages on climate technology policies. As part of its work, the TEC has also analysed policy options emanating from the UNFCCC technical examination process.

Mandate

To enhance climate technology development and transfer, the TEC has the following functions:

- Provide an overview of countries' climate technology needs and analyse policy and technical issues related to climate technology development and transfer
- Recommend actions to promote climate technology development and transfer
- Recommend guidance on climate technology policies and programmes
- Promote and facilitate collaboration between climate technology stakeholders
- Recommend actions to address barriers to climate technology development and transfer
- Seek cooperation with climate technology stakeholders and promote coherence across technology activities

- Catalyse the development and use of climate technology road maps and action plans

To support the TEC to conduct its functions, the TEC has modalities of work. In addition, the COP mandated the TEC to further implement the technology transfer framework. In the context of serving the Paris Agreement, the TEC will also work to accelerate technology cooperation to achieve the Agreement's objectives, particularly the agreement's Article 10 on technology.

Technology Executive Committee

The Technology Executive Committee together with the Climate. Technology Centre and Network, consistent with their respective functions, is mandated to facilitate the effective implementation of the Technology Mechanism, under the guidance of the COP. The Technology Mechanism was established to facilitate the implementation of enhanced action on technology development and transfer to support action on mitigation and adaptation in order to achieve the full implementation of the Convention. Functions of the TEC

- (a) Provide an overview of technological needs and analysis of policy and technical issues related to the development and transfer of technologies for mitigation and adaptation;
- (b) Consider and recommend actions to promote technology development and transfer, in order to accelerate action on mitigation and adaptation;
- (c) Recommend guidance on policies and programme priorities related to technology development and transfer with special consideration given to the least developed country Parties;
- (d) Promote and facilitate collaboration on the development and transfer of technologies for mitigation and adaptation between governments, the private sector, non-profit organizations, and academic and research communities;
- (e) Recommend actions to address the barriers to technology development and transfer in order to enable enhanced action on mitigation and adaptation;
- (f) Seek cooperation with relevant international technology initiatives, stakeholders and organizations, and promote coherence and cooperation across technology activities, including activities under and outside of the Convention;
- (g) Catalyse the development and use of technology road maps or action plans at the international, regional and national levels through cooperation between relevant stakeholders, particularly governments and relevant organizations or bodies, including the development of best practice guidelines as facilitative tools for action on mitigation and adaptation.

7. Questions to be Addressed

1. What kind of harms and benefits might artificial intelligence use have in the future?
2. What policies ought to be put in place by nations to promote artificial intelligence use?
3. What ethical, societal, and financial considerations should nations make when creating artificial intelligence products?

4. What steps may be taken to limit the application and dissemination of AI?
5. Should humans be in charge of managing artificial intelligence?
6. What can nations do to fortify and improve the dependability of AI control systems?
7. How should this procedure be handled in the event that artificial intelligence systems encounter any issues?

8. Further Reading

UK government failing to list use of AI on mandatory register

Technology secretary admits Whitehall departments are not being transparent over way they use AI and algorithms

Not a single Whitehall department has registered the use of artificial intelligence systems since the government said **it would become mandatory**, prompting warnings that the public sector is “flying blind” about the deployment of algorithmic technology affecting millions of lives.

AI is already being used by government to inform decisions on everything from benefit payments to immigration enforcement, and records show public bodies have awarded dozens of contracts for AI and algorithmic services. A contract for facial recognition software, worth up to £20m, was put up for grabs last week by a police procurement body set up by the **Home Office**, reigniting concerns about “mass biometric surveillance”.

But details of only nine algorithmic systems have so far been submitted to a **public register**, with none of a growing number of AI programs used in the welfare system, by the Home Office or by the police among them. The dearth of information comes despite the government announcing in February this year that the use of the AI register would now be “a requirement for all government departments”.

Experts **have warned** that if adopted uncritically, AI brings potential for harms, with recent prominent examples of IT systems not working as intended including the Post Office’s Horizon software. AI in use within Whitehall ranges from Microsoft’s Copilot system, which is being widely trialled, to automated fraud and error checks in the benefits system. One recent AI contract notice issued by the Department for Work and Pensions (DWP) described

“a mushrooming of interest within DWP, which mirrors that of wider government and society”.

Peter Kyle, the secretary of state for science and technology, has admitted the public sector “hasn’t taken seriously enough the need to be transparent in the way that the government uses algorithms”.

Asked about the lack of transparency, Kyle told the Guardian: “I accept that if the government is using algorithms on behalf of the public, the public have a right to know. The public needs to feel that algorithms are there to serve them and not the other way around. The only way to do that is to be transparent about their use.”

Big Brother Watch, a privacy rights campaign group, said the emergence of the police facial recognition contract, despite **MPs warning** of a lack of legislation to regulate its use, was “yet another example of the lack of transparency from government over the use of AI tech.”

“The secretive use of AI and algorithms to impact people’s lives puts everyone’s data rights at risk. Government departments must be open and honest about how they use this tech,” said Madeleine Stone, chief advocacy officer.

The Home Office declined to comment.

The Ada Lovelace Institute recently warned that AI systems might appear to reduce administrative burdens, “but can severely damage public trust and reduce public benefit if the predictions or outcomes they produce are discriminatory, harmful or simply ineffective”.

Imogen Parker, an associate director at the data and AI research body, said: “Lack of transparency isn’t just keeping the public in the dark, it also means the public sector is flying blind in its adoption of AI. Failing to publish algorithmic transparency records is limiting the public sector’s ability to determine whether these tools work, learn from what doesn’t, and monitor the different social impacts of these tools.”

Only three algorithms have been recorded on the national register since the end of 2022. They are a system used by the Cabinet Office to identify digital records of long-term historical value, an AI-powered camera being used to analyse pedestrian crossings in Cambridge, and a system to analyse patient reviews of **NHS** services.

But since February there have been 164 contracts with public bodies that mention AI, according to Tussell, a firm that monitors public contracts. Tech companies including Microsoft and Meta are vigorously promoting their AI systems across government. Google Cloud funded a recent report that claimed greater deployment of generative AI could free up

to £38bn annually across the public sector by 2030. Kyle called it “a powerful reminder of how generative AI can be revolutionary for government services”.

Not all the latest public sector AI involves data about members of the public. One £7m contract with Derby city council is described as “Transforming the Council Using AI Technology” and a £4.5m contract with the department for education is to “improve the performance of AI for education”.

A spokesperson for the department of science and technology confirmed the transparency standard “is now mandatory for all departments” and said “a number of records [are] due to be published shortly”.

Where is the government already using AI?

The Department for Work and Pensions has been using generative AI to read more than 20,000 documents a day to “understand and summarise correspondence” after which the full information is then shared with officials for decision-making. It has automated systems for detecting fraud and error in universal credit claims, and AI assists agents working on personal independence payment claims by summarising evidence. This autumn, DWP started deploying basic AI tools in jobcentres, allowing work coaches to ask questions about universal credit guidance in an attempt to improve the effectiveness of conversations with jobseekers.

The Home Office deploys an AI-powered immigration enforcement system, which critics call a “robo-caseworker”. An algorithm is involved in shaping decisions, including returning people to their home countries. The government describes it as a “rules-based” rather than AI system, as it does not involve machine-learning from data. It says it brings efficiencies by prioritising work, but that a human remains responsible for each decision. The system is being used amid a rising caseload of asylum seekers who are subject to removal action, now at about 41,000 people.

Several police forces use facial recognition software to track down suspected criminals with the help of artificial intelligence. These have included the Metropolitan police, South Wales police and Essex police. Critics have warned that such software “will transform the streets of Britain into hi-tech police line-ups”, but supporters say it catches criminal suspects and the data of innocent passersby is not stored.

NHS England has a £330m contract with Palantir to create a huge new data platform. The deal with the US company that builds AI-enabled digital infrastructure and is led by Donald Trump backer Peter Thiel has sparked concerns about patient privacy, although Palantir says its customers retain full control of the data.

An AI chatbot is being trialled to help people navigate the sprawling gov.uk government website. It has been built by the government’s digital service using OpenAI’s ChatGPT

technology. **Redbox**, another AI chatbot for use by civil servants in Downing Street and other government departments, has also been deployed to allow officials to quickly delve into secure government papers and get rapid summaries and tailored briefings.

This article was amended on 28 November 2024 to clarify that the figure of “up to £38bn”, which a recent report claimed could be freed up across the public sector by 2030, is an annual figure.

<https://www.theguardian.com/technology/2024/nov/28/uk-government-failing-to-list-use-of-ai-on-mandatory-register>

9. Bibliography

https://en.wikipedia.org/wiki/First_emergency_special_session_of_the_United_Nations_General_Assembly

<https://www.ungeneva.org/en/news-media/news/2023/10/86818/explainer-what-un-general-assembly-emergency-special-session-and-why>

<https://journals.law.harvard.edu/ilj/2022/04/uniting-for-peace-the-emergency-special-session-on-ukraine/>

<https://community.trustcloud.ai/docs/grc-launchpad/grc-101/risk-management/risks-and-consequences-of-irresponsible-ai-in-organizations-the-hidden-dangers>

<https://www.theguardian.com/technology/2024/nov/28/uk-government-failing-to-list-use-of-ai-on-mandatory-register>

<https://unctad.org/topic/commission-on-science-and-technology-for-development/about>

<https://unfccc.int/ttclear/tec>

https://unfccc.int/ttclear/misc_/StaticFiles/gnwoerk_static/TEC_documents/5be1bf880cc34d52a4315206d54a711b/b490336895f44125a41eb9326c739986.pdf

<https://unctad.org/topic/commission-on-science-and-technology-for-development>

