LIBERA UNIVERSITÀ MARIA SS. ASSUNTA

DIPARTIMENTO DI GIURISPRUDENZA, ECONOMIA, POLITICA E LINGUE MODERNE

CORSO DI LAUREA MAGISTRALE IN RELAZIONI INTERNAZIONALI
CLASSE INTERNATIONAL RELATIONS IN THE DIGITAL ERA

CATTEDRA DI

DIGITAL DIPLOMACY: NEW MEDIA AND NEW TECHNOLOGIES IN INTERNATIONAL RELATIONS

Titolo Tesi

Artificial Intelligence for Diplomacy: innovation, ethics, and governance challenges

Intelligenza artificiale nella diplomazia: innovazione, etica e sfide di governance

RELATORE
Prof. Ignazio Ingrao

CORRELATORE
Prof. Raniero Romagoli

CANDIDATO
Vittoria Rose Valerio
Matricola 33865/410

ANNO ACCADEMICO 2025-2026

Alla mia famiglia

parte centrale della mia vita sempre pronti

a combattere con me e per me.

To Grant my number 1 supporter

and my favorite person on the heart.

A me stessa per aver sempre

combattuto e non essermi

mai arresa davanti ai no.

Artificial Intelligence for Diplomacy: innovation, ethics, and governance challenges

Index

Introduction ***1***

Chapter 1 — From Digital Diplomacy to AI-Driven Diplomacy **2**
- ***1.1 The Evolution of Digital Diplomacy*** 2
- ***1.2 The Emergence of Artificial Intelligence in Diplomatic Practice*** 7
- ***1.3 AI as a Decision-Support Tool for Diplomats*** 8
- ***1.4 Open Sources and the Transformation of Diplomatic Information Gathering*** 10
- ***1.5 Limits and Constraints of AI Integration in Diplomacy*** 23

Chapter 2 — Artificial Intelligence and Power in Contemporary Diplomacy **30**
- ***2.1 Data and Informational Power in International Relations*** 30
- ***2.2 AI and Strategic Competition Among States*** 36
- ***2.3 Implications for Diplomatic Influence and Negotiation*** 45

Chapter 3 — Ethical and Political Challenges of AI in Diplomacy **48**
- ***3.1 Bias, Opacity, and Algorithmic Risks*** 48
- ***3.2 Accountability and Responsibility in AI-Assisted Decision-Making*** 51
- ***3.3 AI, Human Rights, and Democratic Legitimacy*** 54

Chapter 4 — Governing AI in Diplomacy: The European Approach in a Global Context **62**
- ***4.1 The European Union's AI Governance Framework (AI Act, GDPR)*** 62
- ***4.2 Comparing Governance Models (EU, US, China)*** 74
- ***4.3 Limits of Global AI Governance and Multilateral Coordination*** 87
- ***4.4 Towards AI-Augmented Diplomacy*** 95
- ***4.5 Sustainability Challenges in AI Governance*** 109

Conclusion ***129***

Bibliography ***133***

Introduction

The world of diplomacy is experiencing a radical change in the present-day international system due to digital technologies. Unlike in the past, where the diplomatic agenda was full of traditional security issues, trade, and interstate political relations, in the new times, there are new policy arenas that have thrown the traditional paradigms in international relations. Artificial intelligence governance (AI governance) is one of these problems, which are increasingly being discussed at multilateral negotiating tables because of the growing role of cyberspace and new technologies in the political, economic, and social life of countries[1].

Artificial intelligence is transforming the way diplomatic activities are conducted, influencing decision-making processes, strategic communication, and policy formulation. Among these emerging technologies, artificial intelligence has rapidly become one of the most significant drivers of change in diplomatic practice and international governance.[2]

Bringing artificial intelligence into diplomacy raises significant questions about how algorithms influence decisions, the ethical risks involved, and what this means for global governance. It is not only the technical aspects of these new areas that are strategic in relevance, since they are involved with national security, human rights, economic competitiveness, and supranational legal orders.

These dynamics need to have coherent international agreements that will balance between rights protection, economic development, and collective security[3].

Simultaneously, artificial intelligence is both a threat and an opportunity for modern diplomacy. AI can not only be discussed in the context of negotiation but also serves as an instrument in the process of international policy formulation, analysis, and implementation, forming an overlap of technological governance and international relations.

This thesis aims to analyze the role of AI in contemporary diplomacy, examining both its operational impact on diplomatic practice and the key ethical and governance challenges that arise from it, and understanding the change from digital diplomacy to AI-driven diplomacy.

---

[1] DiploFoundation. (2025). *Digital diplomacy in 2025: Geopolitics, new topics & tools*. Recuperato da https://www.diplomacy.edu/topics/digital-diplomacy/

[2] Horowitz, Michael C. **"**Artificial Intelligence, International Competition, and the Balance of Power." *Texas National Security Review* 1, no. 3 (2018): 37–57. https://muse.jhu.edu/pub/15/article/964758/pdf

[3] DiploFoundation. (2025).

Chapter 1 — From Digital Diplomacy to AI-Driven Diplomacy

1.1 The Evolution of Digital Diplomacy

Historically, the concept of diplomacy has been perceived as the art and practice of international relations, which relies mainly on interactions between sovereign states through bilateral and multilateral negotiations and conflict-prevention and resolution mechanisms. The primary goals of diplomacy, according to this classical vision, were the safety of national security, the advancement of economic interests, and the preservation of international stability by formal and institutionalized tools. In a broader sense, diplomacy may be seen as institutionalized communication in which legitimate actors regulate relationships and generate public goods within the international system[4].

Diplomatic activity had long existed in the context of a system of direct relations between state representatives, clearly defined formal procedures, and a system highly focused on the role of states. This model started to gradually change in the 20th century, as mass media disseminated, and, later, with the growth of digital technologies. Radio, television, and most importantly, the Internet have contributed to the reduction in time between the event that transpired within an international setting and the political reaction to it, which has contributed to the creation of a more responsive and quicker decision-making process. In this sense, diplomacy has progressively evolved to be in an informational space defined by immediacy and continual updates, leading to what is being described as a kind of real-time diplomacy, where decision-makers are under growing pressure to respond to global events in real-time[5].

The emergence of information and communication technologies in the environment of changes introduced by the process of digitalization has increasingly moved the boundaries of the functioning of diplomacy, both geographical and institutional. Such a process has given rise to the so-called digital diplomacy, which is viewed as a collection of practices, according to which international actors employ digital technology, such as social media, online platforms, and sophisticated communication systems, to both achieve foreign policy goals and engage with

[4] Corneliu Bjola e Markus Kornprobst, *Understanding International Diplomacy: Theory, Practice and Ethics*, 2ª ed. (London: Routledge, 2018), 3–8.

[5] Philip Seib, *Real-Time Diplomacy: Politics and Power in the Social Media Era* (New York: Palgrave Macmillan, 2012), 1–10.

diverse state and non-state actors[6]. This is not merely an update of communication channels, but a more fundamental change of the logic and methodology by which international policies are made, negotiated, and implemented in a world that is highly interconnected and where information flows very fast[7].

Social media has been a major facilitator of this change. They're not only new communication avenues but a new distribution dynamic between diplomacy and traditional media to address international public opinion. Foreign ministries and embassies can use these online communication channels to engage directly with audiences online, with foreign nationals, the media, opinion makers, and other diplomatic circles, while enhancing their output and message through the power of these media, and all while foreign ministries and embassies are increasingly stepping into the role of journalists and increasingly communicating with media, and finally by adapting their communication style to the media logic of social networks and digital platforms.[8]

Moreover, social networking affects diplomacy in the sense that it facilitates the swiftness of diplomatic communication. In this regard, diplomatic communication is forced to comment almost in real time on events and respond to possible opposite interpretations, and to quickly generate an interpretive frame before other actors forge another reading of the facts.[9] Despite this shift, new vulnerabilities emerge as states use social media to build narratives, thus legitimizing their policies, delegitimizing their opponents, and shaping the international perception of international events.[10] The same platforms that provide more direct lines of communication can also lead to disinformation, foreign information manipulation, and deep polarization of public conversations. Social media, indeed, is perceived by the European External Action Service as an arena where foreign policy and security of the European Union is contested and threatened by a campaign of

[6] Corneliu Bjola e Ilan Manor, eds., *The Oxford Handbook of Digital Diplomacy* (Oxford: Oxford University Press, 2024), pp,19–20.
[7] Ivi, pp,18–22.
[8] Ilan Manor and Rhys Crilley, "The Mediatisation of MFAs: Diplomacy in the New Media Ecology," *The Hague Journal of Diplomacy* 15, no. 1 (2020): 66–92, pp. 67–68, https://doi.org/10.1163/1871191X-15101051. Manor and Crilley, "The Mediatisation of MFAs," pp. 68–70, 72–74.
[9] Manor and Crilley, "The Mediatisation of MFAs," pp. 71–74.
[10] Manor and Crilley, "The Mediatisation of MFAs," pp. 75–82; Ilan Manor and Zhao Alexandre Huang, "Digitalization of Public Diplomacy: Concepts, Trends, and Challenges," *Communication and the Public* 7, no. 4 (2022): 167–175, pp. 168–170, https://doi.org/10.1177/20570473221138401

Foreign Information Manipulation and Interference, which can polarize societies, twist democratic processes, and impede the achievement of the European Union and its partners' political agenda.[11] The COVID-19 pandemic was perhaps the final impetus that launched digital diplomacy as a force to reckon with. The digital world was given a kick in the pants during the COVID-19 pandemic. Many foreign ministries, embassies, and international organizations have adapted traditional diplomatic events to a digital format, due to international mobility restrictions, border closures, and the lack of possibility to meet in person. In this regard, COVID-19 wasn't just a new set of tools; it was a reminder of the importance of digital communication networks in modern-day diplomacy. Manor and Pamment note that the pandemic had a profound impact on social, cultural, and political life, at the same time as it made the world "less international" in terms of "physical mobility" and made it more reliant upon "digital connections, forms of public communication, and crisis management.[12]

The first impression was in public and consular communication. As the pandemic developed in the early days, many foreign ministries took to social media not only to communicate with foreign but also their own citizens about travel, restrictions, repatriation, consular assistance, and international cooperation. In addition, domestic digital diplomacy played both two-way and explanatory roles, as described by Manor and Yarchi, with the Federal Foreign Office of Germany taking on communicative roles towards the public during the crisis to help them understand it: a) finding out how big the crisis really is and b) informing the public about repatriations, international relief, cooperation within Europe, and multilateral responsibilities.[13]

Bjola's analysis of the German case, for instance, reveals how domestic digital diplomacy took the shape of both an interpretive and an operational function, as it was enacted towards both the domestic public and foreign authorities at both European and global levels, because of the uncertainty and fear permeating the global crisis and because regulations changed so rapidly.[14] The second effect was related to the operation of multilateral diplomacy and international

[11] European External Action Service, "Information Integrity and Countering Foreign Information Manipulation & Interference (FIMI)," 17 marzo 2026, https://www.eeas.europa.eu/eeas/tackling-disinformation-foreign-information-manipulation-interference_en?s=2803.

[12] Ilan Manor and James Pamment, "At a Crossroads: Examining Covid-19's Impact on Public and Digital Diplomacy," *Place Branding and Public Diplomacy* 18 (2022): 1–3, pp. 1–2, https://doi.org/10.1057/s41254-021-00249-9

[13] Ilan Manor and Moran Yarchi, "From the Global to the Local and Back Again: MFAs' Digital Communications During COVID-19," *International Journal of Communication* 17 (2023): 860–881, pp. 860–862, 866–877.

[14] Corneliu Bjola, "Digital Diplomacy as World Disclosure: The Case of the COVID-19 Pandemic," *Place Branding and Public Diplomacy* 18 (2022): 22–25, pp. 22–24, https://doi.org/10.1057/s41254-021-00242-2.

negotiations. COVID-19 caused the emergence of 'virtual venues '; digital venues that were hosted on platforms like Zoom, Microsoft Teams, Webex, etc. COVID-19 resulted in the growth of digital venues, conferencing venues called 'virtual venues ' on platforms such as Zoom or Microsoft Teams or Webex, or others.[15] As such these settings enabled meetings, summits, and negotiations to proceed even in the face of health conditions and restrictions, but they also had an impact on the nature of diplomacy. Virtual venues are not just used to adjust the format of negotiations but can also affect the ability to "read the room", build trust, employ negotiation tactics, and assess the legitimacy of the negotiation process.[16] One example is the United Nations Security Council, which, between March 12 and July 14, 2020, held no in-person meetings due to the WHO's pandemic declaration and the closure of UN headquarters, and instead relied on video-conferencing to carry out its work.[17] But virtualization of diplomacy also exposed some shortcomings. Bramsen and Hagemann point out that the lockdown and the accompanying shift to virtual meetings does not herald a 'no more diplomacy' mentality because it can be argued that the first few aspects of diplomatic negotiation mentioned above are inherently lost when it is conducted virtually.[18] This is why COVID-19 cannot be seen as ending diplomacy in favor of a digital alternative. Rather, it has shown how a hybrid diplomatic approach can be adopted, a combination of digital and on-site means. Thus, the pandemic hastened the evolution of diplomacy, highlighting the fact that digital technologies can maintain the continuity, speed, and inclusion of diplomacy but do not subtract the value of being physically present, of having whatever is known in diplomacy, "the soft polish," or of having the opportunity for informal conversation.

Theoretically, the technological disruption can be viewed as a process of change, which is implemented in three key dynamics: upgrading, the introduction of digital technologies into traditional diplomacy; augmenting, the expansion of diplomacy into the new digital space; and rewiring, the deeper restructuring of diplomacy as a hybrid system involving online and offline

[15] Ivi.
[16] Corneliu Bjola and Michaela Coplen, "Virtual Venues and International Negotiations: Lessons from the COVID-19 Pandemic," *International Negotiation* 27 (2022): 1–25, pp. 1–5, https://doi.org/10.1163/15718069-bja10060
[17] United Nations Security Council, "Highlights of Security Council Practice 2020," https://main.un.org/securitycouncil/en/content/highlights-2020
[18] Isabel Bramsen and Anine Hagemann, "The Missing Sense of Peace: Diplomatic Approachment and Virtualization during the COVID-19 Lockdown," *International Affairs* 97, no. 2 (2021): 539–560, pp. 539–546, https://doi.org/10.1093/ia/iiaa229

dimensions[19]. It has been through this that the tools of diplomacy have been redefined, as well as the actors of diplomacy, its logic, and its modes of interaction as well.

In the context of the changes that the process of digitalization introduces, the notion of cyber diplomacy has been slowly becoming more popular; it can be defined as the complex of diplomatic activities that are directed at the management and regulation of cyberspace. Diplomatic action in this respect is the definition of norms and behavior in using digital technologies, including such problems as cybersecurity, technical collaboration between states, and the prevention of conflicts in the digital context. In this respect, cyberspace ceases to be considered only a useful means of diplomacy but gradually begins to assume the features of a full-fledged sovereign realm of international relations, possessing its own dynamics and logic of work[20].

One of the most important issues related to the change of modern diplomacy is the increase in the scope of participation in the decision-making and negotiation process. Non-state actors, such as technology companies, international organizations, technical communities, and actors of civil society, are assuming an even more important role alongside the states[21]. These actors are actively involved in the formation of the practices and rules of the digital environment, which contributes to the change of the traditional dynamics of the international system.

In that regard, a multi-stakeholder model of governance has been slowly developing, where decision-making is shared among a multiplicity of actors, where states are no longer the sole custodians of digital infrastructure or of the principles of regulation over cyberspace[22]. Although the changes that digitalization has introduced are radical, transitioning to digital diplomacy does not imply that the current practice will be fully substituted. Instead, the modern-day diplomacy may be seen as a hybrid framework, where the old tools, subjects and operational patterns remain and interact with emerging forms of interaction that have been facilitated by digital technologies. In this sense, 21st-century diplomacy retains a strong relationship with its past and continues to evolve to meet the demands of an international environment that is becoming increasingly

[19] Ivi, pp. 19–21

[20] Fatima Roumate, ed., *Artificial Intelligence and Digital Diplomacy: Challenges and Opportunities* (Cham: Springer, 2021), 1–5.

[21] Corneliu Bjola e Ilan Manor, eds., *The Oxford Handbook of Digital Diplomacy* (Oxford: Oxford University Press, 2024), 19.

[22] Corneliu Bjola e Ilan Manor, eds., *The Oxford Handbook of Digital Diplomacy* (Oxford: Oxford University Press, 2024), pp, 19–21.

dominated by the centrality of digital technology and the increasing interconnection among international actors.[23]

1.2 The Emergence of Artificial Intelligence in Diplomatic Practice

Artificial intelligence (AI) has started to have a more prominent role in the political and diplomatic processes in the context of the digital transformation of international relations. In fact, the development of digital technologies has brought about the emergence of new modes of information management and decision support, gradually affecting the dynamics on the basis of which international actors perceive and interpret the global reality. In this respect, AI can be considered a new instrument that is contributing to redefining, albeit slowly, decision-making in the international system.[24] The incorporation of artificial intelligence is also tightly connected with the increased accessibility of data and the possibility to process it fast and effectively. Using AI-based systems, it is possible to implement the processing of large volumes of information, which improves the skills of policymakers to track the complex situation and determine the possible future evolution of the situation. The given application belongs to a larger system where AI performs the major role of analytics, contributing to the increase in the knowledge of international phenomena, but does not deprive the human role in decision-making.[25]

At the same time, the spread of artificial intelligence is part of the broader process of the digitization of diplomacy, which has led to the growing use of technological tools by diplomatic institutions. In this regard, digital technologies and intelligent systems make it possible to increase the potential of gathering, systematizing, and analyzing information, thus supporting the work of decision-makers of the population and enhancing the efficiency of diplomatic actions.[26]

The next important one is the influence of artificial intelligence on the frameworks and competencies necessary in the diplomatic institutions. In fact, the increased role of digital technologies has caused most governments to incorporate new technical skills in their

[23] Ivi, pp, 20–21.
[24] M. L. Cummings et al., *Artificial Intelligence and International Affairs: Disruption Anticipated* (London: Chatham House, 2018), pp. 1–2.
[25] Ivi, pp. 4–5.
[26] Fatima Roumate, ed., *Artificial Intelligence and Digital Diplomacy: Challenges and Opportunities* (Cham: Springer, 2021), pp, 1–3.

administrations, which explains why efforts should be made to adjust diplomatic practices in an environment where technological innovation is increasingly taking center stage.[27]

Nevertheless, it is crucial to point out that the existing uses of artificial intelligence are still confined to specific tasks and do not allow them to act independently at a level that is like the performance of humans. The current technologies are also defined by their narrow application to specific tasks, including the analysis of data or decision support, and are not types of general intelligence. Therefore, AI is not a replacement for human judgment in diplomacy but should be seen as a support system.[28]

To conclude, the advent of artificial intelligence in diplomacy is an important phenomenon in the overall process of digital transformation of international relations. Although it provides new opportunities in information analysis and management, it demands institutions and skills to constantly vary, yet keep the human intervention as the central point of the decision-making process.[29]

1.3 AI as a Decision-Support Tool for Diplomats

Artificial intelligence is becoming one of the most important technologies to rely on when it comes to assisting the decision-making process of the diplomatic actors in the context of the digital transformation of international relations. Instead of substituting human judgment, AI-based technologies contribute to the enlargement of the analytical potential of decision-makers, which impacts the way the information is gathered, processed, and utilized in the development of international strategies. AI, in this regard, is to be considered as an enabling technology, which can change the situation in decision-making without displacing human action.[30]

Among the key applications of artificial intelligence in the diplomatic domain is its capability to process massive data volumes and detect significant patterns in more complicated settings. These tools enable one to automate repetitive information analytical processes and enhance efficiency in information management, and the decision-maker can concentrate on the higher-level strategic

[27] Ivi, pp, 6–8.
[28] Cummings et al., *Artificial Intelligence and International Affairs*, pp, 2–3.
[29] Ivi, pp, 5–6.
[30] Michael C. Horowitz, "Artificial Intelligence, International Competition, and the Balance of Power," *Texas National Security Review* 1, no. 3 (2018): p, 38.

analysis. In this regard, AI aids in redefining the process of decision-making, which has become more informed and can possibly be fast, although it is still subject to human intervention.[31]

Besides its analytical purpose, artificial intelligence can also be predictive and help to evaluate potential progress in international situations. The level of uncertainty, which defines the international environment, can be minimized with the assistance of AI systems, as they can give insights into the implications of some policy decisions. Nonetheless, they do not generate autonomous decisions, but, instead, provide probabilistic assistance that should be interpreted and assessed by human decision-makers.[32]

These capabilities are also causing an ever-growing adoption of artificial intelligence as a working tool to facilitate institutional operations in the diplomatic sphere. As mentioned in the literature, AI may be utilized in different fields of diplomatic practice, such as consular services, negotiations, crisis management, and diplomatic mission security, and it can enhance the efficiency and quality of decision-making. In that way, AI can be perceived as a kind of toolbox that can assist diplomats, as it can offer tangible assistance to their everyday work.[33]

Simultaneously, the application of artificial intelligence can be placed in a larger framework of change in the perception of world reality by international actors. Digital technologies and intelligent systems not only transform the perception of the world and the ways decision-makers can interpret it but also assist in redefining the cognitive processes underlying political decisions. In this respect, AI not only provides technical means but also affects the process of knowledge building and strategic evaluation.[34]

The other crucial aspect is that the effects of AI are not only determined by its technological nature, but also by how it is implemented and incorporated by state actors. The advantages of the use of AI, as observed in the literature, are directly connected with the capacity of states to establish suitable expertise and integrate these technologies into their decision-making and strategic frameworks. As a result, AI also contributes to reconsidering power relationships in the

[31] M. L. Cummings et al., *Artificial Intelligence and International Affairs: Disruption Anticipated* (London: Chatham House, 2018), pp, 2–3.
[32] Ivi, p, 3.
[33] Karim Hamidouche, "Artificial Intelligence: A New Tool for Diplomats," in *Artificial Intelligence and Digital Diplomacy*, ed. Fatima Roumate (Cham: Springer, 2021), p, 25.
[34] Cummings et al., *Artificial Intelligence and International Affairs*, p, 3.

international system, affecting the competition between states and the allocation of strategic resources.[35]

Although it has potential, it is important to note that, at this point, artificial intelligence cannot substitute for human judgment in making diplomatic decisions. The current technologies have fixed boundaries and are subject to human interpretation, particularly in situations where there is great political, cultural, and ethical complexity. In that way, the role of AI should be perceived as supplementary to the role of decision-makers, with human responsibility being the core in decision-making related to international relations.[36]

To conclude, artificial intelligence is one of the primary aspects of the modern development of diplomacy as a decision-support system that can enhance the analysis, prediction, and control of information. However, it should remain strictly subject to human oversight, highlighting the need to maintain a balance between technological innovation and political responsibility.[37]

1.4 Open Sources and the Transformation of Diplomatic Information Gathering

The gathering, decoding, and relay of information has always been a fundamental facet of diplomacy. As before the negotiations and before it needs to be represented or communicated in public, diplomatic activity requires that one attune oneself to external contexts; that one understands what is driving the political scene; that one interprets signals and deliver a dependable sign-out to decision-makers of the international situation. This intelligence function remained mainly through secret methods, such as diplomatic documents, reports from embassies, individual networks, confidential sources, local contacts, and classified intelligence, for centuries. This balance, however, has been changing in recent years as more information has become public.

One concept that has come to the fore in this context is open-source intelligence, or OSINT, as it's often abbreviated. OSINT is more than reading the newspaper or perusing websites or social media posts. A definition that is too general could lead to confusion with routine world observation and information gathering for intelligence research, as Ludo Block notes. Therefore, Block suggests that OSINT should be understood as the process of systematic information gathering and

[35] Horowitz, "Artificial Intelligence," p, 41.
[36] Cummings et al., *Artificial Intelligence and International Affairs*, pp, 2–3.
[37] Hamidouche, "Artificial Intelligence," 25; Horowitz, "Artificial Intelligence," p, 41.

processing from open sources to fulfill a given information or intelligence need.[38] This is a useful elucidation of the concept of diplomacy for a study of modern diplomacy, since it enables us to distinguish between information made available and information that can be turned into something which could be of strategic value.

In fact, the critical part is from a source to a function. A source is only public because it's public, not necessarily intelligence-producing. A video posted online that a commercial satellite image, a Telegram post, a statement from the official website, a digital map, or an open database is not evident to diplomacy until it has been gathered, checked, and put into perspective for a decision-making process. Indeed, OSINT is not just a listing of sources; it is a method. It needs to be directed, collected, processed, analyzed, disseminated, and fed back upon all aspects of the intelligence cycle. Based on the definition that comes from the Office of the Director of National Intelligence, Böhm and Lolagar[39] define OSINT as intelligence that comes from publicly available sources, is collected, analyzed, and disseminated to an appropriate audience in the right time and manner of the appropriate audience to satisfy an intelligence requirement.[40]

This point feeds into diplomacy as it has come to be important in the wake of OSINT. A diplomat plays many roles beyond that of a political messenger; he or she is an interpreter of information. Historically, only being on location or the capacity of the state to collect information via official or covert sources has provided access to information regarding conflict, military movements, social interaction, humanitarian crises, or propaganda[41]. Now, however, a significant amount of this information is accessible in close to real-time in the public domain. It doesn't imply that such information will be "true" or "current," but it does imply that knowledge of foreign happenings no longer has the prerogative of the State.[42]

This change is not new and has been witnessed before. There is a longer history to OSINT than what is commonly thought, according to Block. In many respects, it is the night of World War II that can be said to herald the dawn of OSINT, as it was during this period that the BBC Monitoring

---

[38] Ludo Block, "The Long History of OSINT," *Journal of Intelligence History* 23, no. 2 (2024): 97-98, https://doi.org/10.1080/16161262.2023.2224091.

[39] Isabelle Böhm and Samuel Lolagar, "Open Source Intelligence: Introduction, Legal, and Ethical Considerations," *International Cybersecurity Law Review* 2 (2021): 318, https://doi.org/10.1365/s43439-021-00042-7.

[40] Ivi.

[41] Block, 2024, pp. 97–98; Böhm and Lolagar, 2021, pp. 318–319

[42] Böhm and Lolagar, 2021, pp. 325–326; Magdalene Karalis, "Open-Source Intelligence in Ukraine: Asset or Liability?," *Chatham House*, December 16, 2022, https://www.chathamhouse.org/2022/12/open-source-intelligence-ukraine-asset-or-liability

Service in England was established (1939) and the Foreign Broadcast Monitoring Service in the United States (1941). But the systematic collection and use of open sources are well documented in the U.S., dating back to the mid-19th century, and in Europe as early as the early 20th century.[43] This is significant in that it does not allow one to conclude that OSINT is a phenomenon that arose with the advent of the Internet.[44] Contemporary innovation is not in the existence of an open source, but in the volume, speed, global reach, and ability of the open sources to be collectively verified.[45]

Pre-dating the digital era, newsprint, bulletins, radio, the press, and official publications were good sources of insights into the intentions and capabilities of other political players. Monitoring what other States made public, in or out of purpose, has always been an aspect of diplomacy. But these were still constrained by practicalities: publication dates, the movement of physical records, language barriers, archival availability, monitoring capabilities, and intermediaries when it came to access to open sources. The advent of the Internet has radically changed these conditions. From a public information perspective, the Internet has changed the nature of information; it's now more accessible, more available, quicker, more varied, Böhm and Lolagar say, and offers more people more choices in how they access it. The Internet presents a range of new forms of publicly available source materials, in addition to digitized newspapers and magazines, including personal websites, forums, online encyclopedias, newsletters, blogs, discussion groups, search engines, social media sites, business directories, and other digital services.[46] This avalanche of sources has affected directly the work of the diplomat in intelligence gathering. Confidential reports or institutional sources are no longer the only sources to which Embassies and foreign ministries must depend for information; now, there is public data flowing from citizens, the media, technology firms, NGOs, researchers, local communities, and digital platforms. A picture by a soldier, a geo-locatable video, a commercial ad, a flight-tracked movement, a commercial satellite shot, or a forum post could all be useful for comprehending an international crisis. Thus, today's diplomacy takes place in a more open information environment, one that is also becoming noisier and more intractable.

OSINT is not making diplomacy any easier. To the contrary, it makes them more complicated. When information is abundant, it's not so much about accessing information, but about selecting

[43] Block, “The Long History of OSINT,” pp. 95-96.
[44] Ivi.
[45] Block, 2024, pp. 95–98; Böhm and Lolagar, 2021, pp. 318–319
[46] Böhm and Lolagar, 2021, pp. 318–319

it, verifying, and interpreting it. Böhm and Lolagar differentiate between data, information, and intelligence. Data are facts that are gathered without explanation or analysis; information is a product of the processing, classification, and contextualization of the facts, and intelligence is a product of the integration, evaluation, and analysis of information in relation to a specific information need.[47] It's an important distinction, because it isn't data that's missing, it's 'political' knowledge. What diplomacy requires is not simply more data, but data that is verified, not only organized, but made a basis for strategic judgment.

So from this viewpoint, OSINT is a practice situated at the nexus of diplomacy, intelligence and publicity. It helps to support traditional diplomacy on one hand by offering information about current events, places, actors and dynamics, on the other hand. However, it can also have public influence, as much of the evidence collected from open sources can be publicly exhibited, shared, and discussed both in and outside of classified information environments. This is important because classified intelligence can only be used by decision makers, while OSINT can be used to affect public opinion, allies, international organizations, and the media. That is, open sources are not only changing the way that information is collected but also changing the dynamic between knowledge, credibility, and diplomatic legitimacy.

This is also acknowledged at the institutional level now and the importance of OSINT is acknowledged. The Office of the Director of National Intelligence (ODNI) and the Central Intelligence Agency (CIA) have released the Intelligence Community OSINT Strategy for 2024-2026 in March 2024. Official definition: OSINT is intelligence that comes from sources available to the public or the commercial market, and can answer specific priorities, requirements or intelligence gaps. The strategy also underscores OSINT's importance to the mission of the U.S. intelligence community, stating that "open sources are no longer viewed as a fringe capability but are becoming a key element of the modern intelligence enterprise".[48] This institutional recognition is a sign that open sources are no longer thought of as a complementary or "marginal" tool of intelligence, but as a strategy to the contemporary intelligence eco-system. These strategic priorities are also applicable to diplomacy: Coordinate collection and sharing of open-source data

[47] Ivi, p. 325–326.

[48] Office of the Director of National Intelligence and Central Intelligence Agency, "ODNI and CIA Release the Intelligence Community OSINT Strategy for 2024-2026," March 8, 2024, https://www.odni.gov/index.php/newsroom/press-releases/press-releases-2024/3784-odni-and-cia-release-the-intelligence-community-osint-strategy-for-2024-2026.

to build a sturdy foundation of data collection and sharing; Establish an integrated management of open-source collection; Promote innovation in OSINT to fill capability gaps; and train the next generation of OSINT professionals and methodologies.[49] This is to show how the so-called art of OSINT work is an organized, professional, technical, and oriented collection of information and skills. OSINT is not only any sort of 'Google-ing,' but constructing institutional capabilities to acquire, validate, translate, and incorporate information from complicated, digital environments. This process also alters the dynamics between diplomats and other information actors. In the past, only states could collect and process strategic information on an international level, but today, there are others that can also.[50] Information that relates to foreign policy can come out of investigative reporting, research by university scholars, nongovernmental organizations, private analytics companies, digital communities, or individuals such as reporters.[51] The case of Bellingcat, analyzed in second part of this part using the cases of MH17 and Syria, is a clear demonstration of this: Non-State Actors can contribute to the reconstruction of complex international events, by geolocating, verifying the images, analyzing the videos, cross-referencing public information, and crowdsourcing, but they do not supplant States, rather they participate to the reconstruction of factual truth addressed to which diplomacy is running.[52] This shifting can be seen in today's wars in particular. During a conflict, information from the open-source community can help illuminate troop movements, the destruction of military equipment, attacks on civilians or bombings, violations of international law and disinformation.[53] But they also have a diplomatic dimension: they can help bolster claims to public attribution, the difficulty of denying responsibility, the pressure for sanctions or accountability measures, and help the development of international coalitions.[54] In addition to shaping internal decision-making, analyses of the war in

[49] Ivi.
[50] Block, 2024, pp. 97–98; Böhm and Lolagar, 2021, pp. 318–319.
[51] ODNI and CIA, 2024; Karalis, 2022.
[52] Bellingcat, *MH17: The Open Source Evidence: A Bellingcat Investigation*, 2015, pp. 1–8; Bellingcat, *MH17: The Open Source Investigation Three Years Later*, 2017, pp. 1, 30–36, 72; Eliot Higgins, "The Open Source Hunt for Syria's Favourite Sarin Bomb," *Bellingcat*, April 21, 2020, https://www.bellingcat.com/news/2020/04/21/the-open-source-hunt-for-syrias-favourite-sarin-bomb/
[53] Karalis, "Open-Source Intelligence in Ukraine"; Chris Oxendine, "Ukraine: Open-Source Data Aided Response and Documents Damages and Atrocities," *Esri Blog*, June 14, 2022, https://www.esri.com/about/newsroom/blog/ukraine-open-source-intelligence
[54] Huw Dylan and Thomas J. Maguire, "Secret Intelligence and Public Diplomacy in the Ukraine War," *Survival* 64, no. 4 (2022): 33–74, https://doi.org/10.1080/00396338.2022.2103257; Bellingcat, *MH17: The Open Source Investigation Three Years Later*, p. 72; Organisation for the Prohibition of Chemical Weapons, "OPCW Releases First Report by Investigation and Identification Team," April 8, 2020, https://www.opcw.org/media-centre/news/2020/04/opcw-releases-first-report-investigation-and-identification-team.

Ukraine have shed light on the fact that systematic use of commercial satellite imagery, social media, geospatial data, and collaborative fact checking have enabled a government, international organization or investigative group to present publicly verifiable evidence, thus constructing external pressure as well as informing internal decision-making.[55] Moreover, analyses of the armed conflict in Ukraine have shown that, beyond shape-crafted decisions, a government, an international organization or an investigative group is able to present publicly verifiable evidence, thus building external pressure through the use of commercial satellite imagery, social media, geospatial data and collaborative fact-checking.[56]

This public side of OSINT is among the most inventive aspects of OSINT, which differs from traditional intelligence. Effective covert intelligence might be correct, but it frequently cannot be reported with any degree of honesty without intruding upon sources, methods, and so on.[57] With OSINT, the information is open for all to use. A geolocated photo, a backed-up video, commercial satellite tracking, or a timeline built using available public materials can be used as diplomatic leverage.[58] In this manner, OSINT takes two steps between classified knowledge and public argumentation. Diplomacy doesn't have to stop at saying it 'knows' something; it can, in some instances, show how the knowledge has been built.[59] But this openness gives rise to new risks too. More public sources do not necessarily mean that the more public truth is available. There's also real evidence in the digital world, mistakes, fabrication, propaganda, misinformation, misquotes, and more in the virtual ocean than we have noticed. In the virtual ocean, there are also real news, inaccuracies, fabrication, propaganda, misinformation, misquotes, coordinated disinformation campaigns, etc., in the digital environment.[60] Therefore, OSINT brings in its need for certain methodical skills, verification standards, and limitations awareness. The power of an open source is precisely owing to its public nature; it can be hazardous when unmonitored. Recent studies have also highlighted how OSINT ecosystems themselves can be vulnerable to interpretive errors,

[55] Karalis, 2022; Oxendine, "Ukraine: Open-Source Data Aided Response and Documents Damages and Atrocities".
[56] Oxendine, "Ukraine: Open-Source Data Aided Response and Documents Damages and Atrocities"; Dylan and Maguire, "Secret Intelligence and Public Diplomacy in the Ukraine War."
[57] Dylan and Maguire, 2022.
[58] Bellingcat, 2015, pp. 1–8; Esri, 2022.
[59] Bellingcat, 2017, pp. 30–36, 72; Dylan and Maguire, 2022.
[60] EUvsDisinfo, "About," https://euvsdisinfo.eu/about/; European External Action Service, "Information Integrity and Countering Foreign Information Manipulation & Interference," https://www.eeas.europa.eu/eeas/information-integrity-and-countering-foreign-information-manipulation-interference-fimi_en

confirmation bias, and the spread of unverified information.[61] The diplomatic task, therefore, is not to passively accept what circulates online, but to transform the abundance of information into verified and politically responsible knowledge.[62]

OSINT is, in this way, a metamorphosis of the Diplo collection. It expands the scope of sources, speeds up the observation rate, enriches the participants and makes some kinds of attribution public.[63] Concurrently, it places new demands such as verifying, situating, safeguarding sensitive information, mitigating the risks of attributing information to people, and differentiating between useful information and manipulations.[64] Part I has thus established the conceptual framework: open sources have not been around that long, but the opening of the information has affected the scale and role of information in diplomacy.

The shift from the conceptual to the real-world stage in the production of open-source information-to-diplomatic intelligence is more clearly marked in the applied than in the conceptual. The Russia-Ukraine war is but one example of how OSINT has moved beyond the realm of a tool of specialized analysts to become an integral part of the information environment that governments, international organizations, journalists, investigative communities, and the public draw upon for understanding and interpreting international crises.[65] The first significant one is Ukraine. The Russian invasion in February 2022 brought a new era in how certain military operations, and particularly the Russians, have been viewed by the public using open-source intelligence.[66] The use of intelligence that extends beyond the context of intelligence to influence the external segments of the audience, both publicly and privately, before and after the Russian invasion of Ukraine, is nothing new and is a shift in the relationship between intelligence and diplomacy, public communication, and, even more importantly, the credibility it lends to friends, the public, and the international community. OSINT has also helped to confirm this pattern. Open-source intelligence has come into the spotlight thanks to Russia's invasion of Ukraine and the "growing speed of digital technologies," believes Chatham House.[67] In some respects, the war in Ukraine can be seen as a first partial step

[61] Johannes Niu, Mila Stillman, and Anna Kruspe, "OSINT — or BULLSHINT? Exploring Open-Source Intelligence Tweets about the Russo-Ukrainian War," 2025, pp. 1–4.
[62] Böhm and Lolagar, 2021, pp. 325–326; Karalis, 2022.
[63] Block, 2024, pp. 97–98; Böhm and Lolagar, 2021, pp. 318–319; ODNI and CIA, 2024.
[64] Böhm and Lolagar, 2021, pp. 334–335; Karalis, 2022; Niu, Stillman, and Kruspe, 2025, pp. 1–4.
[65] Magdalene Karalis, "Open-Source Intelligence in Ukraine: Asset or Liability?," *Chatham House*, 16 dicembre 2022, https://www.chathamhouse.org/2022/12/open-source-intelligence-ukraine-asset-or-liability
[66] Dylan e Maguire, 2022
[67] Karalis, 2022

towards the democratization of the intelligence-gathering process, with citizens, journalists, researchers, and the OSINT community, as well as commercial satellite imagery and geolocation platforms, providing information on troop deployments, bombings, infrastructure destruction, war crimes, military asset procurement, and disinformation campaigns.[68] But this also has its drawbacks in terms of the spread of 'squalor' and populist politics. For diplomacy, it is a central challenge, as although OSINT can support the quick identification of what is being seen and thus facilitate quick public verification of some information, it can also create noise, speed up misunderstandings, and lead to the disclosure and exposure of civilians or vulnerable sources.[69] Even though a video, map, or photograph is available, it does not reduce the need; it's still necessary.[70]

Geospatial technologies have especially great importance in Ukraine. Before the invasion of Ukraine, these types of information would typically have been classified or passed only through diplomatic channels, but the situation in Ukraine has seen these technologies used to provide information not only for Ukraine to prepare for the invasion, but for the rest of the world to understand the devastation it has caused.[71]

For example, the U.S. State Department backed the Conflict Observatory, which is designed to gather, analyze and publish information around war crimes and atrocities in Ukraine. This case illustrates that OSINT can be directly utilized in accountability diplomacy, not just to raise the awareness and information of governments and public, but as a record for legal and institutional processes that might be required in the event of future civil and/or criminal proceedings.[72]

The literature produced on social media is also an indicator of the networked nature of the OSINT in the Russia-Ukraine war. Niu, Stillman, and Kruspe review the processes of analyzing OSINT on Twitter, breaking them down into the real OSINT and "BULLSHINT".[73] Their data spans from January 2022 to July 2023 and consists of almost two million tweets from around 1,040 users discussing real time military events, analyses, and misinformation related to the conflict.[74] This demonstrates not only that OSINT is no longer constrained to a single institution but is distributed,

[68] Karalis, 2022; Esri, 2022
[69] Karalis, 2022; Böhm e Lolagar, 2021, pp. 325-326.
[70] Böhm e Lolagar, 2021, pp. 325-326.
[71] Esri, 2022.
[72] Ivi.
[73] Niu, Stillman e Kruspe, 2025, pp. 1-2.
[74] Ivi.

spread over individual OSINT accounts, Telegram channels, digital communities, images, videos, in updates of the maps, explanatory threads, and joint fact-checking activities.[75]

The second case of importance for understanding change in OSINT is related to MH17. For understanding the evolution of OSINT, the case of MH17 is important. On July 17, 2014, Malaysia Airlines (MAL) flight MH17, carrying 298 passengers, was shot down in eastern Ukraine.[76] Thanks to information that is now freely accessible, valid, and verifiable, it is now possible to answer questions about the Buk launcher's journey into eastern Ukraine (who arranged and transported it, its routes, and even its physical characteristics), writes Bellingcat in its three-year anniversary report of MH17: The Open Source Investigation Three Years Later.[77]

The case of MH17 is a pertinent example of how open-source evidence can support public attribution, especially when it comes to diplomacy. Based on photographs and videos, social media posts, geolocation, satellite imagery, and witness accounts, Bellingcat was able to reconstruct the route of the Buk.[78] The quality of this investigation was strengthened not only by the volume of data retrieved but also by the ways this data was verified through reproducible means, such as comparing the images with buildings, roads, shadows, etc., road signs, geographic coordinates, and satellite images.[79]

Bellingcat also identified the Buk system as Russian. Using what damage, side marking, obscured numbers, and white symbols the investigators saw in the vehicle, they correctly determined that it was the launcher Buk 332 from the 53rd AA Missile Brigade of the Russian army[80] The 30–35 combination of the very features that are visible on 30–35 JSM are unique among the Buks photographed and filmed in Russia or Ukraine in 2013–2014.[81] After the report, the only credible evidence that led to the conclusion of the launcher responsible for shooting down MH17 was the Russian Buk 332 in the possession of the 53rd Anti-Aircraft Missile Brigade.[82]

There's no doubt that this kind of inquiry has clear diplomatic ramifications. If there is clear evidence in the public domain from which a state can be said to be responsible, attribution will be harder for the state to refute; moreover, inclusion provides a public forum for governments, official

[75] Niu, Stillman e Kruspe, 2025, pp. 1-2; Karalis, 2022.
[76] Bellingcat, 2017, p. 1.
[77] Ivi
[78] Bellingcat, 2015, pp. 1-8.
[79] Ivi.
[80] Bellingcat, 2017, p. 72.
[81] Ivi.
[82] Ivi.

investigators, journalists, and civil society to compile different analyses on accusations of state responsibility.[83] In this way, OSINT can be used as a diplomatic tool: as a supporting basis for sanctions campaigns, accountability work, to international bodies, and to the public for examples of alternative stories to state narratives.[84]

The third case involves the use of chemical weapons in Syria. As a good example, the connection between OSINT and international organizations and accountability becomes evident here. The OPCW, organization for the prohibition for chemical weapons, (2020) explains the report issued by their Investigation and Identification Team (IIT) on 24, 25 and 30 March 2017 in the city of Ltamenah was prepared after a thorough review of available information, including interviews, examination of samples and debris, victims and medical staff reports of symptoms, imagery analysis, satellite imagery data, and expert consultations.[85]

Today's international fact-finding is not a one-sided affair with official documents and diplomatic testimony; the OPCW case is proof of that. The IIT is not strictly an OSINT organization, but all the evidence outlined, including the use of samples, interviews, and, of course, specialized technical expertise, confirms that international investigations are coming to rely increasingly on the use of a wide range of evidence to interpret complex events.[86]

However, the dynamics of open-source investigation and institutional accountability can be seen further on Bellingcat's investigation of the chemical attack in Syria.[87] OSINT not only generates an alternate narrative but also engages with the work of an international organization, in this instance the OPCW IIT's findings confirm the munitions type detected by Bellingcat and is the sarin bomb used in the attacks on Al Lataminah and Khan Shaykhun in 2017.[88]

The diplomatic importance of this encounter is quite significant. When non-state actors are able to provide verifiable analyses and international actors, such as the OPCW, are able to conduct official technical investigations, there is a multi-layered type of accountability: public, investigative, and

[83] Bellingcat, 2017, pp. 30-35
[84] Bellingcat, 2017, p. 72; Dylan and Maguire, 2022
[85] Organisation for the Prohibition of Chemical Weapons, "OPCW Releases First Report by Investigation and Identification Team," 8 aprile 2020, https://www.opcw.org/media-centre/news/2020/04/opcw-releases-first-report-investigation-and-identification-team
[86] Ivi.
[87] Eliot Higgins, "The Open Source Hunt for Syria's Favourite Sarin Bomb," *Bellingcat*, 21 aprile 2020, https://www.bellingcat.com/news/2020/04/21/the-open-source-hunt-for-syrias-favourite-sarin-bomb.
[88] Higgins, 2020; OPCW, 2020.

institutional, that can be used to support condemnations, sanctions, legal actions, and other forms of international engagement: pressure campaigns.[89]

In Ukraine, MH17, and Syria, these three situations show that OSINT can perform various functions in the process of changing the nature of diplomatic intelligence gathering. In Ukraine, OSINT helps create situational awareness, public mobilization, and real-time accountability.[90] In the MH17 case, it proves to be a tool for the hindsight analysis of events and enhances public attribution.[91] All three examples share a common quality: the reconfiguration of evidence, as information becomes public and ready-made for institutional investigation of chemical weapons use, it emerges as a diplomatic tool.[92]

The beneficial features of OSINT for diplomacy also lend themselves to vulnerabilities, however. Information, although freely available, is not necessarily accurate; not wholly complete and not free of risks for its use.[93] While information is vast, it should be mediated by diplomacy to transform it into contextualized and valid, political knowledge.[94]

The Russian disinformation case is a big story. EUvsDisinfo was launched by the European External Action Service (EEA) 's East StratCom Task Force in 2015 to raise awareness among the public of the Kremlin's information operations, as well as to build an information-resilient society against information manipulation and digital disinformation.[95] EUvsDisinfo monitors media in 15 languages and gathers more than 18,000 examples of pro-Kremlin disinformation in an open-source database.[96] Leite, Razuvayevskaya, Bontcheva and Scarton are able to provide an estimate of its scale by building a multilingual data set of debunking articles from EUvsDisinfo. The dataset comprises 18,249 articles, of which 10,682 fall under the disinformation category and 7,567 are classified as legitimate, 2,946 publishers, 508 topics, 42 languages and is distributed over an eight-

[89] Ivi
[90] Karalis, "Open-Source Intelligence in Ukraine"; Oxendine, "Ukraine: Open-Source Data Aided Response and Documents Damages and Atrocities."
[91] Böhm e Lolagar, "Open Source Intelligence," pp. 325-326; Bellingcat, *MH17: The Open Source Investigation Three Years Later*, p. 72.
[92] Böhm and Lolagar, 2021, p. 325–326; Bellingcat, 2017, p. 72
[93] Böhm e Lolagar, "Open Source Intelligence," pp. 325-326, 334-335.
[94] Böhm e Lolagar, "Open Source Intelligence," pp. 325-326; Karalis, "Open-Source Intelligence in Ukraine.
[95] EUvsDisinfo, "To Challenge Russia's Ongoing Disinformation Campaigns: Eight Years of EUvsDisinfo," https://euvsdisinfo.eu/to-challenge-russias-ongoing-disinformation-campaigns-eight-years-of-euvsdisinfo/
[96] EUvsDisinfo, "About"; João A. Leite, Olesya Razuvayevskaya, Kalina Bontcheva e Carolina Scarton, "EUvsDisinfo: A Dataset for Multilingual Detection of Pro-Kremlin Disinformation in News Articles," in *Proceedings of the 33rd ACM International Conference on Information and Knowledge Management (CIKM '24)*, 2024, pp. 1-3, https://doi.org/10.1145/3627673.3679167.

and-a-half-year period (including from January 2015 to August 2023).[97] From a diplomatic perspective, it can be said that, besides the ability of recognizing false content, open source intelligence (OSINT) is essential to comprehend the narrative strategy, its targets, and how its content is amplified[98] (EEAS, Information Integrity and Countering Foreign Information Manipulation & Interference).[99] The second one is related to Foreign Information Manipulation and Interference (FIMI). The European Union has been enhancing its capacity to detect, analyze and counter FIMI since 2015, especially in Ukraine as one of the targets of Russian aggression.[100] This exemplifies the monitoring of the information space no longer being solely a communication tool, but rather a security, diplomatic and democratic process protection activity.[101]

Other risks include pseudo-OSINT. In Niu, Stillman and Kruspe's Grummanisation of our OSINT: The influence of manipulative content, they reveal how the method of OSINT can imitate the methodology of open-source intelligence and cloud the authority of the information.[102] There is pro-Russian and pro-Ukrainian content in their resource for the Russia-Ukraine war, with less pro-Ukrainian and more pro-Russian content overtly spreading more misinformation, and acknowledging that there will be times when misinformation is not an intentional act, but rather a supporter of disinformation.[103] It is important at this point as it demonstrates that even communities that have been formed to verify information are capable of being imitated, infiltrated or contaminated by distorted content. Thirdly, there is a privacy restriction and the protection of personal data. Although it's based on publicly gathered information, that's not to say using publicly gathered information is always legally and ethically sound.[104] The use of public data can be acceptable if it relates to the computer system, objects or locations, but is more complicated if the investigation involves individuals due to the restrictions on the collection and processing of personal data as well as complex ethical issues.[105] According to the authors, not all technically

[97] Leite et al., "EUvsDisinfo," pp. 1-3.
[98] Ivi, p. 3.
[99] Leite et al., "EUvsDisinfo," pp. 1-3; European External Action Service, "Information Integrity and Countering Foreign Information Manipulation & Interference."
[100] European External Action Service, "Information Integrity and Countering Foreign Information Manipulation & Interference."
[101] Niu, Stillman e Kruspe, "OSINT — or BULLSHINT?," pp.3-4.
[102] Ivi.
[103] Ivi, pp. 1-4.
[104] Ivi.
[105] Böhm e Lolagar, "Open Source Intelligence," pp. 334-335.

possible is also morally right.[106] This is also confirmed by the literature on Open Source Social Network Intelligence. Cascavilla, Beato, Burattin, Conti and Mancini demonstrated how OSINT can be used to retrieve data from social networks, data people think is private, such as their current location, where they are from, the university they went to, or their social connections.[107] This information can be used to identify civilians, refugees, journalists, dissidents or witnesses in contexts of conflict or political repression, enabling civilians to be targeted by surveillance and retaliation.[108]

The fourth constraint is related to civilian and on-ground actors' exposure. While online publication of videos, photos, posts, and geotagged messages can document infrastructure damages, military movements, and crimes in this conflict in Ukraine, it can also cause tangible harm if it is not used carefully, as it reveals the location of a military unit, shelter, hospital, humanitarian convoy, or witness.[109] For these reasons, strict standards on verification are necessary as well when addressing open source information. Diplomatically responsible OSINT collection demands triangulation, an archive of sources, documentation of methodology, evaluation of confidence level, protection of personal data, and independent review.[110] Reliability of such information grows in step with its political importance; regarding public attribution, sanctions, international proceedings, allegations of serious violations of international law, the greater the political weight, the more rigorous the verification method becomes.[111] The profession of spying on open sources thus needs new diplomatic skills. While diplomats do not have to be digital investigators or forensic analysis experts, they must be aware of the dynamics of the modern information environment: Knowing how to geo-locate a video, how to manipulate an image, how to co-ordinate an account, how to recognize a biased data set, how public information can contain personal information and how viral information often lacks reliability.[112] Thus, OSINT literacy is

[106] Ivi.
[107] Giuseppe Cascavilla, Filipe Beato, Andrea Burattin, Mauro Conti e Luigi Vincenzo Mancini, "OSSINT — Open Source Social Network Intelligence: An Efficient and Effective Way to Uncover 'Private' Information in OSN Profiles," arXiv:1611.06737, 2016, pp. 1-2, https://arxiv.org/abs/1611.06737.
[108] Böhm e Lolagar, "Open Source Intelligence," pp. 334-335; Cascavilla et al., "OSSINT," pp. 1-2.
[109] Karalis, "Open-Source Intelligence in Ukraine"; Böhm e Lolagar, "Open Source Intelligence," pp. 334-335.
[110] Böhm e Lolagar, "Open Source Intelligence," pp. 325-326, 334-335.
[111] Böhm e Lolagar, "Open Source Intelligence," pp. 325-326; Bellingcat, *MH17: The Open Source Evidence*, pp. 1-8.
[112] Böhm e Lolagar, "Open Source Intelligence," pp. 325-326, 334-335; Niu, Stillman e Kruspe, "OSINT — or BULLSHINT?," pp. 1-4.

part of today's diplomatic professionalism.[113] To end the conclusion, open sources will have an impact that justifies their adoption for the collection of diplomatic intelligence as they break down the monopoly that the state had on the production of strategic knowledge. Today information that could be important for foreign policy (FP) gathering and verification is collected from journalists, researchers, digital communities, and nongovernmental organizations (NGOs), as well as by citizens.[114] Dismantling the opacity and obfuscation surrounding the definition of critical incidents could improve transparency, accountability, and public attribution, as was the case in Ukraine, in the investigation of MH17, and in the Syrian crisis.[115] However, the limited use of GIS in these contexts contains clear dangers, including the potential for disinformation, manipulation, civilian exposure, subversion of privacy rights, interpretative errors, or political exploitation of unverified.[116] Rather, it is not the job of modern diplomacy to disallow traditional seclusion and require radical openness, but to create a hybrid capacity, one that combines not only open sources with the cooperation of classified intelligence but also diplomats' craft and qualifications as verifiers of ethical information.[117]

1.5 Limits and Constraints of AI Integration in Diplomacy

While it is true that artificial intelligence has many potential uses, its application in the field of diplomacy has a row of limitations and challenges that limit its application. Specifically, the existing technologies cannot be used in a fully autonomous mode, but must be supervised by a human factor, particularly in the context of a high level of political and strategic complexity. The truth is that AI cannot make ethical and political decisions in the same way that a human can, and therefore, constant oversight would be required during the process of decision-making.[118]

---

[113] ODNI e CIA, "ODNI and CIA Release the Intelligence Community OSINT Strategy for 2024-2026"; Böhm e Lolagar, "Open Source Intelligence," pp. 325-326.
[114] Ludo Block, "The Long History of OSINT," *Journal of Intelligence History* 23, no. 2 (2024): 97-98, https://doi.org/10.1080/16161262.2023.2224091; Böhm e Lolagar, "Open Source Intelligence," pp. 318-319; Karalis, "Open-Source Intelligence in Ukraine."
[115] Oxendine, "Ukraine: Open-Source Data Aided Response and Documents Damages and Atrocities"; Bellingcat, *MH17: The Open Source Investigation Three Years Later*, p. 72; OPCW, "OPCW Releases First Report by Investigation and Identification Team."
[116] EUvsDisinfo, "About"; European External Action Service, "Information Integrity and Countering Foreign Information Manipulation & Interference"; Böhm e Lolagar, "Open Source Intelligence," pp. 334-335.
[117] Dylan e Maguire, "Secret Intelligence and Public Diplomacy in the Ukraine War"; Böhm e Lolagar, "Open Source Intelligence," pp. 325-326
[118] M. L. Cummings et al., *Artificial Intelligence and International Affairs: Disruption Anticipated* (London: Chatham House, 2018), pp,2–3.

This restriction takes on greater clarity when evaluating some of the mental abilities associated with diplomatic decision-making. The diplomat not only is being supplied, but he must be picking out and interpreting an event and making sure he appropriately puts it in a larger political context.[119] In international crises, for example, decision makers face issues where they only have a small amount of time in which to make a series of decisions, and information is either incomplete or overwhelming because of the scarcity of information available; they might resort to "cognitive shortcuts, heuristics and non-compensatory strategies," regularly discarding all the options they deem politically unacceptable and only subsequently analyzing the remaining options.[120]

The introduction of AI can directly impact these cognitive processes. It can help to weed out information overload, organize documents, highlight patterns, and facilitate the evaluation of options on the other hand, on the other hand, it can lead to another form of cognitive dependency.[121] In the public sector, it means that there is an "automation bias", sometimes using an automated system to the exclusion of other inputs, information, or intuitive feeling, on one hand, and "selective adherence", on the other, to the suggestion coming from an automated machine, which in either scenario undermines the correct functioning of the automated decision maker.[122] In the context of diplomacy, it means that AI is not a substitute for the human judgment of the person, but an adjunctive instrument, and that the diplomat must have responsibility for his or her interpretation, awareness of ambiguity, political risk assessment and deep questioning of what he or she is receiving from the machine.

Among the primary constraints, the question of accountability and transparency is an issue. The fact that automated systems are used in decision-making processes provokes a major concern about the question of attribution of responsibility in the case of errors or harmful decisions. Specifically, some algorithms are often described as opaque (so-called black boxes), hindering insight into the

[119] Ole R. Holsti, "Cognitive Process Approaches to Decision-Making: Foreign Policy Actors Viewed Psychologically," *American Behavioral Scientist* 20, no. 1 (1976): 11–32, https://doi.org/10.1177/000276427602000103

[120] Steven B. Redd, "Key Concepts in the Poliheuristic Theory of Foreign Policy Decision Making: A Comparative Examination Using Systemist Theory," *Social Sciences* 12, no. 8 (2023): 446, pp. 1, 3–5, https://doi.org/10.3390/socsci12080446

[121] Raja Parasuraman and Victor Riley, "Humans and Automation: Use, Misuse, Disuse, Abuse," *Human Factors* 39, no. 2 (1997): 230–253, https://doi.org/10.1518/001872097778543886

[122] Saar Alon-Barkat and Madalina Busuioc, "Human–AI Interactions in Public Sector Decision Making: 'Automation Bias' and 'Selective Adherence' to Algorithmic Advice," *Journal of Public Administration Research and Theory* 33, no. 1 (2023): 153–169, pp. 154–156, https://doi.org/10.1093/jopart/muac007

criteria by which recommendations are developed and restricting the prospects for effective control by human decision-makers.[123]

Anthropic and its models, Claude, can help convey such limitations. Claude has also been made available to the US government in the three branches of government as an extension of its offerings: Anthropic has offered Claude for science, public goods, national security, and admin productivity.[124] Anthropic has progressively moved towards offering Claude as a tool for government, including solutions such as "Claude for Government" designed to provide security, FedRAMP High authorization access, government infrastructure, data controls, auditability, and encryption.[125] Additionally, Claude has been made available through Amazon Bedrock in AWS GovCloud with FedRAMP High and DoD Impact Level 4/5 access, ideally suited for governmental environments and a range of tasks, including complex document processing and analysis, information summarization, processing of sensitive government data like Controlled Unclassified Information (CUI), and analysis of large amounts of data.[126] It's a case study showing the possibilities and the boundaries of using AI in diplomacy and public administration. In one way, systems such as Claude can help the official and institutions with writing, research, translation, program creation, sourcing under enormous loads of information and analyze it. Meanwhile, concerns about the lack of accountability, security, transparency, and over-reliance grow stronger when the government uses frontier models. There's the risk it may make mistakes, hallucinate, or put out memorable wrong answers, but there's also the danger that increasingly the public institutions can be sucked into private infrastructures and public models and that they don't have full control over their results. Yet these are, again, in nature more corporate governance mechanisms to protect us from increasingly stronger models, rather than public oversight of AI's developments via democratic processes.[127]

---

[123]European Parliament Research Service, *Artificial Intelligence and Diplomacy in the EU Context* (Brussels: European Parliament, 2022), p, 8.

[124] Anthropic, "Model System Cards," https://www.anthropic.com/system-cards

[125] Anthropic, "Constitutional AI: Harmlessness from AI Feedback," 15 dicembre 2022, https://www.anthropic.com/research/constitutional-ai-harmlessness-from-ai-feedback

[126] Anthropic, "Claude in Amazon Bedrock: Approved for Use in FedRAMP High and DoD IL4/5 Workloads," 11 giugno 2025, https://www.anthropic.com/news/claude-in-amazon-bedrock-fedramp-high

[127] Anthropic, "Model System Cards," https://www.anthropic.com/system-cards; Anthropic, "Anthropic's Responsible Scaling Policy," 19 settembre 2023, https://www.anthropic.com/news/anthropics-responsible-scaling-policy

The Anthropic models and their counterparts, called Claude, are good examples in understanding these limitations. Claude has also been rolled out for use by the three branches of the US government, offering a technology for scientific research, public services, national security, and administration productivity.[128] Also, Anthropic has made Claude available via Amazon Bedrock to FedRAMP High and DoD Impact Level 4 or 5 workloads, making the eTool suitable for environments with top-tier security requirements, FedRAMP High authorization, DoD secure infrastructure, data controls, auditability, and encryption.[129]

This case is an example of the opportunities and constraints of using AI in diplomatic and public administration. However, on the other hand, systems such as Claude can help officials and institutions for writing, researching, translation and programming, document classification, and analyzing large amounts of information. Conversely, the increased usage of government frontier models does run the risk of increased concerns on issues such as accountability, security, transparency, and government over-reliance. The dangers are not only that the system might deliver errors, hallucinate, or output something that may be wrong or misguided, but that these sensitive analytical functions, ones that are too important for the government to lose, may be increasingly handed over to private models and infrastructure that the government does not fully control. It is notable, however, that Anthropic publishes "system cards" to outline the capabilities, safety evaluations, and responsible deployment decisions around its models and has implemented a "Responsible Scaling Policy" based on AI safety levels, which are used to regulate the risks of increasingly powerful models, but ultimately, it is corporate governance and not democratic public oversight.[130]

Similar concerns can be found when it comes to consular and migration services, where AI systems may seem less politically charged than in foreign policy or security contexts, but can be equally pertinent to personal rights, mobility, vulnerabilities, and access to protection. This distinction puts a spotlight on consular and immigration processes, where actions such as issuing passes or visas and granting permits, along with other administrative aspects of legal status and family reunification, refugees, and international mobility, are practical as well as deeply life-altering yet

[128] Anthropic, “Constitutional AI: Harmlessness from AI Feedback,” 15 dicembre 2022, https://www.anthropic.com/research/constitutional-ai-harmlessness-from-ai-feedback

[129] Aman Priyanshu, Yash Maurya e Zuofei Hong, “AI Governance and Accountability: An Analysis of Anthropic’s Claude,” arXiv:2407.01557, 2024, pp. 1–3, 5–8, https://arxiv.org/abs/2407.01557

[130] Anthropic, “Detecting and Countering Misuse of AI: August 2025,” 27 agosto 2025, https://www.anthropic.com/news/detecting-countering-misuse-aug-2025

highly administrative issues.[131] The Canadian case is interesting; the IRCC (Immigration, Refugees and Citizenship Canada) claims certain tools can assist it in approving portions of the process in the case of routine applications, but still not reject applications or advise against their rejection.[132] This is an example of one of the potential use cases of an AI in consular services, one that's not meant to replace the officer, but is an enabler for triage, administrative support, and workload management. But in this instance, it also illustrates its primary restriction: clear separation and distinction between administrative support and substantive decision-making. In the Canadian system, the machine-learning tools support application processing by applying preestablished rules developed by officers or machine learning, but all decisions about the refugee aspect of the application go to an officer who still must review, monitor, and oversee the decisions made by the automated tool.[133] Again, every learner must recognize that there are certain things that AI should not do, and in consular and immigration services, that includes making decisions about rights, a person's vulnerability, or a person's legal status.[134] The British case of the "visa streaming tool" is the other extreme threat. In this case, dealing with a simple and transparent system, such as the Home Office system that classifies visa applications, using a traffic-light risk system would have had significant effects in that it would cause changes in how applications were evaluated, where they would go, and if an application was to a channel that is subject to heightened scrutiny.[135] From this perspective, algorithmic triage cannot be completely apolitical. For this reason, algorithmic transparency becomes an essential requirement for digitized consular services. UK standards are vital as they enable the distinction between good algorithmic tool use, decision

[131] Katharina E. Höne et al., *Mapping the Challenges and Opportunities of Artificial Intelligence for the Conduct of Diplomacy* (Geneva: DiploFoundation, 2019), pp. 14–16, 25–26.
[132] Immigration, Refugees and Citizenship Canada, "How We Use Advanced Analytics, Automation and Other Technologies: Immigration, Refugees and Citizenship Canada," Government of Canada, ultima modifica 20 febbraio 2026, https://www.canada.ca/en/immigration-refugees-citizenship/corporate/transparency/digital-transparency-advanced-data-analytics.html
[133] Immigration, Refugees and Citizenship Canada, "Advanced Analytics and Automation for Processing Applications," Government of Canada, ultima modifica 18 novembre 2024, https://www.canada.ca/en/immigration-refugees-citizenship/corporate/transparency/digital-transparency-advanced-data-analytics/processing-applications.html
[134] Immigration, Refugees and Citizenship Canada, "Advanced Analytics and Automation for Processing Applications.". Immigration, Refugees and Citizenship Canada, "Automation to Help Process Privately Sponsored Refugee Applications," Government of Canada, 16 dicembre 2022, https://www.canada.ca/en/immigration-refugees-citizenship/news/notices/automation-private-refugees.html.
[135] Electronic Immigration Network, "Home Office Suspends Use of Digital Streaming Tool for Visa Applications after Legal Action by JCWI and Foxglove," 9 agosto 2020, https://www.ein.org.uk/news/home-office-suspends-use-digital-streaming-tool-visa-applications-after-legal-action-jcwi-and

support, and overthrust of machines by public bodies.[136] These risks extend to the public sector, including the consular and migration services, where algorithms can affect vulnerable people; the negative effects of algorithms, including "automation bias" and "selective adherence," or the reliance on algorithmic recommendations or only doing so when they uphold existing stereotypes, have also been recognized, as are biases in AI human resource systems that might cause negative outcomes for candidates with disabilities or those from specific ethnic backgrounds.[137]

To summarize, it is evident that consular services encompass the double-edged sword quality of AI in diplomacy. On the other hand, automation and sophisticated analysis tools can speed up processes, minimize backlogs, facilitate the structure of a case, and release time for officials. Conversely, they can complicate transparency and verifiability, lack or fail to have human oversight, impact assessments, and challenges to decisions, and exacerbate bias, obscurity, and inequalities. This is why the use of AI in consular work should also be assistive and supervised: The machine should assist in organizing and processing information, but the decision, especially in a stressful situation, should always be up to the human authority.

The other imperative issue is the international security risks. The implementation of AI in military and strategic scenarios can help reduce the conflict threshold, enabling the use of quicker and perhaps less thought-out decisions. The automation of decision-making processes, as it has been pointed out in the literature, may lead to an increased risk of escalation, decreasing the time to assess the situation and to commit errors or misinterpret the data.[138]

Simultaneously, the adoption of AI raises significant issues of nature ethics and normative. The international organizations, such as UNESCO, have highlighted the importance of developing systems of IA that would respect the key principles such as transparency, accountability, and the respect of human rights. Here, we see the need to ensure that the use of AI is consistent with democratic values and that it does not undermine the protection of fundamental freedoms.[139]

---

[136] Government Digital Service, "Algorithmic Transparency Recording Standard — Guidance for Public Sector Bodies," GOV.UK, aggiornato 8 maggio 2025, https://www.gov.uk/government/publications/guidance-for-organisations-using-the-algorithmic-transparency-recording-standard/algorithmic-transparency-recording-standard-guidance-for-public-sector-bodies

[137] Saar Alon-Barkat and Madalina Busuioc, "Human–AI Interactions in Public Sector Decision Making: 'Automation Bias' and 'Selective Adherence' to Algorithmic Advice," *Journal of Public Administration Research and Theory* 32, no. 1 (2022): pp. 1–4, 12–13, https://doi.org/10.1093/jopart/muac007

[138] European Parliament Research Service, *Artificial Intelligence and Diplomacy in the EU Context* (Brussels: European Parliament, 2022), p, 9.

[139] UNESCO, *Recommendation on the Ethics of Artificial Intelligence* (Paris: UNESCO, 2021), pp, 15–17.

On the European level, these concerns have led to the creation of a specific normative framework to regulate the use of intelligent robots. Specifically, the European Union has presented and subsequently adopted the Artificial Intelligence Act, which proposes an approach based on the risk, differentiating between low-risk, medium- and high-risk applications, and anticipates stricter requirements for those that are more risk-sensitive. This practice is evidence-based, as it shows the need to balance technological innovation with the rights protection and safety of the citizens.[140]

One other restriction is the technological gaps of countries. The development and implementation of artificial intelligence systems may result in augmented asymmetries in the international system because not all countries have identical resources and expertise to do so. This way, AI, besides mirroring, even magnifies the existing inequalities, impacting the power distribution and competition among the actors of the global arena.[141]

Lastly, it is worth keeping in mind the restrictions concerning the quality of data and the possible bias of artificial intelligence systems.

However, algorithms learn using the data that they are trained on, and thus can recreate or amplify existing distortions, which may have counterproductive effects on the decision-making process. This may lead to biased or inaccurate evaluations in the diplomatic arena that undermine the success of the implemented strategies.[142]

Overall, the incorporation of artificial intelligence in diplomacy has numerous challenges that should be approached in a balanced and considerate manner. Even though AI has provided new tools that can be used to improve efficiency and analytical skills, one should also provide sufficient approaches to control, regulation, and monitoring the use of artificial intelligence, and be willing to provide a responsible and consistent approach to the application of the principles of relations between states.[143]

[140] European Commission, *Proposal for a Regulation Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act)* (Brussels: European Commission, 2021), pp,2–4.
[141] Michael C. Horowitz, "Artificial Intelligence, International Competition, and the Balance of Power," *Texas National Security Review* 1, no. 3 (2018): pp,42–43.
[142] Cummings et al., *Artificial Intelligence and International Affairs*, pp, 4–5.
[143] European Parliament Research Service, *Artificial Intelligence and Diplomacy*, pp, 8–10.

Chapter 2 — Artificial Intelligence and Power in Contemporary Diplomacy

2.1 Data and Informational Power in International Relations

In the modern international environment, power cannot be analyzed only through the prism of military might, economic potential, or territorial dominance anymore. While these elements remain crucial for understanding the balance of power in the international community, the ongoing informatization of politics, economics, and society has resulted in the creation of an entirely new infrastructure of power built on the basis of information, data, and processing of those by means of cutting-edge artificial intelligence technology. In this regard, informational power should not be seen as just a new kind of soft power but as a revolutionary transformation of international relations in terms of how actors gain power, wield influence, and engage in global governance. As UNCTAD (UN trade and development) stresses, in the modern world, data are the key strategic resource of the global economy and international competition. Indeed, the regulation of cross-border flows of data has a direct bearing on the distribution of power between states and private actors.[144]

Increasingly, the rise of data's prominence has also begun to redefine the concept of sovereignty itself. Whereas modern sovereignty was traditionally defined by control over a particular territorial space and its natural resources, digital sovereignty may be thought of as control over information infrastructures, communications networks, algorithms, and data flows. This means that the exercise of power in the digital era involves not only coercion and persuasion but also control over the technical infrastructure that enables information flows. In doing so, states and actors possessing digital infrastructure in turn gain an advantageous position in determining the terms of international engagement. According to Dev, this trend has led to the formation of an entirely new "AI World Order" wherein control over computation and AI serves as the basis of global power.[145] This change has direct implications for diplomacy. Diplomatic practice has always been founded on intelligence gathering, mediation between conflicting interests, and state representation within

[144] UNCTAD, *Digital Economy Report 2021: Cross-border Data Flows and Development: For Whom the Data Flow* (Geneva: United Nations Conference on Trade and Development, 2021), iv–v.
[145] Rishik Dev, "AI World Order: How Artificial Intelligence Is Reshaping Global Authority and Legitimacy" (2025), 1–2.

the international arena. Nevertheless, in the modern digital context, all this is undergoing radical changes. Diplomacy does not operate only through official channels and secret diplomacy; it now operates through digital means of communication, social media, opinion tracking systems, and algorithmic tools.[146] Bjola and Holmes describe how digital diplomacy has led to a fundamental change in traditional diplomacy, enhancing the significance of information, communication, and perception management in international decision-making.[147]

The information assumes its strategic importance not only because it contributes to better knowledge of the international environment but also because it enables one to have an impact on how this environment is understood. According to Bjola, digital diplomacy can be defined as the process of "world disclosure" in which diplomats play an active role in producing the political meaning of international events.[148] As a result, informational power consists in the capacity to set narratives, frame issues, and define what problems need to be addressed in the international discourse. In this sense, information is not a neutral concept; rather, it is the arena of political and symbolic struggle. The development of technology only enhances this process. According to Manor, the contemporary public diplomacy involves growing direct engagement of states with foreign public audiences due to the emergence of digital means of communication and social media.[149] Consequently, the effectiveness of diplomatic activity will not only depend on the capacity to negotiate with other states, but also on the opportunity to interfere into the flow of global communication and influence international public opinion. Herein lies the importance of artificial intelligence, which facilitates the processing of large volumes of data about the behavior and preferences of foreign populations.

AI, therefore, serves as an amplifier of informational power. It does not just increase the pace at which information analysis takes place; it changes the character of strategic knowledge. With the help of machine learning and predictive analytics programs, actors will be able to discern certain correlations and patterns that could not have been discerned using conventional means. Cummings et al. point out that AI is changing the understanding of the international situation by decision-makers, who are now able to develop political and military strategies based on novel forms of

[146] Manor 2019, 4–5

[147] Corneliu Bjola and Marcus Holmes, *Digital Diplomacy: Theory and Practice* (London: Routledge, 2015), 1–9.

[148] Corneliu Bjola, "Digital Diplomacy as World Disclosure: The Case of WikiLeaks," *International Studies Perspectives* 16, no. 3 (2015): 315

[149] Ilan Manor, *The Digitalization of Public Diplomacy* (Cham: Palgrave Macmillan, 2019), 4–5.

forecasting and simulation.[150] This informational power is also an energy element, simultaneously. The governance of informational power is therefore intrinsically energy governance as well: having access to data and algorithms is not enough; it also depends on the physical power, cooling, and energy supplies needed to maintain the infrastructure that enables it.[151] This is reflected in the case of Ireland, where data centers used roughly one in five tons of electricity in the country in 2024, a figure that is projected to increase tenfold by mid-century.[152] Nonetheless, this revolution is causing deep-seated inequalities. Information advantage is not equally distributed but is often controlled by entities endowed with sophisticated IT architecture, high computing power, and data access advantages. The UNCTAD study warns against the danger of increasing inequalities in the world due to the development of the digital economy, resulting in a gap between developed countries and technologically underdeveloped nations.[153] This inequality is not only associated with access to technology but also with participation in shaping the rules of the global digital playing field. In this regard, there is also a rise of another kind of technological dependency. States that do not have the ability to collect and analyze data will be dependent on technology created by some other actor, mostly private corporations or the dominant technological powers.[154] In this respect, the problem of asymmetry of power created by the concentration of data and computation becomes relevant since not all international actors have the ability to turn data into power due to the lack of proper resources. As a result, the concentration of digital power may create imbalances in the international order and make the dominant actors even more dominant.[155]

The problem of data control relates to the position occupied by large technology companies. One of the main changes in current power dynamics is the rise in the importance of private actors in global information governance. Digital platforms and large technology companies own infrastructure that is crucial for global communication; gather huge volumes of data; and create

[150] M. L. Cummings et al., *Artificial Intelligence and International Affairs: Disruption Anticipated* (London: Chatham House, 2018), 1–3.
[151] International Energy Agency, *Energy and AI*, 50, 65, 216; OECD, *Measuring the Environmental Impacts of Artificial Intelligence Compute and Applications: The AI Footprint*, OECD Digital Economy Papers, no. 341 (Paris: OECD Publishing, 2022), 5–6.
[152] Central Statistics Office Ireland, "Data Centres Metered Electricity Consumption 2024: Key Findings," June 10, 2025.
[153] UNCTAD, *Digital Economy Report 2021*, iv–v
[154] United Nations Conference on Trade and Development, *Digital Economy Report 2021*, v; pp, 1–3
[155] Michael C. Horowitz, "Artificial Intelligence, International Competition, and the Balance of Power," *Texas National Security Review* 1, no. 3 (2018): 36–57, pp.38–40 and 45–46

artificial intelligence technologies whose abilities outperform many countries.[156] According to Dev, AI transforms the connection between state power and private power, establishing a regime where technology companies gain the capabilities of influencing the world in ways that were possible only for sovereign countries before.[157] This change is essential for understanding informational power in contemporary times because it demonstrates how control over digital infrastructure creates political power beyond the scope of the state. Connected to this process is the informational power of digital platforms, which derives not only from the fact that platforms have access to the data but also from the possibility of controlling the visibility of information.[158] Platforms are not just neutral intermediaries of communications; rather, using selection, recommendation, and moderation algorithms, platforms help select what information will flow, what narrative will prevail, and who will get more visibility in the digital public sphere.[159] In this respect, today's informational power goes beyond the mere production of information, as it implies controlling the communicative environment within which the information will be consumed and interpreted.[160] The latter becomes especially important for the practice of diplomacy since the digital visibility of diplomatic messages affects the potential to influence foreign public opinion and narratives.[161]

These trends create serious problems for international governance. There have been attempts from international and regional organizations to formulate principles and guidelines to regulate the use of artificial intelligence technology, trying to balance between innovation and the protection of fundamental rights. For instance, the OECD stresses the need for transparency, accountability, robustness, and human rights protection when developing and deploying AI systems.[162] The White Paper of the European Commission similarly underlines that artificial intelligence can produce

[156] United Nations Conference on Trade and Development, *Digital Economy Report 2021: Cross-border Data Flows and Development — For Whom the Data Flow* (Geneva: United Nations, 2021), 1–4, 123–125; OECD, *A Blueprint for Building National Compute Capacity for Artificial Intelligence*, OECD Digital Economy Papers, no. 350 (Paris: OECD Publishing, 2023), 9–11.
[157] Dev, "AI World Order," 1–2
[158] Ilan Manor, *The Digitalization of Public Diplomacy* (Cham: Palgrave Macmillan, 2019), 4–5
[159] Rishik Dev, "AI World Order: How Artificial Intelligence Is Reshaping Global Authority and Legitimacy" (2025), 1–2.
[160] Corneliu Bjola, "Digital Diplomacy as World Disclosure: The Case of WikiLeaks," *International Studies Perspectives* 16, no. 3 (2015): 315
[161] Manor, *The Digitalization of Public Diplomacy*, 4–5; Philip Seib, *Real-Time Diplomacy: Politics and Power in the Social Media Era* (New York: Palgrave Macmillan, 2012).
[162] OECD, *Recommendation of the Council on Artificial Intelligence*, OECD/LEGAL/0449 (Paris: OECD, 2019; updated 2021), 3–4

many positive effects, but at the same time, there is a risk of opacity, discrimination, and informational power accumulation.[163] The regulation of informational power becomes extremely complicated due to the global character of digital technologies. Information flows across national borders, making it challenging to apply the traditional regulatory approach based on territoriality.[164] Besides, different states have their own interests in terms of the governance of data and artificial intelligence. While some countries are interested in national security-based approaches, other states see protection of rights and democratic governance as more important.[165] On the diplomatic plane, the change of informational power implies significant repercussions in terms of symbolism and communication. Modern diplomacy is characterized by the emergence of real-time communication processes, which means that events are immediately assessed and disseminated through global digital networks.[166] According to Seib, this process is referred to as "real-time diplomacy" since the pace of digital communication makes strategic thinking more challenging for policymakers.[167] Artificial intelligence intensifies the process and becomes an element of information management, thus adding value to diplomatic competition. The acceleration mentioned above has its benefits and drawbacks. First, AI technologies allow diplomats to observe international conflicts, assess public opinion, and foresee geopolitical events more effectively.[168] On the other hand, the speed and sophistication of algorithmic systems imply lower transparency of decision-making and increasing dependence on the use of technological instruments.[169] According to Cummings et al., the increase in the sophistication of AI technologies makes full human control rather unlikely.[170]

---

[163] European Commission, *White Paper on Artificial Intelligence: A European Approach to Excellence and Trust* (Brussels, 2020), 2

[164] United Nations Conference on Trade and Development, *Digital Economy Report 2021: Cross-border Data Flows and Development — For Whom the Data Flow* (Geneva: United Nations, 2021), 1–4, 123–125; OECD, "Cross-border Data Flows," accessed May 19, 2026.

[165] OECD, "Artificial Intelligence," accessed May 19, 2026; OECD, "AI Principles," accessed May 19, 2026.

[166] Corneliu Bjola and Marcus Holmes, eds., *Digital Diplomacy: Theory and Practice* (London: Routledge, 2015), 1–5.

[167] Philip Seib, *Real-Time Diplomacy: Politics and Power in the Social Media Era* (New York: Palgrave Macmillan, 2012

[168] Corneliu Bjola and Marcus Holmes, eds., *Digital Diplomacy: Theory and Practice* (London: Routledge, 2015), 1–5; Miles Brundage et al., "The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation," Future of Humanity Institute, University of Oxford, Centre for the Study of Existential Risk, University of Cambridge, Center for a New American Security, Electronic Frontier Foundation, and OpenAI, 2018, 16–18.

[169] Madeline Carr, "Power Plays in Global Internet Governance," *Millennium: Journal of International Studies* 43, no. 2 (2015): 640–659; Chatham House, *Artificial Intelligence and International Affairs: Disruption Anticipated* (London: Chatham House, 2018), 10–13.

[170] Cummings et al., *Artificial Intelligence and International Affairs*, 2–3.

It becomes clear that the informational aspect of power is closely related to the problem of trust.[171] International diplomacy has always relied on the possibility of establishing trust relations between political actors.[172] But the appearance of opaque algorithmic systems may undermine this trust, particularly in cases when political actors do not fully comprehend how information technologies operate within decision-making processes.[173] Thus, transparency and explainability become key aspects of the governance of informational power.[174] One of the other important aspects of informational power relates to its democratic dimension.[175] The opportunity to control the data and communication flow can be used not only for improving international coordination but also for manipulating public opinion and influencing political processes.[176] While digital diplomacy allows diplomats to communicate in innovative ways and engage citizens, it can also become a powerful instrument of control and manipulation.[177] As Manor notes, digitalization has changed the nature of state-public relations, making diplomatic communication more transparent yet more susceptible to disinformation.[178] Such ambivalence constitutes one of the most significant features of informational power in contemporary politics.[179] While on the one hand, data and artificial intelligence might improve the efficiency of dealing with complicated issues and coordination of policies at the global level, on the other hand, they could also increase inequalities, strengthen existing power centers, and reduce democratic transparency.[180] Thus, informational power cannot be viewed either negatively or positively per se – it is rather a field of competition and conflict where various interests collide.[181] The advent of informational power should be seen as one of the major developments in contemporary international relations.[182] Indeed, nowadays, the capacity to influence data, digital infrastructure, and artificial intelligence systems constitutes a vital element of geopolitical rivalry and diplomatic strategy.[183] Yet, such a development implies new

[171] OECD, *Recommendation of the Council on Artificial Intelligence*, OECD/LEGAL/0449 (Paris: OECD, 2019; updated 2021), 3–4; European Commission, *White Paper on Artificial Intelligence: A European Approach to Excellence and Trust* (Brussels, 2020), 2.
[172] Bjola and Holmes 2015, 1–6; Seib 2012, 1–5
[173] European Commission 2020, 2; OECD 2019/2021, 3–4; Cummings et al. 2018, 2–3
[174] OECD 2019/2021, 3–4; European Parliament and Council 2024, recitals 1–6
[175] Manor 2019, 4–5; Bjola 2015, 315
[176] UNCTAD 2021, iv–v; Manor 2019, 4–5
[177] Manor 2019, 4–8; Seib 2012, 1–5
[178] Manor 2019, 4-5
[179] Bjola 2015, 315; Manor 2019, 4–5
[180] UNCTAD 2021, iv–v; OECD 2019/2021, 3–4; European Commission 2020, 2
[181] Dev 2025, 1–2; Horowitz 2018, 38–41
[182] UNCTAD 2021, iv–v; Dev 2025, 1–2
[183] Horowitz 2018, 38–41; Cummings et al. 2018, 1–3; Dev 2025, 1–2

asymmetries, new dependencies, and new power concentrations that require critical evaluation of global information governance. In turn, artificial intelligence increases the strategic importance of data, but at the same time makes it particularly important to establish clear rules and principles that would guarantee transparency, accountability, and fairness in the international system.[184]

2.2 AI and Strategic Competition Among States

In the current international system, artificial intelligence represents one of the major sources of transformation in the competitive dynamics of states.[185] The latter is no longer a technological innovation but an enabling capability, having the potential to reshape the economy, security, diplomacy, intelligence, defense, and governance.[186] As such, AI needs to be considered an enabling capability, with the potential to shape not only the instruments through which states compete but also the very context of international competition.[187] According to Horowitz, artificial intelligence cannot be likened to just any weapon system, but rather to a general-purpose technology, such as electricity and the internal combustion engine, because its significance lies in its capacity to revolutionize several sectors at once.[188] The strategic implications of AI arise largely from its capability to turn information into a strategic asset. Countries that have access to large amounts of data, powerful computing facilities, and skills in developing algorithms can develop a strategic edge in their comprehension of the international system and the actions of other players. As stressed by Cummings et al., artificial intelligence has already started to impact the interpretation of the world and decision-making by policymakers, as it enables faster processing of vast amounts of data compared to conventional approaches.[189] Therefore, modern strategic competition not only involves military dominance and economic power but also information processing and strategic knowledge generation.[190]

---

[184] OECD, *Recommendation on Artificial Intelligence*, 3–4; UNCTAD, *Digital Economy Report 2021*, iv–v; Dev, "AI World Order," 1–2.
[185] Horowitz 2018, 38–41; Dev 2025, 1–2
[186] Horowitz 2018, 38; Cummings et al. 2018, 1–3
[187] Horowitz 2018, 38–41; UNCTAD 2021, iv–v; Dev 2025, 1–2
[188] Michael C. Horowitz, "Artificial Intelligence, International Competition, and the Balance of Power," *Texas National Security Review* 1, no. 3 (2018): 38
[189] M. L. Cummings et al., *Artificial Intelligence and International Affairs: Disruption Anticipated* (London: Chatham House, 2018), 1–3
[190] Cummings et al. 2018, 1–3; Horowitz 2018, 38–41; UNCTAD 2021, iv–v

Such a change intensifies the relationship between data and power. According to the UNCTAD report, data has become an important strategic resource since it impacts value generation, economic governance, and development for countries.[191] In relation to the struggle between the states, this implies that managing the flow of data is an integral element of power.[192] Countries that control digital infrastructure, platforms, cloud computing, and computing capabilities can not only improve their economy but also enhance their geopolitical power.[193] Thus, the struggle for artificial intelligence is also the struggle for control over the infrastructure that supports power.[194] Artificial intelligence technologies are also transforming the hierarchies of the international system.[195] According to Dev, artificial intelligence creates a new "AI World Order," which can be described as a combination of two opposing processes: while on the one hand it concentrates power in the hands of states and corporations with high technological development, on the other hand it allows for a decentralization of some capacities, allowing even non-state actors to deploy advanced instruments of manipulation and misinformation.[196] This paradox is crucial, since artificial intelligence not only empowers dominant actors but also destabilizes the international system by increasing the number of actors who have access to means of exercising influence.[197] The competition for AI technologies is especially apparent in the rivalry between major powers.[198] As noted by Horowitz, the effect of artificial intelligence on the balance of power will depend on the speed with which states implement AI and incorporate it into their operations.[199]

Thus, the strength lies both in technological innovation and in the ability to assimilate innovations within the institutions. In this regard, the US, China, as well as Russia and the EU to a certain extent, compete not only in terms of creating cutting-edge technologies, but also in terms of developing ecosystems for research, investment, infrastructure, and strategic use of innovations. In particular, the matter is especially acute in the sphere of the defense industry. According to the

[191] UNCTAD, *Digital Economy Report 2021: Cross-border Data Flows and Development: For Whom the Data Flow* (Geneva: United Nations Conference on Trade and Development, 2021), iv–v
[192] UNCTAD 2021, iv–v; Dev 2025, 1–2
[193] UNCTAD 2021, iv–v; Dev 2025, 1–2; Horowitz 2018, 38–41
[194] Horowitz 2018, 38–41; UNCTAD 2021, iv–v
[195] Dev 2025, 1–2; Horowitz 2018, 41
[196] Rishik Dev, "AI World Order: How Artificial Intelligence Is Reshaping Global Authority and Legitimacy" (2025), 1–2
[197] Dev 2025, 1–2; Bjola 2015, 315
[198] Michael C. Horowitz, "Artificial Intelligence, International Competition, and the Balance of Power," *Texas National Security Review* 1, no. 3 (2018): 41; Sebastian Clapp, *Defence and Artificial Intelligence* (Brussels: European Parliamentary Research Service, 2025), 1
[199] Horowitz, "Artificial Intelligence, International Competition," 41

2025 EPRS briefing, artificial intelligence transforms contemporary warfare through its applications in the areas of intelligence, autonomous technologies, and cybersecurity; moreover, the briefing indicates that a full-scale race between countries for military AI has already begun, with China and the US trying to become leaders and Russia making investments in artificial intelligence.[200]

The incorporation of AI into defense will also have consequences for deterrence and superiority. AI technology can be deployed to enhance intelligence, speed up targeting and logistics, provide better cybersecurity, and facilitate command and control. Nevertheless, such acceleration carries considerable dangers. Cummings et al. note that while the deployment of AI in the military will help speed up decision-making, it will limit the ability of humans to oversee the process and make judgments.[201] As a result, in a crisis, the danger of escalation may become higher due to the increased reliability of automated decision-making. The competition over AI technologies, then, is not just about who will be able to build better AI, but who will be able to regulate its risks.

Artificial Intelligence usage in the military is one of the most chivalrous fields of the modern strategic rivalry. AI systems can be used to improve intelligence, surveillance, pattern recognition, cyber defense, logistics, command and control, targeting support, and battlefield analysis. In this respect, it is not just another piece of technology, but also an operating "force multiplier" for army organizations, where information can be gathered, processed, and used at a much faster pace than previously. NATO's AI strategy 2024 reflects the same attitude: NATO considers AI to be a pivotal technology for defense and security that will be early adopted by the Alliance, while ensuring adherence to core principles in the respectful use of AI, including its legal nature, its accountability, its traceability, its reliability, its governability, and the reduction of bias.[202] The rapid acceleration in military processes, however, can generate serious risks thanks to the speed of AI. Automated systems can speed up data collation, target identification, and order fulfillment, but can also reduce decision-making time. This time, wrapping can add to the potential for escalation in an international crisis scenario when decisions of political and military implications may be made based on algorithmic outputs that might not be fully understood, interpreted, and verified. The

[200] Sebastian Clapp, *Defence and Artificial Intelligence* (Brussels: European Parliamentary Research Service, 2025), 1.
[201] Cummings et al., *Artificial Intelligence and International Affairs*, 7–8.
[202] NATO, "Summary of NATO's Revised Artificial Intelligence (AI) Strategy," 10 luglio 2024, https://www.nato.int/en/about-us/official-texts-and-resources/official-texts/2024/07/10/summary-of-natos-revised-artificial-intelligence-ai-strategy

challenges here are not only about the efficiency of the military AI system, but also the need to ensure that meaningful human control is maintained. Indeed, the International Committee of the Red Cross (ICRC) states that ASs present a humanitarian, legal, and ethical dilemma, as they can program and identify targets after activation, without relying on any human intervention.[203] This was exemplified in a debate about the development of autonomous weaponry. But on the other hand, it can lead to losses in the military as well as in terms of speed, precision, and force-protection. On the other hand, if a system is extremely chaotic (or, when human beings lose the ability to limit, control, and stop the machine's actions), it becomes difficult to ensure that compliance will be respected with international humanitarian law.[204] The ICRC has recently recommended three new, legally binding rules in its report by ICRC-SIPRI on the limits of autonomy: banning unpredictable autonomous systems; banning the use of autonomous systems that target human beings; and establishing limits with respect to target types, duration, geographical scope, scale of use, and human oversight.[205] The Report on the Limits of Autonomy also focuses on the need for human control, which has three dimensions: control over weapon parameters; control over the operational environment; and human supervision. The US is also supporting an effort for the Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy of the United States' promotion, using this approach. The declaration has been followed by some sixty States subsequently signing an "aspirational blueprint" for AI responsible usage in the field of military action in Seoul, proving that this matter is now firmly on the international political agenda.[206] But the fact that China is resisting signing that document also indicates its biggest drawback: there is no universal agreement yet on shared and enforceable rules.[207]

Artificial Intelligence (AI) is another factor linked to hybrid warfare. AI in this context can be applied to an information environment, perception manipulation, the destabilization of societies,

[203] International Committee of the Red Cross, "ICRC Position on Autonomous Weapon Systems," 12 maggio 2021, https://www.icrc.org/en/document/icrc-position-autonomous-weapon-systems
[204] Ivi.
[205] International Committee of the Red Cross and Stockholm International Peace Research Institute, "Limits on Autonomy in Weapon Systems," 2 giugno 2020, https://www.icrc.org/en/document/limits-autonomous-weapons
[206] U.S. Department of State, "Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy," 2023, https://2021-2025.state.gov/bureau-of-arms-control-deterrence-and-stability/political-declaration-on-responsible-military-use-of-artificial-intelligence-and-autonomy/
[207] Joyce Lee, "Sixty Countries Endorse 'Blueprint' for AI Use in Military; China Opts Out," *Reuters*, 10 settembre 2024, https://www.reuters.com/technology/artificial-intelligence/south-korea-summit-announces-blueprint-using-ai-military-2024-09-10/

and the erosion of trust in institutions, not military targets.[208] AI-related technologies can, in fact, augment both the success of influence operations and the disinformation campaign, fake profiles, the generation of artificial consensus, the manipulation of online conversations, and the production of deepfakes, all of these to go hand-in-hand with pursuing political objectives, which is an invitation to 'gray zone' warfare.[209] This plays a specific role since hybrid warfare is not just a dimension of war, but also one of civil society, elections, public opinion, and democratic trust. The 2024 guidance on information threats issued by the NATO member states indicates that an information threat can be directed against national security interests and comprises information operations, information manipulation and interference, and disinformation. Thus, in NATO's view, AI and deepfakes can exacerbate manipulative techniques, have the potential to create confusion in the information environment, affect perceptions, and undermine trust in public communications, and consequently impact the EU's foreign policy and security.[210] The EEAS also believes Foreign Information Manipulation and interference are becoming a threat to the European Union's foreign policy and security efforts, as it has the potential to polarize societies, affect democratic processes, and impact the EU's capacity to pursue its foreign policy interventions.[211] What follows is that strategic competition in AI isn't only about military means; it becomes about any means. It also includes information sphere protection, maintenance of public trust, attribution of hostile operations, and democratic resilience. AI can thus revolutionize conventional and hybrid warfare: both increasing intelligence, targeting, and command and control, and manipulating, disinforming, and causing psychological influence. Therefore, AI diplomacy needs to tackle questions of military innovation along with international norms, human oversight, humanitarian law, information security, and cooperation amongst states, international organizations and civil society.

[208] Ralph Thiele, "Artificial Intelligence – A Key Enabler of Hybrid Warfare," Hybrid CoE Working Paper 6, marzo 2020, https://www.hybridcoe.fi/publications/hybrid-coe-working-paper-6-artificial-intelligence-a-key-enabler-of-hybrid-warfare/ Nicolas Mazzucchi, "AI-Based Technologies in Hybrid Conflict: The Future of Influence Operations," Hybrid CoE Paper 14, giugno 2022, https://www.hybridcoe.fi/publications/hybrid-coe-paper-14-ai-based-technologies-in-hybrid-conflict-the-future-of-influence-operations/.
[209] NATO, "NATO's Approach to Counter Information Threats," 18 ottobre 2024, https://www.nato.int/en/about-us/official-texts-and-resources/official-texts/2024/10/18/natos-approach-to-counter-information-threats.
[210] European External Action Service, "Information Integrity and Countering Foreign Information Manipulation & Interference (FIMI)," 17 marzo 2026, https://www.eeas.europa.eu/eeas/tackling-disinformation-foreign-information-manipulation-interference_en?s=2803.
[211] Ivi.

The EU's Artificial Intelligence Act is grounded in the recognition that while AI promises substantial gains, it also poses security, human rights, democratic, and rule of law risks.[212] Even though the EU regulation does not cover the application of AI for military purposes, it is indicative of a wider political attitude: technological power needs to be regulated by certain rules and principles.[213] But when it comes to international politics, the problem is that not all states have the same values or regulatory systems.[214]

This difference in regulations is one of the elements of strategic competition.[215] While some parties advocate for an AI development strategy that relies on innovation, industrial competitiveness, and technological superiority, the European Union aims to develop an AI that will be based on trust, respect for human rights, and humanity.[216] The European Commission White Paper from 2020 argues that AI should be developed in such a manner as to facilitate innovation while mitigating the risks associated with decision-making transparency, discrimination, and violation of people's privacy rights.[217] It can be seen that competition over AI is not only about material resources but also about regulation. In this respect, AI can be considered as an instrument of regulatory power as well. Those who manage to establish international standards will thus have a say in how the technology will evolve and be utilized around the world. Transparency, robustness, safety, and accountability are listed by the OECD as the core principles of responsible AI.[218] Such principles are far from being impartial because they reflect a certain political understanding of AI based on the assumption that technology should be in line with democratic values.[219]

In addition, technological competition is redefining the role of private actors. Big technology firms have data, cloud infrastructure, research facilities, and AI systems that outperform many states in

[212] European Parliament and Council of the European Union, *Regulation (EU) 2024/1689 of 13 June 2024 Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act)*, Official Journal of the European Union, 2024, recitals 1–6

[213] European Parliament and Council 2024, recitals 1–6; European Commission 2020, 2

[214] European Commission 2020, 2; OECD 2019/2021, 3–4

[215] Horowitz 2018, 38–41; Dev 2025, 1–2

[216] European Commission 2020, 2; European Parliament and Council 2024, recitals 1–6; OECD 2019/2021, 3–4

[217] European Commission, *White Paper on Artificial Intelligence: A European Approach to Excellence and Trust* (Brussels, 2020), 2.

[218] OECD, *Recommendation of the Council on Artificial Intelligence*, OECD/LEGAL/0449 (Paris: OECD, 2019; updated 2021), 3–4

[219] European Commission 2020, 2; OECD 2019/2021, 3–4; European Parliament and Council 2024, recitals 1–6

some ways.[220] According to Dev, Big Tech is an independent source of geopolitical power because it has the infrastructure and the resource base that are essential for the development of AI.[221]

This creates a significant shift, like international competition: states compete not only against each other but also must deal with private actors who have strategic resources at their disposal. This makes technological sovereignty relevant, as reliance on third parties in terms of the platform or infrastructure can diminish the strategic autonomy of states. Another factor that contributes to this trend is the power of digital platforms, which is not simply a matter of having data or infrastructure but is associated with the ability to regulate information visibility and to control access to the digital space that shapes political influence.[222] However, platforms cannot be seen simply as neutral instruments; instead, they are actors who create the information environment via algorithms, recommendations, and computational facilities.[223] Consequently, strategic competition over AI technologies will inevitably include private actors that have the capacity to affect public communication, narrative circulation, and digital influence of states.[224] Such competition can be observed in the race for semiconductors, cloud computing capabilities, advanced language models, and computing infrastructure, even though these aspects do not necessarily fall within the scope of discussion in the articles above. Nevertheless, they are the logical extensions of the arguments made by UNCTAD and Dev: whoever owns digital infrastructure owns the power to build large-scale AI technologies.[225] In the modern world, technological power has become a question of strategic independence.

The competition among nations for superiority in AI technology is also leading to the emergence of a new geography of global inequality. Advanced nations can collect data, build infrastructure, and recruit skilled personnel, whereas less developed nations might end up as mere consumers of the technologies developed elsewhere. The UNCTAD reports that the data economy contributes to the widening inequality gap between nations because data creates wealth in a few selected economies and a few selected platforms.[226] Such an inequality also leads to a diplomatic imbalance, whereby technologically weaker nations lack bargaining power at the international

[220] UNCTAD 2021, iv–v; Dev 2025, 1–2
[221] Dev, "AI World Order," 1–2
[222] Ilan Manor, *The Digitalization of Public Diplomacy* (Cham: Palgrave Macmillan, 2019), 4–5
[223] Dev 2025, 1–2
[224] Bjola 2015, 315; Manor 2019, 4–5
[225] UNCTAD 2021, iv–v; Dev 2025, 1–2
[226] UNCTAD 2021, iv-v

negotiation table. Technological inequality also leads to a diplomatic imbalance. Nations with more advanced computational capacity and data collection capability can participate in international diplomacy with superior analytical, technical, and predictive skills.[227]

On the contrary, nations that lack adequate digital assets could find themselves subordinated not only from an economic perspective but also within the process of setting international standards and technological governance.[228] In this regard, AI can contribute to deepening imbalances in the international environment, turning the technological gap into a diplomatic and regulatory one.[229]

Indeed, strategic competition in AI can be characterized by various aspects, including economic, military, regulatory, and diplomatic.[230] Economically, AI is a source of productivity and innovation.[231]

Military-wise, AI is an enabler for intelligence, defense, and autonomy.[232] Regulatory-wise, AI is the field where competing regulatory models will clash.[233] Diplomatically speaking, AI is the topic to be negotiated and used for exerting influence.[234] Thus, the multidimensional nature of strategic competition in AI means that AI is not just another technological sector, but rather a strategic systemic technology.[235]

Yet another important aspect to consider is the potential danger posed by the militarization and securitization of artificial intelligence.[236] Whenever the use of any technology is understood through the prism of national security, countries are likely to engage in its development in a competitive manner.[237] Consequently, there may be fewer prospects for international cooperation and even a rise in the level of perceived threats.[238] As stated in the EPRS briefing, the absence of an international regulatory framework for military AI creates problems when trying to reach an agreement at the global level because of strategic differences between great powers.[239] AI is not

---

[227] Ivi.
[228] Cummings et al. 2018, 1–3
[229] Dev 2025, 1–2; UNCTAD 2021, iv–v
[230] Horowitz 2018, 38–41; Cummings et al. 2018, 1–3; OECD 2019/2021, 3–4
[231] UNCTAD 2021, iv–v
[232] Sebastian Clapp, *Defence and Artificial Intelligence* (Brussels: European Parliamentary Research Service, 2025), 1; M. L. Cummings et al., *Artificial Intelligence and International Affairs: Disruption Anticipated* (London: Chatham House, 2018), 7–8.
[233] European Commission 2020, 2; European Parliament and Council 2024, recitals 1–6
[234] Bjola and Holmes 2015, 1–9; Manor 2019, 4–8
[235] Horowitz 2018, 38–41; Dev 2025, 1–2; UNCTAD 2021, iv–v
[236] Clapp 2025, 1; Cummings et al. 2018, 7–8
[237] Horowitz 2018, 38–41; Clapp 2025, 1
[238] OECD 2019/2021, 3–4; Clapp 2025, 1; Cummings et al. 2018, 7–8
[239] EPRS 2025, 1

always a source of conflict, however. In contrast, it may also become an object of cooperation through diplomacy, particularly in terms of standardization, cybersecurity, data governance, and risk management. The OECD claims that international cooperation is crucial to ensure that there is trust in AI, which should be consistent with shared values.[240] This example shows that strategic competition does not imply the absence of cooperation but makes it more complex. Countries may compete for technological supremacy while having an interest in establishing common standards to avoid systemic risks.

The problem here lies in the fact that collaboration on AI requires trust, while strategic competition hinders it.[241] States may not be willing to collaborate on information-sharing, restriction of certain technology development, and global standardization if they consider AI to be an important aspect in ensuring their national security.[242] In this sense, there is a structural contradiction: while AI needs to be regulated, the strategic significance of AI means that states must defend their competitive edge.[243] This is one of the major issues in AI diplomacy. As theory suggests, artificial intelligence is neither a replacement for traditional sources of power nor something new; instead, it is a redistribution of traditional sources of power.[244] Military sources of power remain significant; however, they are accompanied by analysis, automation, and prediction tools.[245] Economic sources of power continue to be important; however, they are contingent upon data governance and technology governance.[246] Soft power remains relevant; however, it operates within the digital information environment.[247]

These issues have profound consequences for diplomacy. States will be forced to acquire new diplomatic skills related to negotiations on technical issues, understanding the strategic value of AI, and participating in international standard-setting.[248] This implies that AI diplomacy will require a combination of various skills: political, technical, legal, and strategic.[249] Artificial

[240] OECD, *Recommendation on Artificial Intelligence*, 3–4
[241] OECD 2019/2021, 3–4; Clapp 2025, 1
[242] Clapp 2025, 1; Horowitz 2018, 38–41.
[243] European Commission 2020, 2; OECD 2019/2021, 3–4; Horowitz 2018, 38–41.
[244] Horowitz 2018, 38–41; Dev 2025, 1–2
[245] Cummings et al. 2018, 7–8; Clapp 2025, 1
[246] UNCTAD 2021, iv–v; OECD 2019/2021, 3–4
[247] Manor 2019, 4–8; Bjola 2015, 315; Seib 2012, 1–5
[248] Roumate 2021, front matter; OECD 2019/2021, 3–4
[249] Roumate 2021, front matter; Bjola and Holmes 2015, 1–9

Intelligence and Digital Diplomacy: A New Relationship? Edited by Roumate emphasizes the strong link between AI and digital diplomacy today because of new technology.[250]

The crucial point, therefore, is that strategic competition in AI goes beyond technological superiority to include the ability to translate such a superiority into political authority.[251] States and private actors who control data, infrastructure, and standards benefit not only from an operational advantage but also gain the power to shape the rules by which other actors engage within the international system.[252] For these reasons, artificial intelligence should not be seen as merely an instrument of state competition but one of the structural conditions through which international power is produced, distributed, and contested today.[253]

To conclude, the advancement of artificial intelligence is among the most crucial spheres of strategic competition.[254] AI influences the processes of gathering information by states, processing of information, creating capabilities, constructing narratives, and setting standards within the international community.[255] On the contrary, AI technology does not neutrally impact the world. It is biased toward sophisticated actors, benefits large technology corporations, increases inequalities, and poses risks for escalating tensions and mistrust.[256] The competition around AI, therefore, goes beyond technological advantages.[257] Rather, the battle over AI defines the shape of the international order in the future. It means that artificial intelligence should be understood as a tool of power, both a tool that can improve the position of states and a tool used to decide upon who will shape the future international order.[258]

2.3 Implications for Diplomatic Influence and Negotiation

The use of artificial intelligence (AI) in diplomacy must not be seen as merely a technological advancement, but as an ontological transformation of the diplomatic process that profoundly alters how influence and negotiation are exercised. In a world of "Real-Time Diplomacy," Philip Seib

[250] Fatima Roumate, ed., *Artificial Intelligence and Digital Diplomacy: Challenges and Opportunities* (Cham: Springer, 2021), front matter
[251] Horowitz 2018, 38-41
[252] Dev 2025, 1-2
[253] OECD 2019/2021, 3-4; Horowitz 2018, 38-41
[254] Horowitz 2018, 38–41; Dev 2025, 1–2
[255] Cummings et al. 2018, 1–3; Bjola 2015, 315; OECD 2019/2021, 3–4
[256] UNCTAD 2021, iv–v; Dev 2025, 1–2; OECD 2019/2021, 3–4
[257] Horowitz 2018, 38–41; UNCTAD 2021, iv–v
[258] Horowitz 2018, 38–41; Dev 2025, 1–2; UNCTAD 2021, iv–v

(2012) points out how the advent of social media has established a tempo of action that undermines the time horizon for reflection, forcing diplomats into a "continuous present" in which image and immediate response replace long-term strategy.[259]AI amplifies this issue to a systemic level: influence is no longer limited to the efficacy of national discourse, but to the management of the "logic of algorithms" that regulate the visibility and credibility of messages in the global arena. Digital diplomacy has become a domain in which the reputation of states is constantly defined through the management of a huge data flow, turning soft power into a highly volatile commodity at the mercy of automated narratives and bots that can sway public opinion in microseconds.[260]

This power shift is symbolized by the rise of what Rishik Dev (2025) describes as an "AI World Order", in which the authority of the international system is concentrated in the hands of those in control of technological sovereignty, thus establishing a new hierarchy of international power based on computing capacity and the capacity to impose normative structures through algorithmic advantage.[261]

Turning to the practice of negotiations, the use of AI is clear in predictive analytics and decision-support systems, which are already being used to revolutionize multilateral negotiations. A notable application is, for example, highly technical negotiated trade liberalization, where AI is employed to parse thousands of pages of legal documents and tariffs to detect zones of possible agreement (ZOPAs) beyond human comprehension. Volker Stanzel and Stefan Voelsen (2022) report that in the arena of international mediation and conflict resolution, AI has been used to track violations of temporary ceasefires and the language of the disputants to help predict potential escalations, giving UN peace negotiators a greater "situational awareness".[262]

But as the Chatham House report by M. L. Cummings and others (2018) points out, heavy reliance on these predictive algorithms could lead to "negotiation rigidity": if members of a negotiating team solely rely on algorithms that recommend solutions based on past outcomes, they would disregard other factors, including creative thinking and empathy, essential for successful

[259] Philip Seib, *Real-Time Diplomacy: Politics and Power in the social media Era* (New York: Palgrave Macmillan, 2012), pp, 12-18.
[260] Corneliu Bjola e Ilan Manor, a cura di, *The Oxford Handbook of Digital Diplomacy* (Oxford: Oxford University Press, 2024), pp, 45-52.
[261] Rishik Dev, "AI World Order: How Artificial Intelligence Is Reshaping Global Authority and Legitimacy," *APSA Preprints* (2025): pp, 4-9, doi:10.33774/apsa-2025.
[262] Volker Stanzel e Stefan Voelsen, "Diplomacy and Artificial Intelligence: Reflections on Practical Assistance for Diplomatic Negotiations," *SWP Research Paper* n. 1 (2022): pp, 1-4.

diplomatic compromise.[263] Meanwhile, Karim Hamidouche notes that AI is now an indispensable "cognitive assistant" in the management of the technicalities of multilateral treaties (such as those on climate or regulation of cyberspace), which are too complex for diplomats to grasp and analyze individually, but have also created a risk of waning autonomy in the diplomatic decision-making processes.[264]

The systemic security threats of this phenomenon are that diplomacy will become a depoliticized bureaucratic procedure and negotiation an algorithmic process without political and moral content. Underlining the fact that the legitimacy of an agreement lies not only in its technical flavor but also in the mutual recognition of the parties involved, which smartphones are still struggling to implement,[265] Corneliu Bjola and Markus Kornprobst underpin the importance of the preservation of equality of parties.[266] Moreover, security and reliability are constrained by advanced disinformation in these interactions; according to Khedidja Ziani (2025), the stability of cyber-diplomacy rests on the ability to rely on AI technology for data integrity in cyber-negotiations to prevent and detect attempts to alter algorithms to perform functions by hostile actors.[267]

In this context, "data diplomacy" emerges as a new field of geopolitical competition where, as per the UNCTAD report (2021), the ability to manage international data flows and to negotiate international technological standards becomes a new driver of national strategies for securing the interests of a country.[268]

In conclusion, the influence of AI on negotiation and diplomatic practice is irreversible and ambivalent: on one hand, AI provides new management tools to deal with complexity and prevent conflict, on the other hand, it brings risks related to the lack of transparency of decision-making processes and the creation of an "AI Divide" that may push technologically less advanced countries out of the international community, making transparency and trust more difficult to maintain in 21st-century diplomatic processes.

---

[263] M. L. Cummings et al., *Artificial Intelligence and International Affairs: Disruption Anticipated* (London: Chatham House, 2018), pp, 22-26.

[264] Karim Hamidouche, "Artificial Intelligence: A New Tool for Diplomats," in *Artificial Intelligence and Digital Diplomacy*, a cura di Fatima Roumate (Cham: Springer Nature, 2021), pp, 103-110.

[265] Corneliu Bjola e Markus Kornprobst, *Understanding International Diplomacy: Theory, Practice and Ethics*, 2ª ed. (London: Routledge, 2018), pp, 202-205.

[266] Ivi.

[267] Ziani, Khedidja. "Security and Reliability in Digital Diplomacy – Review of Selected Literature." *Humanities and Social Sciences* 32, no. 3 (2025): 139–152. https://doi.org/10.7862/rz.2025.hss.36.

[268] UNCTAD, *Digital Economy Report 2021: Cross-border Data Flows and Development* (Geneva: United Nations, 2021), pp, 15-20.

Chapter 3 — Ethical and Political Challenges of AI in Diplomacy

3.1 Bias, Opacity, and Algorithmic Risks

Introduction of artificial intelligence to the modern-day practice of diplomacy is one such structural change with a fundamental impact on the way information is generated, perceived and used in the international system. The concerns centering on the bias of algorithms, the obscurity of the system of the carried-out decisions, or systemic risks gain prominence in this context not only technically but also on the political or regulatory level. Indeed, these crucial problems have an impact on the quality of diplomatic decisions and, more broadly, the legitimacy of international decision-making processes.[269]

The use of artificial intelligence can be characterized by algorithmic bias as one of the key issues. History teaches AI systems, which in turn are prone to being learned, and are invariably an indicator of the current social, economic, and political structures. This means that any defects in the data can be reproduced and magnified by the models of algorithms, providing discriminating or discriminatory outcomes.[270] This fact acquires a specific meaning in the case of the diplomatic environment, where the creation of representations of the international reality conditions one of the most significant parts of the decision-making process.[271]

The danger at the international level is not just that of creating technical mistakes, but of playing its part in creating warped perceptions about world actors, threats, and strategic opportunities. This might affect political priorities and policies of diplomacy and have immediate consequences to the stability of the international system. Artificial intelligence should be created in line with human and democratic principles, as the OECD stresses, to prevent the discriminatory impact and promote fair usage of technology.[272] This point of view is supported by the European guidelines, emphasizing the necessity to take care of avoiding reproducing bias in information and models to make sure that the decisions made through AI systems do not hurt the vulnerable population.[273]

---

[269] European Commission, *White Paper* (2020), p. 2. European Commission HLEG, *Ethics Guidelines* (2019), pp, 2–3.
[270] High-Level Expert Group on Artificial Intelligence, *Ethics Guidelines for Trustworthy AI* (Brussels: European Commission, 2019), p. 2
[271] Ivi.
[272] OECD (2019/2021), pp. 3–4.
[273] HLEG (2019), p, 2.

One of the issues that is strongly associated with the issue of bias is that of algorithmic opacity. The extensive use of AI systems, particularly those relying on advanced machine learning algorithms, works under procedures that are hard to understand, and it is difficult to explain how decisions are produced.[274] This opaque nature is an enormous impediment to diplomacy, which is rooted in concepts of accountability, justification, and explicit negotiation. This means that in the diplomatic arena, the transparency of AI systems may hurt trust in the decision-making processes since the involved parties cannot fully comprehend the criteria upon which the automated decisions are made.[275] In diplomacy, it would mean that the decision-making processes will undergo a diminished ability to rationalize the policy decisions and be able to clearly attribute responsibility.[276]

In European guidelines, transparency and explainability are mentioned as the basic conditions of the development of trustworthy artificial intelligence.[277] To secure the validity of the decisions and credibility of the stakeholders, it is vital to be able to comprehend the functioning of algorithmic systems. It is even more significant in the situation of digital diplomacy. Manor emphasizes the role of communication and perception management as the key aspects of modern diplomatic activity.[278] Ten In this case, algorithmic opaqueness can not only affect the quality of decisions, but also whether states can create narratives and impact the opinion of the world population.

Equally, Bjola notes that digital diplomacy is a constant process of meaning-making and meaning-interpreting international events.[279] Meanings-making can be changed by opaque algorithmic systems, though, and breaking down transparency and enhancing ambiguity of the decision-making process. In addition to bias and lack of transparency, there is another main problem related to systemic risks regarding the implementation of artificial intelligence. The prevalence of algorithmic systems may also produce cumulative impacts that are challenging to anticipate, particularly the tasks that can be performed with the help of AI negatively, such as the

[274] European Commission, White Paper (2020), p. 2.
[275] Ivi.
[276] OECD (2019/2021), p, 3.
[277] HLEG (2019), pp, 2–3.
[278] Ilan Manor, *The Digitalization of Public Diplomacy* (Cham: Palgrave Macmillan, 2019), pp, 4–5.
[279] Corneliu Bjola, "Digital Diplomacy as World Disclosure: The Case of WikiLeaks," *International Studies Perspectives* 16, no. 3 (2015): p, 315.

manipulation of information.[280] The White Paper also mentions the appearance of harmful uses of AI, including the manipulation of information.[281]

The risks are in a larger context of change in the international system. Cummings emphasizes the impact of artificial intelligence that is reshaping the decision-making process of decision-makers and their cognition of the world.[282] This shift is creating novel kinds of technological reliance, which may heighten the susceptibility of decision-making units. Dev also notes how artificial intelligence is assisting in redrawing the power structures in the world, with those in charge being those who possess command of the technology infrastructure,[283] regardless of whether the consequences of algorithmic risks assume a geopolitical form that affects the nature of competition on the international stage.

Data position is another aspect of complexity. According to the UNCTAD report, data has emerged as a strategic resource, but the centrality of data also presents new vulnerabilities in terms of security, management, and distribution of digital resources. [284]

To overcome the challenges, European guidelines, in turn, lay stress on the principle of human oversight, which will help avoid over-delegation of the decision-making process to AI.[285] But the complexity of algorithmic systems is increasing, and it is hard to retain effective control, making it a contradiction between efficiency and accountability.[286] This tension is one of the central issues of modern diplomacy that must strike a balance between technological innovations and democracy.[287]

The three essential aspects of the challenges related to the use of AI in modern diplomacy are algorithmic biases, opacity, and systemic risks. Their management needs an integrated strategy that involves regulation, technical skills, and moral contemplation to make sure that the

[280] European Commission (2020), p, 2.
[281] Ivi.
[282] M. L. Cummings et al., *Artificial Intelligence and International Affairs: Disruption Anticipated* (London: Chatham House, 2018), 1–3.
[283] Rishik Dev, "AI World Order: How Artificial Intelligence Is Reshaping Global Authority and Legitimacy" (2025), 1–2.
[284] UNCTAD, *Digital Economy Report 2021: Cross-border Data Flows and Development: For Whom the Data Flow* (Geneva: United Nations Conference on Trade and Development, 2021), pp, iv–v.
[285] HLEG (2019), p, 3.
[286] European Commission 2020, 2; European Commission HLEG 2019, 3; Cummings et al. 2018, 2–3
[287] Ivi.

implementation of AI will help to reinforce the transparency, fairness, and legitimacy of international decision-making procedures.[288]

3.2 Accountability and Responsibility in AI-Assisted Decision-Making

The implementation of artificial intelligence into the modern diplomatic decision-making procedures presents some rather fundamental questions concerning the issue of responsibility and accountability as the main issues of the operation of the political establishment and the legitimacy of international action. Implementing algorithmic systems as a decision-support tool brings in the radical change of how decisions are made, and it becomes more difficult to determine who makes the decisions (that were taken) in the first place.[289]

Historically, diplomacy has been conceptualized around a model of agency that has been amenable to institutional accountability, where decisions are the outcome of deliberative processes that can be traced to certain individuals or organizations. Political responsibility in this model is strictly interconnected with the possibility of defining who makes this decision and assessing the action. The advent of artificial intelligence, however, changes this dynamic, with the algorithm systems playing an important role in the generation of the information and suggestions on which decisions to rely.[290]

The result of this transformation is what is commonly termed in the literature as a diffusion of responsibility or responsibility gap, for example, a scenario in which it is hard to establish a clear line of responsibility when making decisions. With the existence of AI systems, in any case, the process of making decisions cannot be ascribed to a specific actor anymore, but to several different actors, such as developers, institutions, and the final users. In this regard, the OECD stresses that traceability and accountability of human actors are crucial to prevent the risk that automation may cause accountability to be lost.[291]

Among the most crucial points is how to tell the difference between decision support systems and automated decision-making systems. In the former, artificial intelligence can offer advice or insights that can assist the human decision-maker, whereas in the latter, the algorithm can play a

[288] European Commission (2020), p, 2; OECD (2019/2021), p, 4.
[289] European Commission, *White Paper on Artificial Intelligence* (2020), p, 2.
[290] M. L. Cummings et al., *Artificial Intelligence and International Affairs* (2018), pp, 1–2.
[291] OECD, *AI Principles* (2019/2021), p, 4.

more prominent role in the decision-making process. Nevertheless, the threat of the so-called automation bias is also high even in the support systems: the decision-makers can be pushed to trust the algorithmic recommendations too much, thus allowing the decreased critical and deliberative role.[292]

Such a phenomenon is especially applicable to the diplomatic scene, where decisions are frequently forced to be made in ambiguity and have restricted time limits. Within these settings, the fact that AI systems can process huge volumes of data in a much shorter time can prove to be both an asset and a drawback because it also poses a high chance of leaving the decision-makers subject to mindless consumption of algorithmic results. This may result in a progressive handing over of decision-making with huge repercussions regarding accountability and legitimacy. This problem is directly addressed within the ethical guidelines of the European Union, which focus on the principle of human oversight that can be viewed as a central aspect in making sure that AI systems are bound to serve humanity as a guide in terms of making sure that artificial intelligence can be used as a means of correcting or reversing automated decisions.[293]

Nevertheless, the reality of human watchdogs presents itself with a lot of challenges. The velocity with which information is handled and the amount of information involved in complex contexts, like in diplomacy, may complicate decision-makers ' ability to effectively maintain control of algorithmic systems. Additionally, the ever-evolving nature of AI technologies complicates comprehension of the inner workings of these systems, thus constraining the possibility of controlling them.[294] It is this efficiency and control dilemma that is one of the most vital facets of including AI in the decision-making process. Artificial intelligence, on the one hand, presents potent mechanisms to enhance the efficiency of the decisions and their speed; on the other, it creates the novel layers of transparency and reliance on technology that may undermine the capacity of decision-makers to maintain effective control.[295]

Within the framework of digital diplomacy, the matter of accountability assumes even more complicated aspects. As Manor remarks, communication and perception management are now the key actors of modern diplomatic action. In this context, we are dealing not only with foreign policy as it is traditionally understood, but also with the establishment of discourses and the formation of

[292] Cummings et al., pp, 2–3.
[293] High-Level Expert Group on AI, *Ethics Guidelines* (2019), p, 3.
[294] European Commission (2020), p, 2.
[295] Ivi.

global opinion.[296] The deployment of the AI systems in this area can enhance the reliance of the decisions made, yet this also complicates pinpointing any adverse consequences. As an illustration, managing diplomatic communications with the help of algorithms may affect what people perceive about international events, yet it is complicated to define who should be held responsible, as a number of actors and decision-making tiers can be involved.[297] What makes the nature of artificial intelligence governance is further compounded by the multilevel nature of artificial intelligence.[298] The responsibility is no longer unilateral to the national level but has been expanded to the international and transnational level, where international organizations, states, and even private actors are involved. In this regard, it is important to have proper accountability mechanisms that are coordinated at various levels of governance and a proper definition of responsibilities.[299]

The White Paper of the European Commission underlines that there is a necessity to create the necessary regulatory framework that could be aimed at guaranteeing the responsible use of artificial intelligence, and it should also balance the role of innovation and regulation, which, at the same time, cannot be easily imposed due to the rapid pace of technological development.[300]

Moreover, the development of artificial intelligence is assisting the substitution of the power balance in the international system, and its direct impacts are related to political responsibility. Dev emphasizes that AI is also setting the concentration of power into the hands of actors controlling technological infrastructure and data, generating new types of asymmetries among states.[301] Accountability, in this context, is not only about the policymakers but also about the power structures that regulate access to technologies.[302] The key to focusing on the data is another level of complexity. The UNCTAD report emphasizes the extent to which data has become a core strategic asset, and it has some big consequences in terms of global governance.[303] But, data management also presents the accountability question, especially when it comes to the security of information and avoiding abuse. [304] In terms of this, there arises a core conflict between efficiency

[296] Ilan Manor, *Digitalization of Public Diplomacy* (2019), pp. 4–5.
[297] Corneliu Bjola, "Digital Diplomacy as World Disclosure: The Case of WikiLeaks," *International Studies Perspectives* 16, no. 3 (2015): 315.
[298] European Commission 2020, p. 2; OECD 2019/2021, p. 4; Dev 2025, pp. 1–2
[299] Ivi
[300] European Commission (2020), 2.
[301] Rishik Dev, *AI World Order* (2025), pp, 1–2.
[302] Ivi.
[303] UNCTAD, *Digital Economy Report* (2021), pp, iv–v.
[304] Ivi

and legitimacy. Artificial intelligence enables the enhancement of the effectiveness of the decision-making process, yet it also endangers transparency and accountability. Such a tension can be viewed as one of the primary problems of modern diplomacy, which needs to establish the equilibrium between technological advancement and democracy.

Lastly, it is necessary to stress that the problem of accountability is not a technical or legal issue only, but a highly political one. The assigning of blame means the consideration of values and goals that underlie diplomacy. Given that, artificial intelligence cannot be an apolitical instrument, but rather it should be placed into a larger context of democratic management. Responsibility and accountability: The adoption of the application of artificial intelligence in diplomatic decision-making activities is a very challenging issue.[305] The intricacy of the systems of algorithms, the circulation of responsibility, and the metamorphosis of power relations entail the need to reinvent classical models of governance. The creation of sufficient control and accountability tools will only be able to make sure that the application of the AI action helps consolidate the legitimacy and effectiveness of modern diplomatic action.[306]

3.3 AI, Human Rights, and Democratic Legitimacy

As humanity adopts Artificial Intelligence for its new diplomacy, what about the preservation of human rights and the democratic legitimacy of decision-making processes? AI presents sophisticated tools to enhance the efficiency and analytical power of institutions, but along with this come immense dangers to the right to freedom and the quality of the democratic process. The technology rights and legitimacy dimension effort springs to mind here as one of the most important dimensions of modern governance.[307]

One such debate is about the implications of AI for fundamental rights, including privacy, freedom of expression, and the right to fair treatment. The way AI systems work is to use a vast amount of data, which is often sensitive, and using these systems allows unprecedented information to be collected and processed. It can generate considerable conflict between security requirements and the right to privacy on the part of individual rights.[308]

[305] Cummings et al. 2018, pp, 1–2; European Commission 2020, p, 2
[306] European Commission 2020, p. 2; OECD 2019/2021, p. 4
[307] European Commission 2020, p. 2, European Commission HLEG 2019, pp. 2–3
[308] OECD 2019/2021, pp. 3–4.

These tensions play out in the diplomatic domain in the scope of surveillance and monitoring of information. AI tools in data analysis can have a positive impact on enhancing the detection of threats and the management of international crises, but they can also negatively impact privacy and the freedom of the individual.[309] The European Commission has emphasized in its White Paper the potential risks of AI to fundamental rights, particularly in sensitive or high-impact applications.[310]

The ethical guidelines set by the European Union stress that the development and application of AI must be fully aligned with the human rights of final users, with an "human-centric" approach.[311] But the real principle put into practice is fraught with difficulties. AI creates novel ways of informational power that can lead to shifting an individual's perception and subsequent behavior. However, AI is not just supporting the decision-making cycles, but the context of decision-making. The phenomenon especially gains significance in the context of digital diplomacy. As mentioned by Manor, this has become a defining feature of today's diplomacy, and in the age of AI, it could be done on a more global and direct scale – affecting the nature and quality of democratic debates.[312]

Likewise, for digital diplomacy, the construction of meaning of international events "… can be supported by the introduction of AI systems, at the same time as the risk of construction of narratives through manipulation or disinformation will grow".[313]

The dangers are even more apparent in the era of fake news, deceptive images (deepfakes), and AI-generated fake news and content. The existence of generative AI solutions makes it possible to create realistic synthetic, fake texts, images, sounds and videos at far lower costs and in the blink of a fairly short time, in other words, to generate disinformation in different formats isn't just about fake information but an issue of democratic legitimacy: if citizens, the press and institutions are no longer able to gauge whether information is genuine or manipulated, trust in information and in the decisions made in the public forum is undermined.[314] Examples of this vulnerability can be seen in recent cases. A widely shared example is the use of a deep fake in Karlov's role during the

[309] European Commission 2020, p. 2; OECD 2019/2021, pp. 3–4
[310] European Commission (2020), p. 2.
[311] European Commission HLEG 2019, p. 2.
[312] Manor (2019), pp, 4–5.
[313] Bjola 2015, p. 315 Bjola 2015, p. 315
[314] Mohamed R. Shoaib, Zefan Wang, Milad Taleby Ahvanooey, and Jun Zhao, "Deepfakes, Misinformation, and Disinformation in the Era of Frontier AI, Generative AI, and Large AI Models," ICCA, 2023, pp. 1–3.

2023 Slovak elections as follows: fabricated audio appeared minutes before election day, appearing to be about electoral fraud and buy-and-sell voting, complicating the possibility of verifying and refuting the messages in a delicate moment of the democratic process.[315] The examples illustrate how AI can directly contribute to diplomacy and democracy through its influence on the information environment in which perceptions and trust, and thus political consensus, are developed. The EEAS also considers Foreign Information Manipulation and Interference as a threat to the EU's foreign policy and security because it seeks to polarize society, trick people into losing faith in democratic institutions, and defraud the EU of serving its goals.[316] In 2024, OpenAI acknowledged that it had suspected political influence operations using its models to produce political commentary, articles, fake social media bios, translations, and multilingual content and had halted such activities.[317] NATO has also described information threats – which encompass information operations, foreign manipulation, and disinformation – as having the potential to harm the national security of member nations, and has noted that AI and deepfakes could enhance manipulative tactics to promote misunderstanding and cause further polarization within society.[318] The EEAS also has defined Foreign Information Manipulation and Interference as one of the EU's threats to foreign policy and security, and it has stated that it "may involve introducing polarization, undermining citizens' trust in democratic processes and institutions, or resorting to deceptive use" of its AI models.[319] The dilemma of diplomacy is thus to balance countering information manipulation and protecting freedom of expression. There are multiple elements to the response beyond the mere deletion of misleading information, such as platform transparency, the tracking of synthetic content, interstate cooperation, media literacy, and speedy verification processes. The debate over countering misinformation of all kinds, including

[315] OpenAI, "Disrupting Deceptive Uses of AI by Covert Influence Operations," May 30, 2024, https://openai.com/index/disrupting-deceptive-uses-of-ai-by-covert-influence-operations/
[316] European External Action Service, "Inside the Infrastructure of Foreign Information Manipulation and Interference (FIMI) Operations," March 19, 2025, https://www.eeas.europa.eu/eeas/inside-infrastructure-foreign-information-manipulation-and-interference-fimi-operations_en
[317] NATO, "NATO's Approach to Counter Information Threats," October 18, 2024, https://www.nato.int/en/about-us/official-texts-and-resources/official-texts/2024/10/18/natos-approach-to-counter-information-threats
[318] NATO, "NATO's Approach to Counter Information Threats," October 18, 2024, https://www.nato.int/en/about-us/official-texts-and-resources/official-texts/2024/10/18/natos-approach-to-counter-information-threats
[319] European External Action Service, "Inside the Infrastructure of Foreign Information Manipulation and Interference (FIMI) Operations," March 19, 2025, https://www.eeas.europa.eu/eeas/inside-infrastructure-foreign-information-manipulation-and-interference-fimi-operations_en; European External Action Service, "Information Integrity and Countering Foreign Information Manipulation & Interference (FIMI)," March 17, 2026, https://www.eeas.europa.eu/eeas/tackling-disinformation-foreign-information-manipulation-interference_en?s=2803

that AI-generated, goes beyond merely an ethical concern; it is also a diplomatic issue, and the question of protection of rights ultimately refers to democratic societies and institutions' claim to credibility in the global public sphere.[320]

This results in a deeper shift in the way technology and democracy interrelate, which could be described as "black box democracy".[321] In this case, the opaque algorithmic systems, which have become increasingly significant throughout political decision-making, make it harder for citizens to scrutinize and understand the decisions made. This opacity threatens the principle of transparency that is central to the very democratic process, diminishing possibilities for effective control over institutions.

Additionally, the European Commission's White Paper notes that opacity in the functioning of AI systems can also have negative impacts on trust in the decision-making process, making it harder to guarantee the legitimacy of decisions.[322] The European guidelines on ethics also point out the importance of making systems transparent and explainable to ensure democratic oversight.[323]

The ethical considerations are even more pertinent when considering AI in the context of diplomacy. Indeed, the ethical considerations around AI are even more critical when considering its application within the realm of diplomacy. Internally, at the local level, AI ethics is largely about how to safeguard individuals with respect to privacy, discrimination, lack of transparency about the algorithms behind AI, surveillance, and a few effective remedies. However, the problem extends beyond national boundaries in the diplomatic realm as well: States need to agree on shared common standards of AI development and use, gain trust in each other, avoid increasing technological dominance over others, and make sure developing and advanced societies and companies don't have a monopoly on how to govern AI use. Ethics, therefore, cannot be limited to a catalogue of abstract principles, but rather requires multiple perspectives and cross-border dialogue, moral framing, and legal and policy frameworks.[324] An important example of this ambition to seeking to make ethics accessible on a common language for universal governance is the UNESCO Recommendation on the Ethics of AI. It states that AI should be developed with

[320] UNESCO, "Guidelines for the Governance of Digital Platforms," November 11, 2023, https://www.unesco.org/en/articles/guidelines-governance-digital-platforms?hub=751

[321] European Commission 2020, 2; European Commission HLEG 2019, 2–3.

[322] European Commission (2020), 2.

[323] European Commission HLEG 2019, pp. 2–3.

[324] Luciano Floridi et al., "AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations," *Minds and Machines* 28, no. 4 (2018): 689–707, pp. 689–690.

human rights and human dignity in mind, principles of transparency, fairness, environmental considerations, and human supervision. Moreover, UNESCO places importance on the principle of proportionality, safety, privacy, accountability, explainability, human oversight, inclusion and non-discrimination.[325] On the diplomatic side of things, this recommendation has the advantage of offering a foundation for a set of principles to be adopted by States with varying political, economic and technological systems. However, it has its drawbacks, because it is a soft law instrument: it can inform strategies which influence national policy and suggest steps towards regulatory convergence, but it will not, in itself, ensure the effective implementation of the principles of ethics. The OECD principles also consider this "two-way" nature.[326] Diplomatic aspects of them include the capacity of the states to which they adhere to apply them to public policies, technical standards and oversight.[327] The challenge is not only around what principles should inform the decision of what to do with AI, but in implementation them into different legal and political systems.[328] An issue at the heart of AI diplomacy is how to achieve the interoperability of regulations without the obliteration of the differences among governance models. The same challenge exists with the European guidelines on 'trustworthy AI'. The requirements that must be satisfied highlight the need to consider beyond just technical performance, to consider how the effects of automated systems can be monitored, explained, challenged and corrected, a capability tied to trust of the system.[329] When AI is applied to decipher crises, build scenarios, manage information flows and help inform public decision-making, the tool or instrument of the trade must be compatible with political accountability, institutional transparency, and democratic oversight.[330] But there is also a story in the literature on this: An extensive listing of ethical principles does not necessarily lead to good governance. Hagendorff throws light on another basic problem, namely that many of the ethical guidelines governing AI are not effectively regulated because there is no enforcement mechanism, meaning they can be

[325] UNESCO, "Recommendation on the Ethics of Artificial Intelligence," 2021, https://www.unesco.org/en/artificial-intelligence/recommendation-ethics
[326] OECD, "AI Principles," https://www.oecd.org/en/topics/ai-principles.html
[327] Ivi.
[328] Ivi.
[329] European Commission, High-Level Expert Group on Artificial Intelligence, "Ethics Guidelines for Trustworthy AI," 2019, https://knowledge4policy.ec.europa.eu/publication/ethics-guidelines-trustworthy-ai_en
[330] Anna Jobin, Marcello Ienca, and Effy Vayena, "Artificial Intelligence: The Global Landscape of Ethics Guidelines," *Nature Machine Intelligence* 1 (2019): 389–399, pp. 1–3.

used as instruments for symbolic self-governance.[331] This is especially true of the diplomatic profession, where consensus on values tends to be too weak. The mechanisms of credible international governance will need to include procedures, mechanisms, institutions, monitoring, evaluation, access to remedies and specific responsibilities. The Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law, adopted by the Council of Europe, is undoubtedly an important development as it is declared the first legally binding international document in this regard.[332] A diplomatic shift, from an 'ethics as a statement of principles' approach to an 'ethics as an institutional architecture' approach, suggests the process of ensuring that activities through the entire lifecycle of the AI system are aligned with human rights, democracy and the rule of law, including considerations of human rights, dignity, autonomy, equality, non-discrimination, privacy, transparency, oversight, accountability, reliability, and prevention and mitigation.[333] In addition, the Council of Europe framework calls for remedies and for permanent precautions to replace temporary ones, such as term records and audits. So, the diplomatic battle is now two-fold. States must guard against the emergence of new inequality, exclusion, and technological domination through the use of AI.[334] On the other hand, they must interpret common values into actual actions. Ethics, values, and morality require processes as well as values: an insight from the model proposed by the Alan Turing Institute for the Council of Europe, which connects human rights, democracy, and rule of law to practical tools such as risk analysis, stakeholder engagement, impact assessment, and impact mitigation, access to remedies, monitoring, and assurance cases.[335] AI diplomacy can't just generate joint declarations; it needs to create systems that are verifiable, contestable, and enforceable.

To stay abreast of the issue of legitimacy becomes even more complicated when considering the international sphere, the number of actors involved, and the variety of levels of government. The democratization of AI can exacerbate inequalities of power between states and their access to global decision-making.[336]

---

[331] hilo Hagendorff, "The Ethics of AI Ethics: An Evaluation of Guidelines," *Minds and Machines* 30 (2020): 99–120, pp. 1–2, 4.
[332] Council of Europe, *Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law*, brochure, 2024, pp. 2–3.
[333] Ivi.
[334] Ivi, pp. 3–4.
[335] David Leslie et al., *Human Rights, Democracy, and the Rule of Law Assurance Framework for AI Systems: A Proposal Prepared for the Council of Europe's Ad Hoc Committee on Artificial Intelligence* (The Alan Turing Institute, 2021), pp. 6–13.
[336] European Commission 2020, 2; OECD 2019/2021, 4; Dev 2025, 1–2.

AI is shaping the distribution of power across the globe.[337] It's empowering those with technology and data to take control. This is a deal because it has an impact on the international system's fairness. May exclude countries that are not adept at technological decision-making.[338]

The UNCTAD report highlights that data is also very crucial today, because they show the inequalities between countries.[339] It can become more unfair, and it goes against the international system. The use of artificial intelligence is creating a divide between nations, and this is a problem for all. The "AI divide" is one of the main challenges facing global governance today, as it exacerbates inequalities and undermines the fairness of the international system. Being able to use technology effectively is a problem of inequality, not access to technology.[340] Less advanced countries may find themselves on the losing end, as some countries do not have as many resources to produce and utilize more complex forms of artificial intelligence systems, thereby giving them an edge in the world.

When that's the case, we're not only talking about what is a fair deal in our homeland, we are also talking about what is a fair deal in the world. Decisions on intelligence must be coordinated such that everybody concerned has a voice, and there are no new inequalities created.

European institutions have begun some initiatives to regulate intelligence use to ensure we can continue to have new ideas and also to protect basic human rights.[341] So far, only some steps have been taken in one region, but we need the possibilities of making decisions together as a global community to govern artificial intelligence and its consequences.

The inequality does not only relate to access to technology but also to the use of the available technologies. States that have the means to create and deploy advanced AI technologies become stronger in the international system, and less developed states become marginalized.[342]

In this regard, democratic legitimacy should not only be discussed from a national perspective but should be analyzed globally. The governance of AI needs to find a way that ensures the participation of all interested parties in its development and prevents the creation of new kinds of inequality.[343]

---

[337] Dev, 2025, p. 1, 2.
[338] Ivi
[339] UNCTAD, pp, IV,V
[340] Ivi
[341] European Commission 2020, p. 2
[342] UNCTAD 2021, pp, iv–v; Dev 2025, pp, 1–2
[343] European Commission 2020, p, 2; OECD 2019/2021, p, 4

In Europe, the authorities have tried to introduce measures to control the development and use of artificial intelligence to ensure a balance between technological innovation and the protection of basic human rights.[344] Such an attempt at regulating the deployment of AI from a regional perspective, however, is not enough to deal with the problems caused by such technologies globally.

Another core aspect of such an approach is the notion of human oversight, which strives to make sure that AI-based decisions do not fall outside of the control of human beings.[345] Such an approach is critical in maintaining accountability and ensuring that artificial intelligence does not substitute the decision-making process.[346]

Nonetheless, human oversight alone cannot ensure legitimacy in decision-making. Governance needs to involve additional aspects such as transparency and participation. What is particularly crucial in this case is making sure that people have access to the necessary information about how artificial intelligence is used in the process of decision-making.[347]

There is a key problem related to human rights protection and the democratic nature of governance that is brought about by AI. Although AI can provide substantial benefits in the form of more efficient and analytically oriented organizations, it also poses potential risks that must be considered. The clash between efficiency and democracy is one of the key aspects of the discussion on AI because it indicates that it is important to create an appropriate governance model that would ensure that technology serves fundamental values in the international system.[348]

[344] European Commission 2020, p. 2
[345] European Commission HLEG 2019, p. 3
[346] Ivi.
[347] European Commission HLEG 2019, 3; OECD 2019/2021, 4; European Commission 2020, 2
[348] European Commission 2020, p. 2; OECD 2019/2021, pp. 3–4

Chapter 4 — Governing AI in Diplomacy: The European Approach in a Global Context

4.1 The European Union's AI Governance Framework (AI Act, GDPR)

In recent years, artificial intelligence has become one of the defining technologies driving the transformation of international relations, not only regarding economic competitiveness amongst states, but also in relation to the exercise of political power, electronic governance and in the context of current diplomacy. The spread of more sophisticated algorithmic systems has indeed brought up challenges that extend beyond the technological layer and encompass basic rights protection, security, transparency of decision-making processes, digital sovereignty, and facilitating the role of international players in setting global regulation. Within this context, the European Union has increasingly emerged as an important developer of rules and standards for the governance and regulation of artificial intelligence.[349]

The EU, unlike other international actors, has not seen artificial intelligence just for economic growth or strategic competition. In contrast, the discussion about technology has slowly become part of the European political project of the European institutions. Coinciding with its ambition to regulate the entire AI ecosystem, the AI Act is the most sweeping horizontal and binding law to come out of an emerging field of technology to date, as Ruschemeier and Bareis explain, "an attempt to define a horizontal and binding framework for AI law, which is the first time such an attempt has been made in a burgeoning technology where innovation can simultaneously run alongside the safeguarding of fundamental rights".[350] Considering that, the European model should not only be understood as one of technological regulation, but also as a political project that balances innovation, rights protection and an impact of international regulation.

The General Data Protection Regulation (GDPR), which came into effect in 2018, is regarded as one of the most significant regulatory frameworks during the digital era within Europe. The regulation was enacted prior to the rise of Generative Artificial Intelligence in Europe, but nonetheless, it served as the underpinning for the bulk of the development in AI governance in

[349] H. Ruschemeier, J. Bareis, *Searching for Harmonised Rules: Understanding the Paradigms, Provisions and Pressing Issues in the Final EU AI Act*, in *Digital Decade: How the EU Shapes Digitalisation Research*, Nomos, 2025, pp. 2-3.
[350] Ivi.

Europe. The European Commission considers the protection of personal data to be a fundamental right, expressly protected by EU law, and that this is one necessary precondition to ensure citizens' trust in the digital space.[351]

The GDPR comes with a series of principles to regulate all processing of personal data, from lawfulness of processing, purpose limitation, data minimization, the accuracy of the information processed, to accountability. The principles are especially relevant in the realm of artificial intelligence, as most artificial intelligence systems today are based on the processing of vast data sets during training and operation. The interplay between AI and personal data is, indeed, structural, with more sophisticated algorithms requiring larger volumes of data, and the impact is visibly seen in the output of these algorithms being dependent on the volume, quality, representation and lawfulness of the data they utilize when training them. However, while significant, the GDPR has gradually begun to show some flaws in its ability to control a new generation of systems based on Artificial Intelligence (AI). The purposes of the regulation were improved protection for personal data and not questions like transparency of algorithms, risk classification, human oversight of automated processes, liability, etc. associated with working with complex generative models. This has, however, brought to the fore the need for an accompanying regime that can deal with the unique issues posed by AI technologies as they develop in sophistication. The ability to train on vast amounts of data and fine-tune them for many different purposes has been the major advantage of so-called foundation models and a major driver in the growth of the potential applications of artificial intelligence, but it has not helped the cause of transparency, the reliability of the resulting output, or the accountability of the developer.[352] But with these changes, it has been taking on the importance of data protection rules being not enough to manage the complicated technological world. This evolution of the Artificial Intelligence Act needs to be comprehended exactly through this development. The regulation comes from a lengthy political and regulatory journey designed to create a framework for governing artificial intelligence by managing the risk and safeguarding fundamental rights.[353] The goal that European institutions are setting themselves is not only to mitigate the potential risks of AI, but also to ensure its technological development that will be compatible with the Union's basic values. The European

[351] European Commission, *Data Protection*, https://commission.europa.eu/law/law-topic/data-protection_en last accessed June 18, 2026

[352] R. Bommasani et al., *On the Opportunities and Risks of Foundation Models*, Stanford University, 2021, p. 1.

[353] Ruschemeier, Bareis, *Searching for Harmonised Rules*, pp. 15-20.

strategy rests on one of its pillars, human-centric AI, which is a way for this to be implemented.[354] This vision demands that AI serve humans, while at the same time keeping its application under well-defined regulatory boundaries. Technology has thus not been seen only as a functioning entity that would be left to the forces of the market as they influence its operation, but as a tool that should be managed in a democratic manner through public institutions and mechanisms that allow for public accountability, and with legal protection. European governance of AI, however, is not a solely internal affair. It assumes an ever more prominent geopolitical dimension. From this angle, the European regulation of AI becomes an ever more sophisticated apparatus of foreign policy and a means of disseminating the European values worldwide.[355]

The implementation of the GDPR, then the AI Act, has indeed proven to be a means for the European Union to have some extraterritorial regulatory power. International companies that want to conduct themselves on the market in Europe are often forced to adhere to the European rules that are established in European countries, which leads to the export of European rules abroad. In this regard, governance of AI may be seen as an illustration of EU influence on global governance, as it seeks to shape it indirectly by way of its economy and regulation, the "Brussels Effect" as it is known.[356] It is especially important in the realm of international competition to set the rules for AI in the future that will shape the next decade. Thus, alongside the development of new technologies, the regulatory turn has shifted competition to the global level into a more complex regulatory competition, in which other nations are competing to build advanced technologies as well as establish their own regulatory norms and practices as models for others in the world.[357]

In this setting, the European Union wants to become a ruleset maker whose rules are going to affect how global rules on AI are made in the future. The GDPR and the AI Act can thus be considered as Europe's endeavors to create a united approach to digital governance. The GDPR remains at the heart of protecting personal data and privacy, while the AI Act adds a mechanism specifically geared towards risk management with AI. The interplay and interdependence of both instruments, therefore, are one of the highlights of the European strategy and its vision is embedded in a model

---

[354] European External Action Service, *Digital Diplomacy for a Secure and Competitive Europe*, https://www.eeas.europa.eu/eeas/digital-diplomacy-secure-and-competitive-europe_en, last accessed June 18, 2026.
[355] Ivi.
[356] A. Bradford, *The Brussels Effect: How the European Union Rules the World*, Oxford University Press, Oxford, 2020.
[357] N. A. Smuha, *From a "Race to AI" to a "Race to AI Regulation": Regulatory Competition for Artificial Intelligence*, «Law, Innovation and Technology», vol. 13, n. 1, 2021, pp. 57-84.

that is meant to balance technological innovation, the safeguarding and promotion of fundamental rights and the ability to lead globally on digital security and regulations for artificial intelligence (AI).[358] This aspect makes the AI Act particularly relevant to the study of contemporary diplomacy. The regulation is not only a legal tool at the domestic level, but a tool for normative diplomacy as well. The European Union aims to influence the global discussion on AI governance by providing the international community with high-level standards on transparency, human oversight, data protection, and risk management, and to advocate for a form of innovation that is in line with European democratic values.[359] From this perspective, European regulation can be used as a means of Brussels influencing the future of global rules of the digital ecosystem.[360] So, the European principle of governing artificial intelligence is not simply a regulation of a new technological gadget, rather, it is a global commitment to governance that the European Union is promoting as a normative establishment in the international digital order. The GDPR and the AI Act therefore fulfill a dual role: The protection of the fundamental rights of European citizens, but also as tools of international influence in defining future rules for the usage of artificial intelligence (AI) worldwide.[361]

A key piece of the Artificial Intelligence Act is the use of a risk-based regulatory framework. The European legislature has decided not to modulate obligations, controls, and restrictions for all technologies of AI, but rather based on potentially safety-compromising, dimensioning, fundamental rights-impacting, and collective interests-affecting repercussions of a particular system. While a uniform regulation applicable to the entire pool of AI technologies may seem like the simpler path, the legislators have opted to make out the differences, depending on which

---

[358] European Commission, *Artificial Intelligence Act*, 2024,https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai, last accessed June 18, 2026

[359] European Parliament and Council of the European Union, *Regulation (EU) 2024/1689 Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act)*, *Official Journal of the European Union*, 2024, http://data.europa.eu/eli/reg/2024/1689/oj; European Commission, *Artificial Intelligence Act*, https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai; European Parliament, *EU AI Act: First Regulation on Artificial Intelligence*, 2023, https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence.

[360] Anu Bradford, *The Brussels Effect: How the European Union Rules the World* (Oxford: Oxford University Press, 2020); Nathalie A. Smuha, "From a 'Race to AI' to a 'Race to AI Regulation': Regulatory Competition for Artificial Intelligence," *Law, Innovation and Technology* 13, no. 1 (2021): 57–84, https://doi.org/10.1080/17579961.2021.1898300

[361] Georgios Pavlidis, "The EU AI Act and the Rights-Based Approach to Technological Governance," *Review of European and Comparative Law*, 2026, https://doi.org/10.31743/recl.19283; Hannah Ruschemeier and Jascha Bareis, "Searching for Harmonised Rules: Understanding the Paradigms, Provisions and Pressing Issues in the Final EU AI Act," in *Digital Decade: How the EU Shapes Digitalisation Research*, ed. Gsenger and Sekwenz (Baden-Baden: Nomos, 2025).

consequences receive which treatment.[362] The decision considers the fact that AI is not a uniform technology, but exists in various forms and investments, some of which pose low risk or which may directly impact one's right to employment, education, access to essential services, justice or personal freedom. The AI Act, therefore, adopts a hierarchical approach to the classification of AIs, declaring some prohibited uses and others to be "high-risk" AIs, which must be subjected to the principles of EU rules on AI transparency, as well as those determined to be "low risk" and which are "transparently used solely for entertainment purposes".[363] It enables the European Union to push the purist requirements when they are most difficult to apply on a legal, social, and ethical basis, namely, to applications that are found to be most problematic. The regulation is not meant to discourage technological innovation but is designed to create a suitable regulatory environment to intervene, primarily, in sectors where AI can have an impact on the individual and society. That of systems viewed as an unacceptable risk is the most strictly regulated one. The scope of the ban is quite understandable: some AI technologies and applications do not allow for a mere regulatory framework or transparency obligations without interfering with the fundamental values of the European people or undermining their autonomy, identity, and democracy.[364]

Seeking to maintain its strictly values-based approach, the reference to these practices is included in this list of prohibited practices. The AI Act isn't just a framework for the regulation of a technology market; it's also political and legal boundaries on the application of AI. The regulation of AI is also the regulation of power as its algorithmic technologies have implications on the distribution of power with regards to the creation of authority for decision-making, access to information, and a person's ability to contest the automated decision.[365]

From this point of view, the European regulation has almost a constitutional format, since it tries to establish legal guarantees and controls to reduce the use of algorithmic power. Attention is paid to so-called high-risk systems that constitute the operational backbone of European regulations. They are sectors where the introduction of AI could significantly influence individual opportunities and the enjoyment of fundamental rights, such as in the education sector, employment, managing critical infrastructure, accessing essential public and private services,

[362] Ruschemeier e Bareis, 2025, pp. 9-11.
[363] Artificial Intelligence Act, 2024, artt. 5-50; Ruschemeier e Bareis, 2025, pp. 15-20.
[364] European Parliament and Council of the European Union, *Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)*, Official Journal of the European Union, 2024, art. 5, disponibile all'indirizzo: http://data.europa.eu/eli/reg/2024/1689/oj
[365] Ruschemeier, Bareis, *Searching for Harmonised Rules*, pp. 4-6.

administration of justice, migration, border control, and law enforcement.[366] The regulation sets forth several stringent requirements for such systems. The aim is to avoid the use of systems that no one understands or can control completely in situations where automated decisions may impact an individual's life.[367] It is a clear example of the significance placed upon human intervention, which is one of the primary concerns in this current AI debate. Indeed, the process of making decisions may be increasingly automated, leading to growing risks of ever more contentious decisions being handed to a system that is hard to explain and resist. As per this view, the man-in-the-middle issue through man is not a technical issue, but rather a political and legal cover. It ensures that there is maintained some human accountability in decisions made through algorithms and that decisions themselves do not bequeath decision-making to algorithmic systems in accordance with the human oversight enshrined in the AI Act: Article 14.[368]

Securing equality in algorithmic discrimination is a special kind of issue. Fairness cannot be automated, as European antidiscrimination law makes use of contextual interpretation-based case-by-case assessments.[369] Also, as the authors explain it, algorithmic systems don't need a significant discriminatory purpose or knowledge of discrimination to discriminate in their operation using correlations, classifications, and inferences which may perpetuate or reinforce inequalities. This is because, as the authors show, removing sensitive variables from datasets is not enough to ensure fair results, as other parts of the data can serve as proxies to the protected characteristics.[370] The insights of Wachter, Mittelstadt and Russell are crucial parts of the design of the AI Act. The European regulation doesn't rely on purely technical solutions for self-regulation by AI. Rather, it acknowledges that legal commitments, institutional controls, and accountability mechanisms beyond the technological system are necessary to manage algorithmic risks. In the first place, the European model differs from one more inclined to private self-regulation of technological innovation, thus placing upon public institutions the job of setting limits for technological innovation. Another cutting-edge feature of the AI Act is the oversight of a novel sort of AI model, widely anticipated to be called foundation models or General Purpose AI Models (GPAI). As several generative systems have emerged in the past few months, like ChatGPT, it became evident

---

[366] European Parliament and Council of the European Union, *Regulation (EU) 2024/1689*, artt. 6-7 e Allegato III.
[367] Artificial Intelligence Act, 2024, Art. 9–15
[368] Ivi.
[369] S. Wachter, B. Mittelstadt, C. Russell, *Why Fairness Cannot Be Automated: Bridging the Gap Between EU Non-Discrimination Law and AI*, 2020, pp. 2-8, https://dl.acm.org/doi/epdf/10.1145/3680127.3680200
[370] Ivi.

along the way that certain rules must be drafted for models that are highly systemic, adaptable, and versatile. A key feature of foundation models is that they are designed to be widely extensible, adaptable, and able to power many applications, while having multifaceted impacts on a variety of economic, social, and institutional domains.[371] Two dynamics that are of particular importance may also be emphasized, namely, "emergence" and "homogenization." The first is about the capability that is not expected when the system was designed; the second is the possibility of inclusion of errors, biases or vulnerabilities from the base model replicated by all applications on top of it.[372] These dynamics make foundation models particularly complex to regulate, as their effects are not limited to a single application but can propagate throughout the entire digital ecosystem.

As more countries realize the significance of these foundation models, the European hitmakers have proposed specializations in its proposal concerning general-purpose models. As Ruschemeier and Bareis point out, this move marks one of the most notable innovations in the final draft of the AI Act – it will not only apply to the end products of artificial intelligence, but also to the models which form the technological backbone for the modern AI landscape.[373]

This development is reflected in the provisions dedicated to GPAIs contained in Title VIII of the regulation.[374] In this way, the European Union seeks to address not only the immediate risks posed by individual systems but also the systemic risks arising from technological concentration and dependence on foundational models developed by a limited number of large private actors. This seems especially relevant to the international governance perspective. The foundation models are not only technological products, but are global digital infrastructures. These can be created in one political and economic setting, but then applied to various legal systems, impacting markets, public administrations, political communication, and decision-making processes. Analyzing the decision to regulate them, it is in response to a desire to protect the rights of users as well as a wish to deal with one of the most dominating modern sources of technological strength. But there are also some difficulties with the European model. As outlined by Ruschemeier and Bareis, the logic of product law is mixed in with the substantive rights protection objectives, a tension exists between technical

[371] R. Bommasani et al., *On the Opportunities and Risks of Foundation Models*, Stanford University, 2021, p. 1.
[372] Ivi, pp. 129-130; pp. 152-154.
[373] Ruschemeier, Bareis, *Searching for Harmonised Rules*, pp. 20-23
[374] Artificial Intelligence Act, 2024, Articles 51–56.

risk regulation and the substantive rights protection.[375] There's a risk that some of the more complicated issues of basic rights protection will become a matter of technical compliance, losing a bit of their political and social nature. The contribution of technical standards and of the standardization processes is a crucial question that also must be addressed. A considerable portion of the practical implementation of the AI Act will rely on the standards-setting processes, which might prove challenging for those directly impacted by the use of algorithmic systems in access to This is important because, beyond the adoption of a rule, the governance of AI is more dependent on how the obligations are implemented in practice, in the form of technical standards, administrative procedures, or oversight processes. Finally, the AI Act's risk-based framework is one of its most important unique elements of the European framework for AI governance. The European Union is working to create a regulatory framework that balances the aspirations of ensuring democratic values with the need to encourage technological innovation by implementing a combination of risk classification, human oversight, transparency, and some regulatory restrictions of the foundation models, along with a focus on fundamental freedoms. Meanwhile, ongoing debates for and against the sufficiency of state-of-the-art controls, as well as on the need for standardization and on the implementation of the fundamental rights into technical elements, reveal that the AI Act is not an endpoint for this process, but rather a new one, concerning the creation of governance for AI in Europe.[376]

It is impossible to answer the question of how to govern artificial intelligence in Europe without reading between the lines in the GDPR and the AI Act itself. This should also involve an analysis of the extent to which these principles have been put into practice by European institutions in relation to the challenges new technologies are bringing. The case of ChatGPT and Clearview AI are two examples and worth highlighting, in terms of their potential and how it has been tricky to square with existing regulatory tools, and why the European Union felt it was necessary to create a dedicated framework to AI. An important case for consideration in the European discussion about Artificial Intelligence concerns the Italian Data Protection Authority's action on ChatGPT in March 2023. The Italian measure received world media coverage since it marked the first concrete application of the European data protection regulation to a large-scale generative AI model.[377]

---

[375] Ruschemeier, Bareis, *Searching for Harmonised Rules*, pp. 25-28.
[376] Ivi, pp. 28-30.
[377] Garante per la Protezione dei Dati Personali, *Provvedimento del 30 marzo 2023, n. 112. Limitazione provvisoria del trattamento dei dati personali nei confronti di OpenAI L.L.C. (ChatGPT)*, doc. web n. 9870832.

In the decision, the Data Protection Authority criticized the absence of a legal basis for the processing of information used to train this algorithm and noted gaps in the provision of information to users and the weak possibility of verifying the age of minors.[378] In addition, the authority pointed out that the system has the potential to produce false information about real individuals, while noting that the principle of accuracy established in Article 5 paragraph 1 sub-paragraph d) of the GDPR is also violated.[379] The ChatGPT scenario is significant as it illustrates how GDPR could prove structurally short of the needs of generative AI systems. While the regulation succeeded in some ways in pinpointing issues concerning the safeguarding of personal data, it did not include specific provisions to tackle issues like privacy of generative AI models, evaluating systemic risks, humans' oversight of foundation models or accountability in the AI development and deployment lifecycle. That is to say, the GDPR could cover parts of the topic of personal data, but it didn't have a specific structure of its own that could deal with this field of AI. It is precisely this regulation gap that can highlight the need for far wider amounts of regulation. The way to the implementation of the AI Act has been one of the subsequent awareness that data protection has become an anchoring pillar, but is no longer adequate to guarantee the new ways in which artificial intelligence has been developed in recent years.[380]

The John Lewis case involving the use of ChatGPT can thus be seen as a practical example of the necessity of developing innovative regulatory instruments focused on the governance of AI technologies in addition to the GDPR. A second case was the one handled by the Italian Data Protection Authority against Clearview AI, a U.S. company. In 2022, the Italian authority fined the company for twenty million euros for further processing of information about Italian citizens and ordered the deletion of information collected by the company.[381]

Clearview AI had created a facial recognition program with automated photo collection of images from the web and social media. The company had used web scraping tools to create a huge repository of photos, which were then turned into biometric templates that could identify people.[382] This activity infringed on many core principles of GDPR, such as lawfulness, fairness, and

[378] Ivi, parr. 5-11.
[379] Ivi, parr. 8-10.
[380] H. Ruschemeier, J. Bareis, *Searching for Harmonised Rules: Understanding the Paradigms, Provisions and Pressing Issues in the Final EU AI Act*, in *Digital Decade: How the EU Shapes Digitalisation Research*, Nomos, 2025, pp. 15-20.
[381] Garante per la Protezione dei Dati Personali, *Provvedimento nei confronti di Clearview AI Inc.*, doc. web n. 9751362, 2022.
[382] Ivi, parr. 2-7.

transparency in data processing, and it was held by the Data Protection Authority (DPA). The importance of the case is due to the size of the fine awarded and, most importantly, the international aspect of the case. While Clearview AI had no EU presence, the processing of personal data about European citizens made the GDPR applicable by virtue of its territorial scope of application, explicitly referring to Article 3 of the Regulation.[383]

The decision is one of the most visible instances of the European Union exerting its regulatory influence outside its borders, and its tangible results in forging a new international standard. The case has also been acknowledged at the European level, as it is international. As such, the Clearview AI case represents a prime example of how the GDPR can be utilized not as a mechanism for protecting the rights of individual citizens but also as a tool to hold global technology firms to account, even if their actions take place on a global stage and not in Europe. Additionally, the Clearview case bears great significance in anticipation of the issues that are further covered by the AI Act.[384] The mass collection of biometric data, security risks of a permanent surveillance society, and the effect on fundamental rights are some of the main difficulties that have been driving the development of the European legal framework on AI.[385]

The problems that have arisen in the ChatGPT and Clearview AI cases are very commented upon in the academic literature. Wachter, Mittelstadt, and Russell acknowledge that achieving and optimizing for algorithmic fairness is not something that can be achieved exclusively through technical means as algorithms may still have discriminatory effects after significant systemic efforts have been made on the part of their users to ensure their neutrality, fairness, and equity.[386] Using a similar theme, Bommasani et al. emphasize that while the adoption of application-level governance frameworks is important, the introduction of foundation models also means the need for governance frameworks to act at a much lower level: technological governance.[387]

---

[383] Ivi, parr. 8-12.
[384] European Data Protection Board, *Facial Recognition: Italian SA Fines Clearview AI EUR 20 Million*, 2022, https://www.edpb.europa.eu/news/national-news/2022/facial-recognition-italian-sa-fines-clearview-ai-eur-20-million_en
[385] European Parliament and Council of the European Union, *Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)*, Official Journal of the European Union, 2024, considerando 32-38 e artt. 5-6, disponibile all'indirizzo: http://data.europa.eu/eli/reg/2024/1689/oj
[386] S. Wachter, B. Mittelstadt, C. Russell, *Why Fairness Cannot Be Automated: Bridging the Gap Between EU Non-Discrimination Law and AI*, 2020, pp. 2-8, disponibile all'indirizzo: https://dl.acm.org/doi/epdf/10.1145/3680127.3680200
[387] R. Bommasani et al., *On the Opportunities and Risks of Foundation Models*, Stanford University, 2021, pp. 129-130; pp. 152-154.

The GDPR and the AI Act in this context should be regarded as parts of a comprehensive digital governance framework.[388] The GDPR continues to be the central regulation in Europe to safeguard personal data and privacy because of its foundational nature.[389] However, the AI Act also provides obligations specifically geared toward the management of risks associated with the use of artificial intelligence.[390] Another of the key issues for European institutions is, as was already revealed by the debate around the so-called Digital Omnibus, to ensure that the various digital regulation acts are coordinated and that rules are not applied in parallel to avoid any overlap and are not implemented inconsistently.[391] The GDPR and the AI Act are therefore two of the most innovative aspects of the European strategy now. The former safeguards personal data and individuals' rights, whereas the latter debates more complex matters of transparency, man oversight and management of system risks.[392] These tools are all working towards the development of a governance model that seeks to balance technological innovation, protection of fundamental rights, and digital security. The EU Governance approach is not of course exclusively European and does not start with the institutions of the European Union. It falls in the wider European regional context where the Council of Europe is also playing a key role in establishing international standards based on human rights.[393] In particular, the Framework Convention on Artificial Intelligence and Human Rights, Democracy, and the Rule of Law bears witness to a key step forward in expressing the commitment of the European Union to champion a legal framework for the regulation of artificial intelligence that goes beyond the internal market and will impact the international debate.[394] In this view, the problem of European AI governance is non-neutral. In this light, European AI governance becomes a clear diplomatic problem. The AI Act isn't just a building block of internal

[388] European Parliament and Council of the European Union, 2016; European Parliament and Council of the European Union, 2024
[389] European Parliament and Council of the European Union, 2016; European Commission, "Data Protection"
[390] European Parliament and Council of the European Union, 2024; European Parliament, "EU AI Act: First Regulation on Artificial Intelligence"
[391] *Digital Omnibus e AI Act: verso un sistema di governance integrato?*, Diritto Bancario, 2025, disponibile all'indirizzo: https://www.dirittobancario.it/art/digital-omnibus-e-ai-act-verso-un-sistema-di-governance-integrato/
[392] European Parliament and Council of the European Union, 2016; European Parliament and Council of the European Union, 2024
[393] Council of Europe, *Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law*, 2024, https://www.coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence; Vladislava Stoyanova, "Council of Europe Framework Convention on Artificial Intelligence: Context, Regulatory Approach and Scope of Obligations," *European Journal of Risk Regulation* 17 (2026): 536–565, https://doi.org/10.1017/err.2025.10070.
[394] Council of Europe, *Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law*, 2024.

control, but also a way for the European Union to gain influence in the international discourse on how to govern artificial intelligence.[395] In setting high standards for regulation, Brussels seeks to encourage an approach to AI that is based on the respect and intrinsic worth of human rights, accountability, transparency, and human oversight.[396] Regulation is then a form of 'regulatory diplomacy' that will enable the European Union to contribute to the global rules of the future of artificial intelligence.[397] Indeed, in addition to the debates about fundamental rights, security and data protection, the European discussion is slowly starting to incorporate questions regarding the environmental impact of AI as well.[398]

As AI systems become more advanced, the study of energy usage has become a growing priority, with data centers drawing in significant attention for their overall energy footprint, and digital infrastructure becoming more sustainable[399]. There is now a more multidimensional character of the governance of artificial intelligence, as observed in its demand for energy use, the influence of data centers, and the sustainability of digital infrastructure.[400] This aspect includes legal and ethical parameters, as well as environmental sustainability targets.[401] This element also enables the while link from the study of the European regulatory framework to future reflection on ensuring

[395] European Parliament and Council of the European Union, *Regulation (EU) 2024/1689 Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act)*, *Official Journal of the European Union*, 2024, http://data.europa.eu/eli/reg/2024/1689/oj; Anu Bradford, *The Brussels Effect: How the European Union Rules the World* (Oxford: Oxford University Press, 2020); Nathalie A. Smuha, "From a 'Race to AI' to a 'Race to AI Regulation': Regulatory Competition for Artificial Intelligence," *Law, Innovation and Technology* 13, no. 1 (2021): 57–84, https://doi.org/10.1080/17579961.2021.1898300

[396] uropean Parliament and Council of the European Union, *Regulation (EU) 2024/1689*; European Parliament, *EU AI Act: First Regulation on Artificial Intelligence*, 2023, https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence; Georgios Pavlidis, "The EU AI Act and the Rights-Based Approach to Technological Governance," *Review of European and Comparative Law*, 2026, https://doi.org/10.31743/recl.19283

[397] Bradford, 2020; Smuha, 2021; Ruschemeier and Bareis, 2025.

[398] Nicolas Alder, Kai Ebert, Ralf Herbrich, and Philipp Hacker, "AI, Climate, and Transparency: Operationalizing and Improving the AI Act," arXiv:2409.07471, 2024, https://arxiv.org/abs/2409.07471; European Commission, *Energy Performance of Data Centres*, https://energy.ec.europa.eu/topics/energy-efficiency/energy-efficiency-targets-directive-and-rules/energy-efficiency-directive/energy-performance-data-centres_en; AI Act Service Desk, *Annex XI: Technical Documentation Referred to in Article 53(1), Point (a)*, https://ai-act-service-desk.ec.europa.eu/en/ai-act/annex-11

[399] nternational Energy Agency, 2025; European Commission, "Commission Adopts EU-Wide Scheme for Rating Sustainability of Data Centres"

[400] International Energy Agency, *Energy and AI*; OECD, *Measuring the Environmental Impacts of Artificial Intelligence Compute and Applications: The AI Footprint*, OECD Digital Economy Papers, no. 341 (Paris: OECD Publishing, 2022); United Nations Environment Programme, *Artificial Intelligence (AI) End-to-End: The Environmental Impact of the Full AI Life Cycle Needs to Be Comprehensively Assessed*, Issues Note, 21 September 2024.

[401] Pavlidis, "The EU AI Act and the Rights-Based Approach to Technological Governance"; Alder et al., "AI, Climate, and Transparency."

environmental protection as part of global governance of AI.[402] This isn't to say that the European approach to artificial intelligence is only reacting to current issues; rather, it's a regulatory approach designed to deal with those challenges.[403] Furthermore, it is a political, geopolitical initiative that is being used by the European Union to project itself in the global digital landscape.[404] Traditionally, the EU has only enforced rulings within its own borders, but the GDPR and the AI Act are intended not just to control its territory but also to help shape the future of international rules for artificial intelligence, further cementing the EU's role as a regulatory power in the face of global competition on the regulation of new technologies.[405]

4.2 Comparing Governance Models (EU, US, China)

The comparative analysis of the AI governance models adopted by the European Union, the United States, and the People's Republic of China illustrates the normative differences and conceptions of the inter-relations of technology, political power, market, and rights that may be observed about how AI is being governed. All three designers are mindful of the strategic significance of AI development, but they speak different regulatory and institutional languages when it comes to dealing with AI challenges. The EU has taken a more rights-based, preventative approach to AI regulation along with a risk management philosophy. In contrast, the U.S. has been more flexible with a philosophy of innovation, competitiveness, and national security, while China is building a model more oriented toward the state, where the regulation of AI is becoming increasingly aligned with digital sovereignty, social stability, and control of the information landscape.[406]

---

[402] Alder et al., "AI, Climate, and Transparency"; Kai Ebert, Boris Gamazaychikov, Philipp Hacker, and Sasha Luccioni, "The Global Landscape of Environmental AI Regulation: From the Cost of Reasoning to a Right to Green AI," 2026, https://arxiv.org/abs/2603.00068

[402] Alder et al., "AI, Climate, and Transparency"; Kai Ebert, Boris Gamazaychikov, Philipp Hacker, and Sasha Luccioni, "The Global Landscape of Environmental AI Regulation: From the Cost of Reasoning to a Right to Green AI," 2026, https://arxiv.org/abs/2603.00068

[403] Ruschemeier and Bareis, "Searching for Harmonised Rules"; European Parliament and Council of the European Union, *Regulation (EU) 2024/1689*.

[404] Bradford, *The Brussels Effect*; Smuha, "From a 'Race to AI' to a 'Race to AI Regulation'"; European External Action Service, *Digital Diplomacy: Secure and Competitive Europe*, https://www.eeas.europa.eu/eeas/digital-diplomacy-secure-and-competitive-europe_en

[405] Ruschemeier, Bareis, *Searching for Harmonised Rules*, pp. 2-3; N. A. Smuha, *From a "Race to AI" to a "Race to AI Regulation": Regulatory Competition for Artificial Intelligence*, «Law, Innovation and Technology», vol. 13, n. 1, 2021, pp. 57-84.

[406] J. Chun, C. Schroeder de Witt, K. Elkins, *Comparative Global AI Regulation: Policy Perspectives from the EU, China, and the US*, 2024, pp. 1-3, disponibile all'indirizzo: https://arxiv.org/abs/2410.21279

The comparison will be organized with a focus on three areas: source of regulatory power, nature of the considered risk, and enforcement tools. Regulatory at the Supranational level is predominantly shaped by supranational law and the guarantee of Fundamental rights in the European scenario, while in the U.S. scenario, by the market, Federal agencies, technical standards and National Security, and in the Chinese scenario, by the possibility of influencing the development of the technology, its administration, and the digital sovereignty. Instead of describing the three systems separately, this allows us to explore the governance of AI as a space of regulatory, economic, and geopolitical competition: This adhering of the three-part was understood, however, not to be taken literally. None of the three models is completely coherent and homogeneous. While fundamental rights will be at the core of its actions, the European Union cannot forget that the path it will have to follow will not drive technological innovation and industrial competitiveness to their detriment. Traditionally linked with a market approach and soft law, the U.S. government is gradually augmenting enforcement and agency powers. Lastly, whilst generally viewed as extremely centralized, China mixes state instructions, local experimentation, industrial growth, and regime (algorithmic) control.[407] Among all these complexities, when it comes to governance, however, the three models clearly show precedent of three different ways of justifying governance: regulation, competition, and sovereignty.

The U.S. model takes a significantly different approach to AI governance from the first European regulatory framework described in the previous section. Although the European Union depends on a horizontal and legally binding framework based on risk classification and fundamental rights protection, the United States has traditionally followed a more fragmented, flexible and innovation-oriented model. In the absence of a general federal regulation akin to the AI Act, U.S. AI governance is comprised of executive orders, voluntary frameworks, federal agency initiatives, antitrust tools, consumer protection enforcement, technical standards, and sector-specific legislative measures.[408]

This approach suggests an alternative notion of the relationship between the state, the market, and technology. In the United States, artificial intelligence is primarily understood as an engine of

[407] E. Hine, *Governing Silicon Valley and Shenzhen: Assessing a New Era of Artificial Intelligence Governance in the United States and China*, «Digital Society», vol. 3, n. 50, 2024, pp. 2-3, disponibile all'indirizzo: https://doi.org/10.1007/s44206-024-00138-7

[408] Hine, 2024, pp. 3–5

economic competitiveness, technological leadership, and national security. Overall, regulation is not viewed as a hindrance to innovation but as a tool for reducing particular risks, safeguarding consumers, maintaining competition, and defending American dominance in strategic technologies. Hine observed that the United States has traditionally operated within a governance structure of industry self-regulation and voluntary instruments, but in recent years, there has been a tendency towards more concrete public interventions and more effective enforcement mechanisms[409].

The main weakness of the US model is thus not the lack of governance but its fragmentation. Unlike the European model which is based on a unified regulatory framework, the American approach uses a variety of institutions, sectors and policy tools.[410] The National Institute of Standards and Technology (NIST) develops technical standards and risk management frameworks, the Federal Trade Commission oversees competition, unfair practices, and consumer protection, the Government Accountability Office assesses federal agencies' use of technology, Congress deals with certain national security issues, and state courts and legislatures address litigation and sector-specific regulations. The result is a system that is flexible and adaptable, but less uniform than the European model.[411]

A key tool of this strategy is the NIST Artificial Intelligence Risk Management Framework, released in January 2023. The NIST AI RMF is not legally binding but is one of the major technical references for AI risk management in the United States. It is defined as a voluntary, rights-preserving, non-sector-specific, and use-case-agnostic tool to help public and private organizations to develop, deploy, and use AI systems responsibly.[412]

The NIST AI RMF describes several important properties for trustworthy AI systems, including validity and reliability, security, resilience, transparency, explainability, privacy protection, and fairness with harmful bias management. Furthermore, the framework organizes risk management around four main functions: Govern, Map, Measure, and Manage.[413] This architecture is a clear example of the logic of the U.S. model: rather than imposing a strict legal categorization ex ante,

[409] Hine, 2024, pp. 1–3
[410] Hine, 2024, pp. 3–5
[411] Ivi.
[412] National Institute of Standards and Technology, *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*, NIST AI 100-1, gennaio 2023, pp. 1–3, disponibile all'indirizzo: https://doi.org/10.6028/NIST.AI.100-1
[413] Ivi, pp. 12–20.

as is the case with the European AI Act, it offers practical ways for organizations to identify, measure, and manage risks related to the specific context of use.

The NIST AI RMF's voluntary status shows how different the United States is compared to Europe with respect to artificial intelligence (AI) regulation; whereas the EU has the AI Act that creates legally enforceable obligations (and a general prohibition against some types of artificial intelligence), the NIST AI RMF only provides a technical standard and non-binding guidelines.[414] This creates a level of flexibility and adaptability within AI development, but whether it will be effective remains to be seen if private companies have an economic incentive to not protect user rights or prevent larger societal risks.[415]

Moreover, it can be seen in the way generative AI is perceived in the two areas. In fact, the EU sees requirements for all general-purpose AI technologies and those technologies that pose systematic risks to the users, whereas the United States perceives competitiveness, protection of the consumers of generative AI, and the amount of technology power that is concentrated in a select few tech firms as opposed to generative AI regulation. In June 2023, the FTC published an article stating that generative AI depends on critical inputs such as data, specialized labor, and computing power to succeed. It was said that critical inputs being owned by a few firms would lead to obstacles to purchase and distortion of the competitive dynamic in AI markets.[416]

This is where one of the crucial considerations when trying to understand the logic behind U.S. regulation of AI comes from. In addition to being a technology that poses a potential threat to human rights and freedoms, AI is also a rapidly developing industry where there is a cluster of data, cloud computing technology, chips, capital, and expertise that could create uneven playing field conditions. The main concern, therefore, is not merely to ban certain practices, as is the case with the EU approach, but to avoid the monopoly of a few companies on some crucial components of the digitized economy infrastructure. And in this way, combining antitrust and innovation paradigms, one may see that it is not about large models of AI per se, but about having a monopoly on their production/distribution.[417]

---

[414] NIST, 2023, pp. 1–3.
[415] Hine, 2024, pp. 1–3.
[416] Federal Trade Commission, "Generative AI Raises Competition Concerns", 29 giugno 2023, disponibile all'indirizzo: https://search.ftc.gov/policy/advocacy-research/tech-at-ftc/2023/06/generative-ai-raises-competition-concerns.
[417] Ivi.

In November 2023, the FTC (Federal Trade Commission) consolidated its stance regarding the investigative use of compulsory process against products and services alleging to use or detect artificial intelligence when it authorized the use of compulsory process for non-public investigations involving the same.[418] By utilizing civil investigative demands (CIDs), which are like subpoenas, FTC staff can obtain documents, information, and testimony regarding consumer protection as well as competition investigations related to artificial intelligence.[419] Even before the creation of a federal AI Act, the United States can utilize existing enforcement mechanisms in the artificial intelligence industry and has begun exercising this authority through biometric technologies. Specifically, in May 2023, the FTC issued a policy statement detailing the risks associated with biometric information misuse, including privacy concerns, data security issues, and discrimination and biases that can arise from biometric data collected using facial, iris, or fingerprint recognition systems.[420] The FTC also explained the many factors that can cause a given practice to be found unfair under the FTCA; for instance, lack of training for employees responsible for collecting or handling biometric data, the collection of biometric data without the knowledge or consent of consumers (i.e., covert collection), and insufficient oversight over the performance of vendors who provide biometric technologies.[421]

In terms of the differences between how Facial Recognition is handled in the U.S. and the EU, the Facial Recognition Regulation and its use as an immigration enforcement tool have become a major controversy.[422] In the E.U., several laws, including the AI Act and EU Reg. 2250 of 2020, explicitly state that these forms of Cognitive Artificial Intelligence (AI) used for Remote Biometric Identification (RBI) will either be banned in some member countries, regulated in others, or must follow privacy-by-design protocols if used ethically.[423] By contrast, the United States has a fractured regulatory environment regarding facial recognition, with various approaches including technical standards, agency oversight, Federal Trade Commission (FTC) regulation, state-led

[418] FTC, 2023
[419] Ivi.
[420] Ivi.
[421] Ivi.
[422] AI Act Service Desk, "Article 5: Prohibited AI Practices," https://ai-act-service-desk.ec.europa.eu/en/ai-act/article-5; U.S. Government Accountability Office, *Facial Recognition Technology: Federal Law Enforcement Agency Efforts Related to Civil Rights and Training*, GAO-24-107372, March 8, 2024, https://www.gao.gov/products/gao-24-107372
[423] European Parliament and Council of the European Union, *Regulation (EU) 2024/1689 Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act)*, *Official Journal of the European Union*, 2024, http://data.europa.eu/eli/reg/2024/1689/oj; AI Act Service Desk, "Article 5: Prohibited AI Practices."

initiatives for regulating the use of facial recognition, and legal actions against organizations that use facial recognition.[424]

An important example of an organization evaluating facial recognition algorithms developed by private and public organizations and providing a technical framework for evaluating the accuracy or efficiency is the National Institute of Standards and Technology (NIST) Face Recognition Vendor Test.[425] Additional evidence exists that this issue has an institutional element, as reported in the Government Accountability Office (GAO) report. The GAO issued a report in 2024 that highlighted that seven (7) different federal law enforcement agencies operating under the Department of Justice or the Department of Homeland Security have utilized facial recognition technologies in support of criminal investigations, but none of the seven agencies required specific training for their employees on the use of these technologies before using them to support criminal investigations. The GAO also found that not all agencies have developed agency-specific policies and guidelines to ensure civil rights and civil liberties are protected when utilizing facial recognition technologies.[426]

The relationship of AI governance to national security is a crucial aspect of the U.S. model. The TikTok case illustrates this dynamic very well. In 2024, the Protecting Americans from Foreign Adversarial Controlled Applications Act was enacted into law through Congress, as part of an overall legislative package, H.R. 815. The law makes it illegal to provide those distributions, maintenance, or updates to any application that is controlled by a "foreign adversary" in the U.S. unless the application owner performs a qualified divestiture within a designated period of time.[427]

---

[424] National Institute of Standards and Technology, "Face Recognition Vendor Test (FRVT)," https://www.nist.gov/programs-projects/face-recognition-vendor-test-frvt; Federal Trade Commission, "FTC Warns About Misuses of Biometric Information and Harm to Consumers," May 18, 2023, https://www.ftc.gov/news-events/news/press-releases/2023/05/ftc-warns-about-misuses-biometric-information-harm-consumers; U.S. Government Accountability Office, *Facial Recognition Technology*, GAO-24-107372.

[425] National Institute of Standards and Technology, "Face Recognition Vendor Test (FRVT)," https://www.nist.gov/programs-projects/face-recognition-vendor-test-frvt

[426] U.S. Government Accountability Office, *Facial Recognition Technology: Federal Law Enforcement Agency Efforts Related to Civil Rights and Training*, GAO-24-107372, March 8, 2024, https://www.gao.gov/products/gao-24-107372

[427] U.S. Congress, *H.R. 815 – 118th Congress: Making Emergency Supplemental Appropriations for the Fiscal Year Ending September 30, 2024, and for Other Purposes*, 2024, disponibile all'indirizzo: https://www.congress.gov/bill/118th-congress/house-bill/815

The Congressional Research Service has specifically identified TikTok and ByteDance as falling under the definition of "foreign adversarial controlled applications".[428]

The European model is especially interesting for comparison. As TikTok shapes EU regulation under the Digital Services Act, priorities include child protection, addictive design, transparent advertising, and data access for researchers. By contrast, the U.S. government primarily views the threat of TikTok as a national security matter or a challenge to foreign data control and geopolitical influence. The same technological object, namely, its use of algorithmic recommendation systems, is thus conceived differently in terms of both its rationalization and its interpretation. When the EU's rationalization and interpretation of this technological object present it as a systemic risk to users and the digital public space, the U.S. context approaches the analysis of it as a potential risk to national security or information sovereignty.[429] The geopolitical aspect of the U.S. model is further reinforced by the U.S. AI Action Plan released in July 2025 by the White House. The prose is rife with self-aggrandizing language about "winning the AI race," which the U.S. intends to do and believes will lead to economic prosperity, technological competitiveness, and national security.[430]

This strategy echoes the fact that AI governance in the US model is intertwined with geopolitical competition. Risk prevention is not the primary purpose of regulation; it is a tool to help innovation, grow strategic infrastructure, export American technology, and establish international alliances. Among other items is the plan to export full-stack AI packages to allies and partners, expand data center and semiconductor infrastructure, simplify permits for strategic infrastructure, and update federal guidelines for procurement of frontier AI systems.[431]

In this view, the US approach can be characterized as a model of governance based on competitive security.[432] It combines trust in private innovation, intervention by federal agencies, antitrust law, consumer protection, technical standards, and a growing focus on national security.[433] It seems less systematic and less preventative than the European model, and more flexible and geared

---

[428] Congressional Research Service, *Regulation of TikTok Under the Protecting Americans from Foreign Adversary Controlled Applications Act: Analysis of Selected Legal Issues*, LSB11127, aggiornato 9 maggio 2024, pp. 1–2, disponibile all'indirizzo: https://crsreports.congress.gov/product/pdf/LSB/LSB11127

[429] The White House, "White House Unveils America's AI Action Plan", 23 July2025, https://www.whitehouse.gov/releases/2025/07/white-house-unveils-americas-ai-action-plan/.

[430] Ivi.

[431] Ivi.

[432] White House, "White House Unveils America's AI Action Plan"; Hine, 2024

[433] FTC, "Generative AI Raises Competition Concerns"; FTC, "FTC Authorizes Compulsory Process for AI-related Products and Services"; NIST, 2023; White House, "White House Unveils America's AI Action Plan"

towards faster innovation.[434] However, it differs somewhat from the Chinese model in its emphasis on the private sector and the market and its appreciation of the reality that AI has emerged as a new level playing platform in global competition.[435]

To sum up, there is no single alternative regulation to the European AI Act in the United States. Instead, it provides a polycentric model, one where AI governance is a result of the interactions between AI technology companies, independent agencies, Congress, Courts, states, and the federal government. It's a strength of its ability to evolve nicely to the development of technology and to support invention. Its shortcoming is that it's small and has no unified political structure. Innovation vs. sectoral enforcement vs. national security is a quintessential characteristic of the US model relative to the European Union and China.

China's approach to artificial intelligence (AI) governance is distinctly different from that of Europe and the United States. Europe emphasizes protecting fundamental human rights as the basis for developing its approach to AI, while the US uses a flexible approach based on promoting innovation, competition, and national security. The Chinese approach emphasizes digital sovereignty, social stability, and political control over the information ecosystem as the basis for developing its AI strategy. The Chinese approach to AI governance can be viewed as either purely repressive or devoid of an innovation component. Rather, China's use of strong state direction and industrial policy focused on achieving technological competitiveness has resulted in a hybrid model where regulation, security, and economic development are closely interconnected.[436]

Over the past few years, Hine writes, the Chinese governance of AI has increasingly shifted from an era of local and scattered experimentation to more centralized control and institutionalization of the process. China has established certain regulations for these new technologies, such as recommendation algorithms, deep synthesis, and generative AI, indicating a fast pace of regulation in strategic and potentially destabilizing areas.[437] This regulation is not, however, shaped to a greater degree by a justification of fundamental individual rights as it is in the EU. Instead, it's an outgrowth of a more general vision where security, social order, industrial development, and

[434] Smuha, 2021; Chun, Schroeder de Witt, and Elkins, 2024; White House, "Public Comment Invited on Artificial Intelligence Action Plan".
[435] Hine, "Governing Silicon Valley and Shenzhen"; Federal Trade Commission, "Generative AI Raises Competition Concerns"; The White House, "White House Unveils America's AI Action Plan."
[436] J. Chun, C. Schroeder de Witt, K. Elkins, *Comparative Global AI Regulation: Policy Perspectives from the EU, China, and the US*, 2024, pp. 9–11, disponibile all'indirizzo: https://arxiv.org/abs/2410.21279.
[437] Hine, 2024, pp. 9–10.

information control are all significant. This is the first feature of the Chinese model: its state-centeredness. AI is seen as an economic technology rather than just an infrastructure for social governance, national security, or digital sovereignty.[438] This seems to be the case in the fashioning of recommendation algorithms, particularly. Currently, the Algorithmic Recommendation Management Provisions of the Internet Information Services (IRMP), effective from 01 March 2022, are among the world's first efforts to have a direct impact upon algorithmic recommendation services.[439]

The Chinese algorithm regulation does not tie the regulation of platforms to information protection and systemic risks, as is done in the European Digital Services Act (2022–2023), which introduces an information order management system for China.[440]

The guidelines for the use of recommendation algorithms issued by China's government forbid the use of such algorithms to "fake" ranks, popular searches, topic lists, or search results to affect public opinion in the digital sphere or to circumvent the regulatory framework. In addition, the Chinese model will require the algorithmic services with the capability of influencing public opinion or inviting social mobilization of users to be assessed for dangerousness, which involves not only individual harm, but also the extent to which it may contribute to social mobilization or to disrupt political stability.[441] While in the West, the management of recommender algorithms has been associated with the principles of accountability, individual autonomy, and information privacy, in China, it is more connected to cyberspace sovereignty, security, and the rule of law in cyberspace.[442] This distinction reflects how early the Chinese authorities chose to put in place regulations on recommender systems, for both reasons: not just consumers need to be protected from the abuse of recommender systems, but most importantly, to prevent recommender systems

[438] Hine, 2024, pp. 9–10; Chun, Schroeder de Witt ed Elkins, 2024, pp. 9–11.
[439] F. Yang, Y. Yao, "A New Regulatory Framework for Algorithm-Powered Recommendation Services in China", *Nature Machine Intelligence*, vol. 4, 2022, pp. 802–803, disponibile all'indirizzo: https://www.nature.com/articles/s42256-022-00546-9
[440] Ivi.
[441] Cyberspace Administration of China, "Provisions on the Administration of Algorithmic Recommendations for Internet Information Services", 2022, https://www.cac.gov.cn/2022-01/04/c_1642894606364259.htm; traduzione inglese in DigiChina, "Internet Information Service Algorithmic Recommendation Management Provisions – Effective March 1, 2022", https://digichina.stanford.edu/work/translation-internet-information-service-algorithmic-recommendation-management-provisions-effective-march-1-2022/.
[442] U. Reviglio, G. Santoni, "Governing Platform Recommender Systems in Europe: Insights from China", *Global Jurist*, vol. 23, n. 2, 2023, pp. 153–158 https://doi.org/10.1515/gj-2023-0013

in the platforms being used as autonomous tools to manipulate public opinions beyond the control of the State.[443]

The TikTok/Douyin case merely further demonstrates the divide in governance models. TikTok is not a platform without its challenges. In the European Union, it has been identified as a source of systemic risks for minors, vulnerable audiences, transparency of advertising and access to data for researchers.[444] In the U.S., TikTok is, on the other hand, portrayed as a risk to the nation's safety because of ByteDance's ties with China, reflected in the adoption of the Protecting Americans from Foreign Adversary Controlled Applications Act and the Congressional Research Service (CRS) analysis.[445] Its twin platform Douyin, however, is operating in a different environment, in which algorithms for the recommendation system are required to be filed, checked, and comply with the Internet Information Service Algorithmic Recommendation Management Provisions in China.[446] Three regimes for regulating the same type of algorithmic infrastructure are also in place; all in their respective contexts as a systemic risk for users in the EU, a geopolitical risk in the United States, and an informational issue in China.[447] The second core theme of the Chinese model relates to ‘generative' AI. The Interim Measures for the Management of Generative AI Services (Effective August 15, 2023) is China's first specific regulation on generative AI services.[448] The draft directives cover all instances of generative AI services provided to the public in the Chinese People's Republic, regarding, for example, the legality of data sources, protection of IP rights, treatment of personal data, quality of training data, security, and moderation of content.[449]

In comparison with the draft, the final version of the measures can be understood to be more open, and the focus is shifted towards avoiding the deadlock of innovation. Comparative AI provides a reminder that the final version focuses on services to the public in China and doesn't cover as much internal research, development, and corporate use. But it is significant that even in so highly regulated a model as China's, the commitment has been renewed to achieve some level of certainty about data, in this case, by providing increased scope for implementing effective measures to

[443] Ivi.
[444] Commissione europea, 2024; Digital Services Act, Regolamento (UE) 2022/2065, artt. 34–40
[445] Congress.gov, 2024; CRS, 2024, pp. 1–2
[446] CAC, 2022; Yang e Yao, 2022, pp. 802–803.
[447] Reviglio e Santoni, 2023, pp. 153–158
[448] Comparative AI, 2023; Hine, 2024, pp. 9–10
[449] Comparative AI, “Interim Measures for the Management of Generative AI Services”, 2023: https://comparativeai.org/en/rules/china/generative-ai-interim-measures/.

improve data quality, data accuracy, objectivity, and diversity.[450] China would like to have a balance between security and development so that its companies can compete in the generative AI business without many restrictions.

Hine's description of this progression traces a pattern in the overall changes to China's politics, to maintain social stability alongside economic development. The final formulation of the generative AI measures offers an implicit recognition of the limitation and practical difficulty in mandating full accuracy for all output from a probabilistic model, as well as the requirement to create incentive for greater innovation China's governance of generative AI can't be just a matter of state control.[451] It's also an industrial technology designed to drive competitiveness in the generative modelling segment, but with high levels of political oversight over content and services provided to the public.[452]

The management aspects of this control are supported by the latest data on the registration of generative services. The number of generative AI (GenAI) services registered with China's Cyberspace Administration (CAC) rose to 346 as of March 31, 2025, a figure that underscores the government's efforts to establish a policy and administrative framework for the technology.[453]

According to data, China is developing its own governance style through both broad principles and specific procedures for administering, such as filing, registering, defining methods, publishing registration numbers, and supervising via competent authorities. The institutional nature of this approach, as noted by Hine, follows the trend of centralization and institutional oversight that Hine has characterized as China's AI governance due to its growing use of regulatory and bureaucratic mechanisms that are intended to align technological advancement with government control.[454]

A third area that is relevant is related to facial recognition and biometric technologies. One of the biggest examples of the widespread adoption of facial recognition is China. As for the penetration of FRT applications, both public and private sectors, including public safety, transportation, online payments, access to residential buildings, and hospital services, have become extremely high in China.[455] But the authors also point out that the regulation of facial recognition in China is

[450] Ivi.
[451] Hine, 2024, p. 9-10.
[452] Comparative AI, 2023
[453] China Daily/Xinhua, "346 Generative AI Services Filed with Cyberspace Administration of China", 9 aprile 2025, https://www.chinadaily.com.cn/a/202504/09/WS67f5e234a3104d9fd381e55e.html.
[454] Hine, 2024: 9-10.
[455] J.-A. Lee, P. Zhou, "FRT Regulation in China", 2023, pp. 242–245, https://doi.org/10.1017/9781009321211.020

‘asymmetrical; it creates a protective framework for personal data and obtains consent in most of the private contexts, meanwhile, it also leaves a wide room for the intervention of the State in cases where public safety is not implicated or raised.[456]

This is an important asymmetry to compare with the figures from the EU and the USA. S. regulatory frameworks, such as the EU AI Act, do not permit real-time, remote biometric recognition in public areas for law enforcement use, unless there are certain exceptions.[457] U.S. regulation is diffuse, and includes the FTC, GAO, NIST, State regulations, and litigation. In China, however, protection of biometric data exists in parallel to the robust exception for security.[458] While the ability to protect individual rights and the need for consent may be recognized in regulations, public safety considerations may substantially limit the implementation of those protections.[459]

The most recent regulatory initiatives confirm this tension. The rules draw a distinction between collecting facial data with consent and with reasonable and affordable alternatives and clearly prohibit the use of facial recognition technology to force anyone to verify their identity.[460] In this instance, as in the previous, the Chinese model has a structural feature: more trust for the relationship between private individuals and public utility users, and more freedom for the State in relation to security and social control.[461]

Numerous factors account for this being a Chinese model, so it can be considered an instance of algorithmic sovereignty. AI systems must be safe or reliable, and this goal is extended to ensure that they are operated within the state's political goals. Personal data protection, user protection, and transparency are not lacking, but are set within a larger framework where social stability, national security, and information control are of primary importance for the country.[462] In this respect, the Chinese approach is completely different from the European. Both are regulatory and interventionist, though the EU is going about its algorithmic regulation with the fundamental rights angle in mind, whereas China will take the same approach with algorithmic power to fall under

[456] Ivi, pp. 245–251.
[457] Regolamento (UE) 2024/1689, art. 5.
[458] Lee e Zhou, 2023, pp. 245–251
[459] Ivi.
[460] Reuters, “China Says Facial Recognition Should Not Be Forced on Individuals”, 21 march 2025, https://www.reuters.com/technology/china-says-facial-recognition-should-not-be-forced-individuals-2025-03-21/.
[461] Ivi.
[462] Reviglio e Santoni, 2023, pp. 153–158; Hine, 2024, pp. 9–10.

the governance of the state.[463] In making this comparison, it is possible to therefore acknowledge the three different kinds of power that prevail in the current digital order: the power possessed by the EU, the power wielded by the U.S., and the power that can be realized by China. A key power of the European Union is its regulatory power. The concept of the GDPR, AI Act, and Digital Services Act is designed to create transferable rules affecting the actions of global companies and corporations in the EU market.

The United States exercises innovative and infrastructural power. They are leveraging the leadership of tech firms, the resources to access capital and talent, command of key components in frontier technologies like clouds, chips, and frontier models, and enforcement mechanisms in the future within the sectors. At the end of the day, however, China has its sovereignty and administration. It aims to provide guidance for technological progress in alignment with the state's priorities and to mitigate the potential for social unrest or compromise of information control through the use of platforms, algorithms, or generative models.[464] These differences have important consequences for the international governance of AI. They show that there is no politically neutral model of regulation for AI, and that each one involves distinct political values, institutional setups, and geopolitical interests. While tensions between the United States and China, and the increasingly pluralistic formation of coalitions around 'democratic values', might make it difficult to develop forms of international cooperation when it comes to AI, it is nevertheless possible to imagine a pluralistic framework based on shared minimum standards.[465]

As demonstrated above, data collected from multiple datasets have corroborated that AI Governance has become one of the major dimensions of international rivalry among countries, with all three countries relying on this type of technology for the establishment of, and control over, the sources of Economic, Political, and Symbolic Power associated with their respective countries. The European Union wishes to eventually convert its regulation of AI Governance into normative regulatory systems with greater international influence; on the other hand, the United States would like to retain its position as the dominant technology leader by pursuing innovative technological advancements through research and development, as well as through use of its market size and national security; and finally, the People's Republic of China will utilize its

[463] Reviglio e Santoni, 2023, pp. 153–158
[464] Hine, 2024, pp. 9–10; Chun, Schroeder de Witt ed Elkins, 2024, pp. 9–11
[465] Hine, 2024, pp. 1–3

governance capabilities of AI as means of both creating and maintaining digital sovereignty, reinforcing political stability and administrative control. The differences among these three countries in terms of the dominant models for regulating this technology are, therefore, critical when establishing the basis for examining the potential limitations of global interactions related to AI governance, as well as the potential for creating multilateral governance systems associated with AI in the remainder of the article

4.3 Limits of Global AI Governance and Multilateral Coordination

The EU, the U.S., and China provide a stark contrast in the range of political and regulatory approaches for AI governance as this article shows. While the EU tends to focus on an approach which claims fundamental rights, preventive regulation and risk management, the United States follows a more polycentric governance model on the fundament of innovation, competition, and national security, and China embraces a governance model centered on the state, in which AI control is tightly coupled with digital sovereignty and social stability. The cross-over nature of these pieces complicates the building blocks of global governance of AI. There is more to it than simply differencing technical standards: political compatibility between systems with different priorities for the rights of others and markets, security, sovereignty, and technological development must also be addressed. Over the past few years, there has been an increasing number of multilateral projects, political statements, codes of conduct, ethical recommendations, and instruments of cooperation on AI. The international community's increasing awareness of the issue of AI is evident in these efforts, which are increasingly moving beyond the scope of mere technological innovation and domestic policy regulation to become a global concern. They signal, however, that the current international framework on AI governance has one significant weakness: there is a lack of binding rules, enforcement mechanisms, and effective coordination among States in implementing consensus concepts important for good governance, such as safety, reliability, transparency, inclusion, and respect for human rights. The adoption of the United Nations General Assembly resolution A/RES/78/265, titled “Seizing the opportunities of safe, secure, and trustworthy artificial intelligence systems for sustainable development” on March 21, 2024, was a key step. The resolution acknowledges that safe, secure, and trustworthy AI systems could help drive progress on the Sustainable Development Goals, while also noting that the potential privacy

risks, threats to information integrity and discriminatory impacts are among those risks posed by the misuse or malicious use of AI, and the role of AI in deepening inequalities and digital divides.[466]

The impact of the UN resolution lies in its perspective on the convergence of AI and wider issues of sustainable development, human rights, and international collaboration. The advancement of AI is not just being seen as a technology to be defended or claimed, but rather a global infrastructure that could shape the distribution of economic opportunities, access to knowledge, developing countries' participation, and safeguarding fundamental freedoms. The resolution aims to bridge the digital divides and enhance the effective engagement of developing countries in international governance mechanisms on AI.[467]

This is a key aspect of the ability of global governance to constrain AI. Whilst the international community could converge on principles, such convergence can be incomplete because in many cases, it does not take place under the premise of addressing the deep asymmetric nature between those states that can build technological capabilities and those that rely on the technologies of other states, requiring access to certain infrastructure, models, and platforms being developed elsewhere.[468] In this regard, the issue of global governance is at the same time normative and material, in that many States are unable to do so on an equal footing, unless they possess the technological capacity to do so, digital infrastructure, expertise, and access to data.[469]

All this comes through in the Global Digital Compact, which has been agreed to as part of the Pact for the Future in 2024, in this regard, the issues of international cooperation and interdependency in AI governance, the role of developing countries, and avoiding fragmentation of technology spheres are closely interconnected with AI governance being directly related to closing digital divides, inclusion in the digital economy, protecting human rights in the digital sphere, responsible data governance, and preventing fragmentation of technology spheres.[470]

---

[466] United Nations General Assembly, *Resolution A/RES/78/265, Seizing the Opportunities of Safe, Secure and Trustworthy Artificial Intelligence Systems for Sustainable Development*, 21 marzo 2024, pp. 2-3, https://digitallibrary.un.org/record/4040897

[467] Ivi, pp. 3-4.

[468] Ivi, p. 3.

[469] Ivi.

[470] United Nations, *Global Digital Compact*, 2024, disponibile all'indirizzo: https://www.un.org/pact-for-the-future/en/annex-i-global-digital-compact

The Global Digital Compact is of particular importance as it aims to address the persisting situation in which national, regional, and sectoral initiatives are not connected. It acknowledges the need for a balanced, inclusive, and risk-based approach to international governance of AI and equal representation of all countries, particularly developing countries.[471] Additionally, the Compact talks about "compatibility and interoperability" among various organizational approaches for AI governance, directly addressing the issue that arose in confrontations between the EU, the United States, and China: if the main players create a variety of their own frameworks, global governance should do its best to find minimum points of compatibility between different structures.[472]

In addition to the Global Digital Compact, the Secretary-General (SG) of the UN's High-Level Advisory Body on Artificial Intelligence (AI AB) plays an important role. In 2023, the UN Secretary-General called together the advisory group to put together recommendations for international governance of AI, and released the final report, Governing AI for Humanity, in 2024.[473] The goal of the report is to outline the essentials of a more inclusive governance framework for AI (distributed and global) and propose recommendations for how to supplement the existing governance framework.[474]

This body is important because of its multi-stakeholder structure, the group's members represented academia, civil society, the private sector, and government, aiming to develop an approach that would encapsulate the multiplicity of views.[475] This is not to mention, however, the fact that one of the shortcomings in international AI governance, a lack of central authority with binding regulatory authority, still exists in the current international system.[476] The United Nations can provide platforms to have a discussion, to advocate a common set of principles, provide support for cooperation, and to design institutions, but it cannot force a consensus on obligations upon states or global tech providers.[477]

One of the more important moves has been the U.N.'s decision to establish an Independent International Scientific Panel on Artificial Intelligence and a Global Dialogue on Artificial

---

[471] Ivi.
[472] Ivi.
[473] United Nations Secretary-General's High-level Advisory Body on Artificial Intelligence, *Governing AI for Humanity: Final Report*, settembre 2024, https://www.un.org/en/ai-advisory-body
[474] Ivi.
[475] United Nations, "About the UN Secretary-General's High-level Advisory Body on AI," United Nations, accessed June 21, 2026.
[476] United Nations Advisory Body on Artificial Intelligence, *Governing AI for Humanity: Final Report* (New York: United Nations, September 2024).
[477] United Nations, *Charter of the United Nations*, arts. 10 and 13.

Intelligence Governance. These are two important provisions because they reflect society's efforts to move beyond the development of a common set of principles towards the establishment of more robust institutional structures for international cooperation on AI.[478]

The Scientific Panel can contribute to mitigating information asymmetries among actors, including States, technology companies, and civil society, by offering independent analysis of the impacts of AI and its social, economic, and institutional changes.[479] The Global Dialogue, on the contrary, can provide a more inclusive forum to explore the political, economic, and social implications of AI, avoiding the international governance of AI remaining confined to the international governance forums in the most technologically advanced countries.[480] They can, however, also reflect the constraints of the global governance structure, as they can spread knowledge, organize and facilitate dialogue, but are yet to evolve into an international regulatory authority with binding power.[481] That is, creating a dialogue within institutions is progress, but not enough to make it a real global enforcement regime.[482]

Even before UNESCO's 2021 Recommendation on the Ethics of Artificial Intelligence, there was an ethical aspect to multilateral AI governance, wherein UNESCO recommended that all member states uphold morals in AI governance. With a focus on human rights, dignity, inclusion, sustainability, and responsibility, the Recommendation is one of the first global initiatives to establish common ethical guidelines for AI. UNESCO has called out that it has developed tools like the Readiness Assessment Methodology, which can guide countries in building their capacity to implement AI ethically and responsibly, based on this Recommendation, adopted by 193 countries in 2021.[483]

Aside from the aspect of global security concerns raised by the challenge of developing AI technologies, the Recommendation also shows that global AI governance is rooted in a wider

---

[478] United Nations General Assembly, *Resolution A/RES/79/325, Terms of Reference and Modalities for the Establishment and Functioning of the Independent International Scientific Panel on Artificial Intelligence and the Global Dialogue on Artificial Intelligence Governance*, 26 agosto 2025, disponibile all'indirizzo: https://www.un.org/pga/wp-content/uploads/sites/110/2025/12/A_RES_79_325-EN.pdf; United Nations, *AI Panel and Dialogue*, disponibile all'indirizzo: https://www.un.org/global-digital-compact/en/ai

[479] United Nations Advisory Body on Artificial Intelligence, *Governing AI for Humanity: Final Report* (New York: United Nations, September 2024)

[480] United Nations, "AI Panel and Dialogue," *Global Digital Compact*.

[481] United Nations, "FAQ," *Global Dialogue on AI Governance* https://www.un.org/global-dialogue-ai-governance/en/faq?utm

[482] United Nations, *Charter of the United Nations*, arts. 10 and 13

[483] UNESCO, *Recommendation on the Ethics of Artificial Intelligence*, 2021; UNESCO, *Global AI Ethics and Governance Observatory*, disponibile all'indirizzo: https://www.unesco.org/ethics-ai/es

consideration of the values that should inform innovation. Fosters a shared moral, human rights and responsibility vocabulary. Even here, however, the results have been limited as the recommendation carries with it strong political and normative implications but is not a binding international treaty. It has potential for informing national and regional policy-making processes, but its realization would rely on the extent to which the states are ready to embrace it and implement it, and the extent of the implementation of the ethical principles into operational rules.
The G7 Hiroshima AI Process is also a major area of multilateral coordination. In 2023, G7 leaders endorsed, as Principles, the International Guiding Principles for Organizations Developing Advanced AI Systems and the International Code of Conduct for Organizations Developing Advanced AI Systems. [484] While the Guiding Principles are intended to meet worldwide standards and are addressed to all actors developing and using the most advanced AI systems, including 'foundation models' and 'generative AI systems', the Code of Conduct is targeted to organizations developing the most advanced AI systems and is a voluntary guideline that asks these organizations to take a risk-based approach across the AI system lifecycle.[485]
The benefits of the Hiroshima Process are to directly engage the actors who are creating new and stimulating technologies.[486] The G7 is one of the more focused approaches, in contrast with other less specific ones taken by the United Nations or UNESCO, to tackle a specific issue: the management of advanced systems and foundation models, i.e., technological infrastructures that, in their use, can bring transnational and systemic impacts. Its flaws are two-fold, though.[487] On the other hand, it is a process dominated by a few advanced economies and, hence, only partially representative of the international community.[488] Concurrently, the code of conduct has a

[484] European Commission, *Hiroshima Process International Guiding Principles for Organizations Developing Advanced AI Systems*, 30 ottobre 2023, disponibile all'indirizzo: https://digital-strategy.ec.europa.eu/en/library/hiroshima-process-international-guiding-principles-advanced-ai-system
[485] [10] European Commission, *Hiroshima Process International Code of Conduct for Organizations Developing Advanced AI Systems*, 30 ottobre 2023, disponibile all'indirizzo: https://digital-strategy.ec.europa.eu/en/library/hiroshima-process-international-code-conduct-advanced-ai-systems
[486] G7, *Hiroshima Process International Code of Conduct for Organizations Developing Advanced AI Systems*, October 30, 2023
[487] European Commission, "Hiroshima Process International Guiding Principles for Advanced AI System," October 30, 2023; European Commission, "Hiroshima Process International Code of Conduct for Advanced AI Systems," October 30, 2023
[488] OECD, "OECD Launches Pilot to Monitor Application of G7 Code of Conduct on Advanced AI Development," July 22, 2024.

'voluntary' character: organizational membership and pressure of reputation are used to make it effective without binding legal measures.[489]

The OECD's effort to monitor the G7 Code of Conduct also reveals this dichotomy. Using the OECD's Global Partnership on Artificial Intelligence as an umbrella, the OECD began a pilot phase to assess an accountability framework for tracking compliance with the actions specified in the Hiroshima Code in July 2024.[490] However, the framework will still be based on voluntary compliance and will not be legally binding.

The Bletchley Declaration was taken from the UK's AI Safety Summit in November 2023, and it developed the topic of advanced systems safety. The declaration highlights, among other risks, those of frontier systems, which are models capable of undertaking very broad tasks and have the potential to cause significant or even catastrophic harm if used incorrectly or run amok.[491] The declaration also addresses many AI risks that are essentially international in nature, meaning they must be dealt with through cooperation between states.[492]

The Bletchley Declaration is significant because it represents the development of an AI security diplomacy among the international community. It provides an opportunity to focus not on general ethical principles but on the danger to frontier systems (sometimes including systems leading to damage), e.g., cybersecurity, biotechnology, disinformation, and possible system damage. The statement, at the same time, is a political document and not binding. It creates the discursive convergence, no legal obligations, sanctions, and international oversight authorities.

The Seoul Declaration for Safe, Innovative, and Inclusive AI, adopted at the next AI Seoul Summit in May 2024, aimed to build on this path to ensure that AI is used responsibly and ethically.[493] It also encourages the establishment or expansion of AI safety institutes, as well as collaboration among them, potentially fostering a network of AI safety technical expertise and safety evaluations

---

[489] European Commission, "Hiroshima Process International Code of Conduct for Advanced AI Systems," October 30, 2023; OECD, "OECD Launches Global Framework to Monitor Application of G7 Hiroshima AI Code of Conduct," February 7, 2025

[490] OECD, 2024

[491] UK Government, *The Bletchley Declaration by Countries Attending the AI Safety Summit, 1-2 November 2023*, disponibile all'indirizzo: https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023

[492] Ivi.

[493] UK Government, *Seoul Declaration for Safe, Innovative and Inclusive AI by Participants Attending the Leaders' Session: AI Seoul Summit, 21 May 2024*, disponibile all'indirizzo: https://www.gov.uk/government/publications/seoul-declaration-for-safe-innovative-and-inclusive-ai-ai-seoul-summit-2024/seoul-declaration-for-safe-innovative-and-inclusive-ai-by-participants-attending-the-leaders-session-ai-seoul-summit-21-may-2024

that is transnational.[494] Seoul marks an important development in this regard as it looks forward to the creation of a joint understanding of artificial intelligence safety through joint collaboration between institutes, exchanges on the information of the models and products, monitoring of incidents, the creation of a common database and analyses of the phenomena related to AI, and the promotion of international standards.[495]

This is alongside the Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law of the Council of Europe, a new and more legally binding development in this landscape, which is dominated by soft law instruments. It will be open for signature in September 2024 and described as the first legally binding international agreement on Artificial Intelligence.[496] The vision is to ensure that activities performed throughout AI system's life cycle are in line with human rights, democracy and the rule of law.[497]

The Convention is not like political statements like Bletchley or Seoul, which make statements, but it is a treaty law that will impose obligations on the states that ratify it. Nevertheless, it is noteworthy that its scope is limited: it does not create any specific product safety regime that mirrors the European AI Act; it puts in place exceptions for national security, defense, and research; and it leaves Member States with discretion when it comes to regulating the activities of private actors.[498] The Convention, therefore, marks a significant progress in advancing more binding AI governance yet does not cover the issues of enforcement, accountability of private entities or regulation of strategically critical AI applications.[499]

This is a space in which, however, the restriction becomes apparent: the new marketplace of institutions, statements and practices is not a global regulatory body.[500] These widespread efforts

---

[494] Ivi.

[495] UK Government, *Seoul Statement of Intent toward International Cooperation on AI Safety Science, AI Seoul Summit 2024*, disponibile all'indirizzo: https://www.gov.uk/government/publications/seoul-declaration-for-safe-innovative-and-inclusive-ai-ai-seoul-summit-2024/seoul-statement-of-intent-toward-international-cooperation-on-ai-safety-science-ai-seoul-summit-2024-annex

[496] Council of Europe, *Council of Europe Opens First Ever Global Treaty on AI for Signature*, 5 settembre 2024, https://www.coe.int/en/web/portal/-/council-of-europe-opens-first-ever-global-treaty-on-ai-for-signature

[497] Council of Europe, *Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law*, brochure informativa, 2024, p. 2.

[498] Council of Europe, *Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law*, brochure informativa, 2024, p. 4; Vladislava Stoyanova, "Council of Europe Framework Convention on Artificial Intelligence: Context, Regulatory Approach and Scope of Obligations," *European Journal of Risk Regulation* 17 (2026): 545–546, https://doi.org/10.1017/err.2025.10070

[499] Stoyanova, "Council of Europe Framework Convention on Artificial Intelligence," pp. 545–546.

[500] Jonas Tallberg et al., "The Global Governance of Artificial Intelligence: Next Steps for Empirical and Normative Research," 2023.

show us that an emerging form of global AI governance is being created.[501] All of these efforts exhibit a common language about AI, such as: "AI will be safe and secure", "AI will be trusted", "AI will be human-centric", "AI will be used responsibly", "AI will be inclusive", "AI will be sustainable", and "AI will respect human rights".[502] Furthermore, the United Nations' system is taking steps toward creating a framework for global coordination on AI, such as with the establishment of the International Scientific Panel and global dialogue on AI governance.[503] Fragmentation, as discussed in the previous paragraph, is not retrievable by the growth of this common vocabulary and the institutional spaces. Contrary to this, it supports its visibility.[504]

While general principles, diplomatic forums, and early institutional arrangements are developing, the state of global AI governance can be said to be interim: there is no global legislation to discipline dominant private actors or to effectively involve countries that are not at the forefront of AI development, and enforcement mechanisms to ensure that unfair strategies in the development and utilization of AI have minimal effect.[505] There is, thus, not only a legal but also a structural restriction regarding global AI governance. These involve geopolitical fragmentation,[506] the AI divide between technologically advanced countries and the Global Majority,[507] the creation of geopolitical concentration of technological power in a handful of large-state corporations,[508] the military applications of AI and autonomous weaponry,[509] and the

[501] United Nations General Assembly, Resolution A/RES/78/265, "Seizing the Opportunities of Safe, Secure and Trustworthy Artificial Intelligence Systems for Sustainable Development," March 21, 2024 https://www.reuters.com/technology/cybersecurity/un-adopts-first-global-artificial-intelligence-resolution-2024-03-21/?utm

[502] G7, "Hiroshima Process International Guiding Principles for Organizations Developing Advanced AI Systems," October 30, 2023; UNESCO, "Recommendation on the Ethics of Artificial Intelligence." https://digital-strategy.ec.europa.eu/en/library/hiroshima-process-international-guiding-principles-advanced-ai-system

[503]United Nations, "AI Panel and Dialogue," *Global Digital Compact* https://www.un.org/global-digital-compact/en/ai?utm

[504] Swati Srivastava and Justin Bullock, "AI, Global Governance, and Digital Sovereignty," 2024. https://arxiv.org/abs/2410.17481?utm

[505] United Nations, *Governing AI for Humanity: Final Report of the High-Level Advisory Body on Artificial Intelligence* (New York: United Nations, 2024), pp. 9–15.

[506] UNESCO, *Global Readiness Assessment Methodology: A Tool of the Recommendation on the Ethics of Artificial Intelligence* (Paris: UNESCO, 2023), pp. 12–15; Mark Coeckelbergh, *AI Ethics* (Cambridge, MA: MIT Press, 2020), pp. 145–150

[507] United Nations Development Programme (UNDP), *Human Development Report 2023/2024* (New York: UNDP, 2024), pp. 87–92

[508] Luciano Floridi, *The Ethics of Artificial Intelligence: Principles, Challenges, and Opportunities* (Oxford: Oxford University Press, 2023), pp. 201–208

[509] United Nations Environment Programme (UNEP), *Navigating New Horizons: A Global Foresight Report on Planetary Health and Human Wellbeing* (Nairobi: UNEP, 2024), pp. 78–82

environmental costs of computational infrastructure.[510] The barriers identify that these obstacles are not with the absence of international initiatives, but rather the differences between consensus on the goals and the ability to implement them with regulatory capacity.[511]

4.4 Towards AI-Augmented Diplomacy

Having examined, in the previous sections, how governance models for AI have broken up and how these are limited by multilateral cooperation, it is important to now look at the level of diplomatic practice. This section adopts another yet related approach to address the question: In what ways could AI change the work of diplomacy in practice? This goes in parallel with section 4.2, in which we saw how the EU, the United States, and China are charting new paths for governance of AI, and with section 4.3, which examined how the world struggles to establish a coherent governance framework for artificial intelligence. Don't view AI-enhanced diplomacy as a strategy for eliminating humans; consider it a marriage of both. Instead, it combines the use of artificial intelligence tools within the diplomatic processes, aiming to become a complement to human capacity for exploring, analyzing, forecasting, communicating, managing, preparing for diplomatic negotiations, and dealing with consular services in general. Rather than AI being a self-starter in foreign policy, from this angle, it is only a device capable of supporting diplomats in facing the growing complexity in high-speed information flow, in a landscape of information and many, many data points, in transnational crises, and in one where technology and international power are increasingly intertwined.

The important issue, therefore, is not whether AI can replace diplomacy, but when it can enhance its human abilities, such as political judgement, responsibility, trust, the sense of context, and uncertainty management, without doing its job in place. The foundation for this approach is in the report of the DiploFoundation, which offers a particularly interesting one between three levels of interaction between AI and diplomacy: AI as a subject of diplomacy, AI as an instrument of diplomacy, and AI as an influence that shifts the field in which diplomacy is carried out.[512] The

---

[510] United Nations General Assembly, *Lethal Autonomous Weapons Systems: Report of the Secretary-General*, UN Doc. A/79/88 (2024); Human Rights Watch, *Stopping Killer Robots: Country Positions on Banning Fully Autonomous Weapons and Retaining Human Control* (2024)

[511] United Nations, *Governing AI for Humanity*, pp. 18–22; Stoyanova, "Council of Europe Framework Convention on Artificial Intelligence: Context, Regulatory Approach and Scope of Obligations," *European Journal of Risk Regulation* 17 (2026): 545–546.

[512] DiploFoundation, *Mapping the Challenges and Opportunities of Artificial Intelligence for the Conduct of Diplomacy* (Geneva: DiploFoundation, 2019), pp. 6-7, https://www.diplomacy.edu/AI-diplo-report

first dimension relates to the negotiation of rules, standards, and principles on AI as discussed in the preceding paragraphs.[513] The other is AI's application to diplomatic work assistance. Lastly, the third is that in the distribution of international power, culture is about to be reshaped by artificial intelligence, affecting economies, security, information, and state capacities.[514]

By dividing the sphere of diplomacy into three parts, the concept of AI-augmented diplomacy can be better contextualized. The ‘augmented diplomacy' is not about AI governance, but about the usage of AI within diplomacy. They are still connected, though: if governments rely on AI systems to make sense of crises, to write and plan negotiations, or even to converse with citizens in other countries, then they have to ask themselves what limits are there on the use of these systems in a regulatory, ethical and democratic way? AI-driven diplomacy, in other words, is in essence inseparable from the question of the governance of AI. The same instruments that help to sharpen the diplomatic analysis can give rise to biases, lack of transparency, over-reliance on technology, and human determinism. This is true for the DiploFoundation report and will be seen again in more recent reflections on the use of AI in shaping foreign policy.[515] A shift from digital diplomacy to AI-enhanced diplomacy is thus a qualitative shift.[516] Primarily, digital diplomacy has been associated with the use of online platforms, social media, digital communication tools, and information technologies to facilitate and streamline diplomatic processes.[517]

In contrast to AI-driven diplomacy, AI-augmented diplomacy entails the science of developing systems that analyze data, generate text, classify documents, interpret patterns, translate content, provide support for predictions, and help with planning movements.[518] It is important to note that digital diplomacy is primarily about altering diplomatic communication routes, whilst augmented diplomacy alters the cognitive and decision-making processes leading to diplomatic communication, negotiation, and diplomatic action.[519] As a result, there is a need to keep the

[513]Ivi.
[514] Ivi.
[515] Ivi, pp. 25-26.
[516] Katharina E. Höne et al., *Mapping the Challenges and Opportunities of Artificial Intelligence for the Conduct of Diplomacy* (Geneva: DiploFoundation, 2019), 6–7, 14–16.
[517] Corneliu Bjola and Marcus Holmes, eds., *Digital Diplomacy: Theory and Practice* (London: Routledge, 2015), 1–10.
[518] Höne et al., *Mapping the Challenges and Opportunities of Artificial Intelligence for the Conduct of Diplomacy*, 25–29; Muneera Bano, Zahid Chaudhri, and Didar Zowghi, “The Role of Generative AI in Global Diplomatic Practices: A Strategic Framework,” 2023, 2–5
[519] Bano, Chaudhri, and Zowghi, “The Role of Generative AI in Global Diplomatic Practices,” 2–5; Höne et al., *Mapping the Challenges and Opportunities of Artificial Intelligence for the Conduct of Diplomacy*, 25–29.

difference between assisted intelligence, augmented intelligence, and automated intelligence clearly defined.[520] Accompanied AI helps with human work; Augmented AI allows a human to augment their work; Automated AI refers to cases where the entire work is conducted by an AI system.[521] To diplomacy, the distinction means that one is not to go down the road of total automation but rather aid and augmentation.[522] In the current state of technology, in fact, AI would be more valuable to the work of a diplomat as a tool to assist and augment rather than supplant his or her judgment.[523]

It is important to note that diplomacy is not only a matter of information, but rather about understanding the intentions of its actors, reading political signals, developing trust, dealing with ambiguity, and being able to adjust to new contexts.[524] There are a lot of tools that can do this, but no AI can carry the same political and moral load as political and diplomatic decisions.[525] In this sense, augmented diplomacy is to be understood as a kind of human-machine partnership in which there should be a strengthening of the machine's capacity to process information; the responsibility for interpretation and decisions will be kept by the human actors.[526] The field of text analysis is one of the most promising areas.[527] Documents: treaties, memoranda, diplomatic reports and declarations, notes verbale, negotiations documents, press releases, internal briefings, and intelligence reports are closely interconnected with diplomatic practice. A large proportion of diplomatic work relies on unstructured text information, and natural language processing (NLP) tools can enable diplomats to handle the vast amount of text, find links between documents, and

[520] Höne et al., *Mapping the Challenges and Opportunities of Artificial Intelligence for the Conduct of Diplomacy*, 25–26.
[521] Ivi.
[522] Ivi, pp. 26-29.
[523] Luigi Di Martino and Heather Ford, "Navigating Uncertainty: Public Diplomacy vs. AI," *Place Branding and Public Diplomacy* 21 (2025): 350–354, especially 350–352; Höne et al., *Mapping the Challenges and Opportunities of Artificial Intelligence for the Conduct of Diplomacy*, 29.
[524] Luigi Di Martino and Heather Ford, "Navigating Uncertainty: Public Diplomacy vs. AI," *Place Branding and Public Diplomacy* 21, no. 3 (2025): 350–352; Volker Stanzel and Daniel Voelsen, *Diplomacy and Artificial Intelligence: Reflections on Practical Assistance for Diplomatic Negotiations*, SWP Research Paper 1 (Berlin: Stiftung Wissenschaft und Politik, 2022), 7–8.
[525] Stanzel and Voelsen, *Diplomacy and Artificial Intelligence*, 6, 31.
[526] Katharina E. Höne et al., *Mapping the Challenges and Opportunities of Artificial Intelligence for the Conduct of Diplomacy* (Geneva: DiploFoundation, 2019), 25–26; Stanzel and Voelsen, *Diplomacy and Artificial Intelligence*, 31.
[527] Stanzel and Voelsen, *Diplomacy and Artificial Intelligence*, 10, 27; Höne et al., *Mapping the Challenges and Opportunities of Artificial Intelligence for the Conduct of Diplomacy*, 25–29.

save time on analytical preparation.[528] This is significant because merely increasing available information does not equate to more and more knowledge.

In contrast, it creates information overload, can slow down the decision-making process, and can increase difficulty in differentiating relevant information signals from information noise.[529] AI can provide an operational benefit when it helps diplomats ask questions of intricate archives, summarize documents, compare texts, call attention to contradictions, and, or, long since reconstruct the evolution of a negotiation topic.[530] But that's only as long as the data is accurate, the system is clear, the user can accurately interpret the output, and the humans have mechanisms in place to oversee the system.[531]

The example given by the DiploFoundation is the Cognitive Trade Advisor, designed to help diplomats to anticipate international trade negotiations. The system was created to address questions regarding the rules of origin of many trade agreements, which is a rather technically complex subject, filled with a significant volume of legal and negotiating documents.[532]

Now, it is important to state here that this case study is not about having an AI to take the place of diplomats in the negotiations, but about how AI can help in managing the complexity of documents that precede and follow negotiations.[533] Instead of being a replacement technology, AI is a preparatory technology.[534] This difference is also acknowledged by the study of Diplomacy and Artificial Intelligence: Reflections on Practical Assistance for Diplomatic Negotiations by Volker Stanzel and Daniel Voelsen.[535] The authors wonder if AI systems can analyze information in ways that create some sense of added value for diplomatic negotiations sometimes, as in other cases, over traditional methods of analysis.[536] Added value is based on two criteria: Firstly, efficiency:

---

[528] DiploFoundation, *Mapping the Challenges and Opportunities of Artificial Intelligence for the Conduct of Diplomacy*, pp. 26-27

[529] Volker Stanzel and Daniel Voelsen, *Diplomacy and Artificial Intelligence: Reflections on Practical Assistance for Diplomatic Negotiations*, SWP Research Paper 1 (Berlin: Stiftung Wissenschaft und Politik, 2022), 5–6.

[530] Katharina E. Höne et al., *Mapping the Challenges and Opportunities of Artificial Intelligence for the Conduct of Diplomacy* (Geneva: DiploFoundation, 2019), 6–7, 25–29; Stanzel and Voelsen, *Diplomacy and Artificial Intelligence*, 24–29

[531] Elham Tabassi, *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*, NIST AI 100-1 (Gaithersburg, MD: National Institute of Standards and Technology, 2023), 13–16, 22, 40–41; Stanzel and Voelsen, *Diplomacy and Artificial Intelligence*, 6.

[532] DiploFoundation, *Mapping the Challenges and Opportunities of Artificial Intelligence for the Conduct of Diplomacy*, pp. 26-27

[533] Volker Stanzel and Daniel Voelsen, *Diplomacy and Artificial Intelligence: Reflections on Practical Assistance for Diplomatic Negotiations*, SWP Research Paper 1 (Berlin: Stiftung Wissenschaft und Politik, 2022), 27–28.

[534] Stanzel and Voelsen, *Diplomacy and Artificial Intelligence*, 5–6, 31.

[535] Stanzel and Voelsen, *Diplomacy and Artificial Intelligence*, 5–6.

[536] Stanzel and Voelsen, *Diplomacy and Artificial Intelligence*, 5.

being able to evaluate information faster and with fewer resources; Secondly, quality of analysis, the ability to produce strategic insights that will contribute to better chances of successful negotiation.[537] Their analysis is relevant as it is related to the most fundamental aspect of diplomatic work, which is negotiation. It therefore becomes interesting whether AI can help diplomats with administrative and/or communication work, but also in negotiating. The authors lay down their guard with caution. On one hand, they demonstrate the capabilities of AI to support scenario construction, state behavior forecasting, huge data analysis, and automatic tracking of media information during the negotiation process. At the same time, they note that certain aspects of negotiations are not objectively and quantitatively measurable, and that artificial intelligence cannot supplant the judgment of senior and experienced diplomats.[538] This prompting for caution is essential: AI tools can help prepare diplomats, but they don't necessarily reduce the uncertainty, contingency, or relational aspect of diplomatic negotiations.

Multilateral negotiations display this tension nicely, in an environment such as the United Nations, where several states are engaged in several difficult negotiations, diplomats must evaluate interests, voting records, public positions, informal signals, etc., and look at potential coalitions. Stanzel and Voelsen examine, among other cases, the General Assembly discussions about the UN resolution on cybercrime and discuss whether using AI tools can enhance predictions about states' behavior in the General Assembly. Though the findings are vague, the authors note that they suggest the future possibility of AI's role in the management of voluminous amounts of diplomatic documents.[539]

This results in the first conclusion: that AI-assisted diplomacy is not automated diplomacy but diplomatic support with systems of analysis. The augmented diplomat doesn't let AI take over to make political decisions; they leverage it to inform their understanding of the situation and to build some alternate scenarios, to see where there might be potential risk, and to come up with a better strategy, if there is one. In this aspect, AI can reinforce diplomacy if it is grafted onto an institutional environment that avoids entangling it with the politics of the decision-making process.

---

[537] Volker Stanzel e Daniel Voelsen, *Diplomacy and Artificial Intelligence: Reflections on Practical Assistance for Diplomatic Negotiations*, SWP Research Paper 1 (Berlin: Stiftung Wissenschaft und Politik, 2022), pp. 5-6, https://doi.org/10.18449/2022RP01

[538] Ivi, p. 7.

[539] Westminster Brief, *Foreign, Commonwealth and Development Office: Artificial Intelligence — Written Question 73676*, risposta del 16 settembre 2025, https://westminsterbrief.co.uk/archive/pq/73676

The difference is also evident in the initial institutional uses of AI by foreign ministries. A parliamentary answer suggests the British Foreign, Commonwealth, and Development Office has applied AI in translation, document categorization, and analysis in the past year to assist its political and consular staff.[540] Examples of the use of AI in modern diplomacy are focused on what are often described as administrative functions, such as making processes more efficient, improving the organization and filing of documents, or supporting political or consular analysis, showing that the concrete application of AI in current diplomacy is limited.[541]

There is also a more ambitious vision of AI in foreign policy that is evident in the British debate: to make intelligent decisions among imperfect people. But this is exactly the reason that augmented diplomacy is all about: the need to do things faster and more effectively, with more tools at our disposal, including AI but no elimination of risk, no replacement of judgment, and no substitute for building trusting diplomatic relationships, UKAuthority wrote in the same report.[542] Why this is so is because of the volatile nature of diplomacy. As Di Martino and Ford elaborate, the study of public diplomacy is full of epistemic and political confusion, while diplomats are frequently required to operate in complicated and unpredictable settings in which the truth or reality of things is subject to a variety of interpretations.[543] Generative AI, conversely, is more likely to generate simple, fluently presented, and seemingly authoritative answers, creating a 'smoothing' effect that obscures questions, disintermediates uncertainty, and presents it as an authoritative, stable factual response, which may be contentious, unclear, or politically sensitive (Di Martino and Ford, 2025, p. That restriction renders a vision of an AI as a replacement for diplomacy not only perilous, but simply illogical.[544]

Thus, three principles must serve as the foundation of the idea of AI-enhanced diplomacy. Accordingly, the idea of AI-supported diplomacy should be based on three principles: The first is the principle of complementarity: AI should support, not replace, the work of humans. The second is the principle of meaningful human control: Outputs generated using AI systems shall be verified,

---

[540] Mark Say, *Foreign Secretary Points to Future of AI in Diplomacy*, UKAuthority, 14 luglio 2025, https://www.ukauthority.com/articles/foreign-secretary-points-to-future-of-ai-in-diplomacy

[541] Ivi.

[542] Luigi Di Martino e Heather Ford, "Navigating Uncertainty: Public Diplomacy vs. AI," *Place Branding and Public Diplomacy* 21 (2025): 350-354, pp. 350-353, https://doi.org/10.1057/s41254-024-00330-z

[543] Ivi, pp. 351-352.

[544] Ivi.

contextualized, and interpreted by responsible officials. The third is to state institutional accountability, which takes AI utilization in diplomacy undeniably hand in hand and attached to data strategies, the regulatory guidelines, security criteria, data protection, and staff training. S. Stanzel and Voelsen contribute by proposing that diplomatic administrations should adopt a foreign policy data strategy and the adoption of regulatory guidelines on the use of AI, in this case, limitations on system security, data protection, human control, and accountability.[545]

Once the term AI-assisted diplomacy is defined as an assistant and enhancement of human diplomatic work, some real-life applications of artificial intelligence in today's diplomacy can be explored. These cases show how AI has already made an inroad into foreign ministry work in the form of an assistive function: translation, document sorting and classification, data interpretation, communication management, dealing with citizens, and heavy parsing and consultation of complex documents. The examples are not because AI has replaced diplomats, but because it helps change the working processes of diplomats within diplomatic institutions with respect to information, service, and processes of public communication.

Firstly, the use of AI internally by foreign ministries. The British case is interesting in this respect because it has become a gradual process of introducing AI, beginning with handling administrative and analytical tasks. The Foreign, Commonwealth and Development Office (FCDO) parliamentary response says they have used artificial intelligence (AI) for language translation, document classification and data analysis to support political and consular teams (Westminster Brief, 2025), and are coordinating the responsible use of AI technology through FCDO.ai, the UK government's wide-ranging trial of the Microsoft Copilot system, and assurance processes that map to UK AI safety principles.[546] The case is significant as it differentiates between digitization by administration and the actual “augmented diplomacy”. Diplomatic functions in the sense of political negotiation are not directly involved in machine translation and document classification but affect the diplomatic functions to operate faster and more efficiently. In a Foreign Office where vast amounts of cables, documents, reports, analyses, and communications circulate in numerous languages, AI can help to streamline workflows, aid access to information, and enhance cross-office cooperation. Because of their specific position in public administration, however, it is

[545] Stanzel e Voelsen, *Diplomacy and Artificial Intelligence*, pp. 6-8.

[546] Westminster Brief, *Foreign, Commonwealth and Development Office: Artificial Intelligence — Written Question 73676*, risposta del 16 settembre 2025, disponibile all'indirizzo: https://westminsterbrief.co.uk/archive/pq/73676

necessary that acceptable use policies, human control and supervision and institutional guarantees are in place in these applications. The FCDO case illustrates how augmented diplomacy relies not just on using the technological tool, but on developing internal governance for the use of AI as well, as confirmed by official British government documentation.[547]

The more strategic dimension of the British case emerges from a speech by Foreign Secretary David Lammy, reported by UKAuthority in July 2025. According to Lammy, AI should be brought to the heart of the FCDO's policymaking machinery to accelerate activities such as drafting briefings and memoranda, modeling the effects of diplomatic crises, red-teaming British responses, and identifying emerging risks that human analysts might not promptly detect.[548]

This vision comes closest to the idea of AI-augmented diplomacy because it concerns not only document management but also support for policymaking.

The most significant element, however, is that the speech itself maintains an explicit distinction between machine analytical capacity and human diplomatic judgment. Indeed, Lammy argues that AI cannot solve foreign policy, eliminate risk, or remove the need for careful human judgment and the building of trust.[549] In this sense, the British case confirms the thesis developed in the previous section: AI can help navigate an international environment characterized by growing uncertainty, but it cannot replace the diplomat's political responsibility. It can detect patterns, simulate scenarios, or accelerate document production, but the decision about the political significance of such outputs must remain human.

A second area of application concerns consular services.[550] Here, the use of AI appears more immediately practical and less controversial than political decision-making or diplomatic negotiation.[551] Consular services are often characterized by repetitive requests, standardized procedures, and the need to provide rapid information to citizens abroad.[552] Precisely for this reason, they constitute a particularly suitable terrain for forms of assistive AI, such as chatbots and

---

[547] Ivi.

[548] Mark Say, *Foreign Secretary Points to Future of AI in Diplomacy*, UKAuthority, 14 luglio 2025, https://www.ukauthority.com/articles/foreign-secretary-points-to-future-of-ai-in-diplomacy

[549] Ivi.

[550] Katharina E. Höne et al., *Mapping the Challenges and Opportunities of Artificial Intelligence for the Conduct of Diplomacy* (Geneva: DiploFoundation, 2019), 14–15.

[551] Volker Stanzel and Daniel Voelsen, *Diplomacy and Artificial Intelligence: Reflections on Practical Assistance for Diplomatic Negotiations*, SWP Research Paper 1 (Berlin: Stiftung Wissenschaft und Politik, 2022), **31**.

[552] OECD, *Governing with Artificial Intelligence: The State of Play and Way Forward in Core Government Functions* (Paris: OECD Publishing, 2025), chapter "AI in public service design and delivery."

virtual assistants.[553] The case of the Greek Ministry of Foreign Affairs clearly illustrates this trend. The Ministry's website describes the Virtual Assistant of the Consular Authority as an artificial intelligence tool capable of answering citizens' written questions, submitted online through the consular authorities' websites, Facebook, Messenger, or Viber, guiding the procedures to follow for each consular service.[554]

This is especially important as the Greek virtual assistant is embedded into the website of consular offices, where the individual services are offered in Greek, English, and a dozen other languages.[555] Both citizens overseas and other parties may require rapid access to passports, certificates, registration, emergency, administrative documents, or consular information. The multilingual chatbot can relieve pressure on consular offices, provide instant answers to common questions, and provide information even if the office is closed. In this regard, artificial intelligence is not intended to replace the consul or official, but rather to provide an initial level of guidance towards the official; this is clear from Greek ministerial documents.[556] This is also a theoretical lesson derived from the Greek case: Purposeful incorporation of AI tends to occur when the task is procedural, repetitive, and defined.[557] Questions about a consular procedure are not questions of a diplomatic crisis and are not amenable to international negotiation.[558] The advantage of AI is that it's available around the clock (which hasn't been the case with humans), it can operate in multiple languages and is much quicker.[559] In the latter, however, there is a significant risk of oversimplification, a danger not encountered by humans.[560] Thus, the difference between low-discretionary and high-discretionary is imperative to the design of responsive augmented diplomacy.[561] When the procedure is well understood and predictable, AI can be used more

[553] Hellenic Republic, Ministry of Foreign Affairs, "Chatbot — Virtual Assistant of the Consular Authority."
[554] Hellenic Republic Ministry of Foreign Affairs, *Chatbot — Virtual Assistant of the Consular Authority*, https://www.mfa.gr/en/services/digital-tools/chatbot/
[555] Ivi.
[556] Ivi.
[557] OECD, *Governing with Artificial Intelligence: The State of Play and Way Forward in Core Government Functions* (Paris: OECD Publishing, 2025), section "AI in public service design and delivery."
[558] Volker Stanzel and Daniel Voelsen, *Diplomacy and Artificial Intelligence: Reflections on Practical Assistance for Diplomatic Negotiations*, SWP Research Paper 1 (Berlin: Stiftung Wissenschaft und Politik, 2022), 31
[559] Hellenic Republic, Ministry of Foreign Affairs, "Chatbot — Virtual Assistant of the Consular Authority"; Information Society S.A., "Transformation of Consular Services Procedures."
[560] Luigi Di Martino and Heather Ford, "Navigating Uncertainty: Public Diplomacy vs. AI," *Place Branding and Public Diplomacy* 21, no. 3 (2025): 350–352.
[561] Elham Tabassi, *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*, NIST AI 100-1 (Gaithersburg, MD: National Institute of Standards and Technology, 2023), 13–16, 22, 40–41.

confidently; when it involves political judgment, cultural sensitivity and context will require interpretation, the computer must be kept firmly under human control.[562]

The third one, being particularly representative, is Vienna in Ukraine, where in May 2024, a "digital persona" created by artificial intelligence (AI) technologies was introduced as a means of conveying consular information to the media. According to the Ukrainian Ministry of Foreign Affairs, Victoria Shi was created to officially comment on consular information, provide news on the work of consuls to protect the rights and interests of Ukrainian citizens abroad, respond to incidents or emergencies, and transmit timely and verified information from the Ministry's Consular Department.[563] This case is significant because it shifts AI from administrative assistance to institutional public communication.

Victoria Shi is not a chatbot that uses simple text communication, but a visual communication interface. Its job is not just to respond to questions, but also to act as the Ministry's voice in a very special, at times stressful, situation, which is war: to inform Ukrainian citizens of the country abroad. It was in this context that then-Minister Dmytro Kuleba introduced the project as a means of saving time and resources, empowering real diplomats to devote their efforts to other project activities serving citizens.[564] The theoretical interest of this case is that it illustrates a new approach to public diplomacy that merges AI and has significant institutional control. The Guardian reports that Victoria Shi reads official statements from the Ministry, but that the texts are not generated autonomously by AI: they are written and verified by real people.[565] The same article clarifies that AI intervenes primarily in the visual part of the communication, while the content remains under human control.[566] The model that emerges is therefore that of an AI-generated interface, but with human-controlled content.[567] For AI-assisted diplomacy, this division is crucial; the aim is to make

[562] Stanzel and Voelsen, *Diplomacy and Artificial Intelligence*, 31; Katharina E. Höne et al., *Mapping the Challenges and Opportunities of Artificial Intelligence for the Conduct of Diplomacy* (Geneva: DiploFoundation, 2019), 25–26, 29.

[563] Ministry of Foreign Affairs of Ukraine, *The Ministry of Foreign Affairs of Ukraine Has Appointed a Digital Person to Inform on Consular Issues*, 1 maggio 2024, https://mfa.gov.ua/en/news/mzs-ukrayini-priznachilo-cifrovu-osobu-dlya-informuvannya-shchodo-konsulskih-pitan

[564] Ivi.

[565] The Guardian, *Ukraine Unveils AI-Generated Foreign Ministry Spokesperson*, 3 maggio 2024,: https://www.theguardian.com/technology/article/2024/may/03/ukraine-ai-foreign-ministry-spokesperson

[566] Ivi.

[567] Ministry of Foreign Affairs of Ukraine, "The Ministry of Foreign Affairs of Ukraine Has Appointed a Digital Person to Inform on Consular Issues," May 1, 2024; Agence France-Presse, "Ukraine Unveils AI-Generated Foreign Ministry Spokesperson," *The Guardian*, May 3, 2024.

things simpler or more transparent for politicians and civil servants, but nothing can outsmart human responsibility and accountability for political and institutional voice verification.[568]

The Ukrainian Ministry has also adopted measures against falsification in digital form activities. This is especially crucial as it shows that video productions already need to be authenticated when adopted in diplomatic communication.[569] In an information world that is full of deepfakes, false information, and audiovisual manipulation, avatars or digital personas-based diplomacy needs to develop tools that ensure authenticity.[570] QR codes do not make things safe to false; but they add the element of institutional control to differentiate official messages from manipulated ones.[571] Victoria Shi's case thus illustrates both possibilities and uncertainties of 'augmented diplomacy'.[572] On the other hand, AI can enable a foreign ministry to save resources, communicate quickly, and try out new public face structures for the foreign ministry.[573]

Conversely, there are concerns around trusting, authenticity, and accountability while using an artificial face for delivering institutional messages.[574] In diplomacy, credibility depends not only on the content of the message but also on the person delivering it, the context in which it is received, and the perceived trustworthiness of the institution.[575] AI-generated communication, therefore, can only be effective when transparent and verified, and with clear attribution of responsibility for the content, as we have seen in the official communication from the Ukrainian Ministry, as well as in independent journalistic verifications.[576]

---

[568] Volker Stanzel and Daniel Voelsen, *Diplomacy and Artificial Intelligence: Reflections on Practical Assistance for Diplomatic Negotiations*, SWP Research Paper 1 (Berlin: Stiftung Wissenschaft und Politik, 2022), 31; Elham Tabassi, *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*, NIST AI 100-1 (Gaithersburg, MD: National Institute of Standards and Technology, 2023), 22, 40–41.

[569] Ministry of Foreign Affairs of Ukraine, *The Ministry of Foreign Affairs of Ukraine Has Appointed a Digital Person to Inform on Consular Issues*

[570] Luigi Di Martino and Heather Ford, "Navigating Uncertainty: Public Diplomacy vs. AI," *Place Branding and Public Diplomacy* 21, no. 3 (2025): 350–352; Ammarah Hashmi et al., "Unmasking Illusions: Understanding Human Perception of Audiovisual Deepfakes," 2024, 1–2.

[571] Ministry of Foreign Affairs of Ukraine, "The Ministry of Foreign Affairs of Ukraine Has Appointed a Digital Person to Inform on Consular Issues," May 1, 2024.

[572] Ministry of Foreign Affairs of Ukraine, "The Ministry of Foreign Affairs of Ukraine Has Appointed a Digital Person to Inform on Consular Issues," May 1, 2024; Di Martino and Ford, "Navigating Uncertainty," 350–354.

[573] Ministry of Foreign Affairs of Ukraine, "The Ministry of Foreign Affairs of Ukraine Has Appointed a Digital Person to Inform on Consular Issues," May 1, 2024; Agence France-Presse, "Ukraine Unveils AI-Generated Foreign Ministry Spokesperson," *The Guardian*, May 3, 2024.

[574] Luigi Di Martino and Heather Ford, "Navigating Uncertainty: Public Diplomacy vs. AI," *Place Branding and Public Diplomacy* 21, no. 3 (2025): 350–352

[575] Di Martino and Ford, "Navigating Uncertainty," 352.

[576] Ministry of Foreign Affairs of Ukraine, "The Ministry of Foreign Affairs of Ukraine Has Appointed a Digital Person to Inform on Consular Issues," May 1, 2024; Agence France-Presse, "Ukraine Unveils AI-Generated Foreign Ministry Spokesperson," *The Guardian*, May 3, 2024.

A fourth case, that is not directly linked to a foreign ministry, but to an international policy setting one, is discussing LLMs in negotiation processes, and document parsing with complex documents. Ziegler, Lothian, O'Neill, Anderson, and Ota describe the case of a BBNJ Question-Answering Bot that is built to ask questions of the Agreement on Marine Biodiversity in Areas Beyond National Jurisdiction (BBNJ Agreement).[577] The case is especially interesting for diplomacy because the BBNJ Agreement is the outcome of long multilateral negotiations in the United Nations and is for a sector where the developed countries have unequal relationships with the developing countries.

The authors believe that LLM-based tools will enable LVM to allow policymakers to undertake tasks that are mundane and labor-intensive, including interpretation and analysis of complex legal documents; production of formal statements; rapid information retrieval to answer specific questions; and support in public consultation processes.[578] In this way, BBNJ can be considered an example of AI as a capacity-building tool, clearly highlighting the capacity to build the BBNJ chatbot.[579] Limited resources delegations may be able to leverage those same tools to more efficiently move through complicated technical negotiations with reduced personnel, less legal expertise, and/or scientific knowledge. This directly addresses the Global Majority problem outlined in section 4.3: AI could curtail several operational asymmetries by enhancing access to information and documents. But the same research indicates that technologies can also perpetuate or deepen inequities. The authors highlight that AI use cases can reflect the views of wealthier and more powerful States as datasets for training and official documents, and discourses available in these States, are often dominated by Northern perspectives.[580] Moreover, many developing countries do not maintain an explicit record of their own position and its arguments but move about in the shadows behind formal diplomatic expressions, which consist of both cautious formulations and euphemisms as well as indirect allusions. With the help of official documents, the chatbot can then emulate the celebratory or consensual tone found in negotiation discourses but cannot understand the actual political tensions that developed over the course of these negotiations.[581]

---

[577] Matt Ziegler, Sarah Lothian, Brian O'Neill, Richard Anderson e Yoshitaka Ota, "AI Language Models Could Both Help and Harm Equity in Marine Policymaking," *npj Ocean Sustainability* 4, no. 32 (2025): pp. 1-2, https://doi.org/10.1038/s44183-025-00132-7

[578] Ivi, pp. 1-2.

[579] Ivi.

[580] Ivi, pp. 2-4

[581] Ivi, pp. 4-5.

The case is pivotal as it demonstrates the non-neutral nature of the use of AI for diplomacy. Despite its role as a service to garner information, it relies on the information sources available, the conceptualizations of phenomena in those sources, and the questions to be answered. Ziegler, et al. point out that biases may come from the subtleties of the linguistic models, the type of document database being used by the application, and the prompt or interface design.[582] This also makes it possible for an AI tool to litigate diplomacy instead of just helping it, and it also gives it a particular perspective on a negotiation, biasing what is heard and what isn't.[583]

The same study suggests a risk of overreliance. Given the anthropomorphic nature of chatbots, their ability to respond seamlessly and suggest seemingly convincing explanations, there is a risk that suggested policies and diplomatic assertions made by chatbots could be accepted by politicians or policymakers as relying exclusively on them, without proper scrutiny.[584] The danger of this kind of risk is greater in a diplomatic setting, where a lot of pertinent information isn't in an official document but rather comes out during informal negotiations, in the corridors, as part of personal relationships of trust, context, or institutional memory. A chatbot may summarize a text, but it doesn't completely know what it's not saying. For just this reason, the BBNJ case highlights the difference between having access to the information and diplomatic skills.[585] While AI can save underfunded delegations valuable time reviewing complex documents that will take time to read, its ability to perform these other political tasks, such as coal-building, reading a negotiator's position, making negotiation strategies, and supporting the country's position, cannot be matched. The same authors point out, however, that simply adopting AI tools would not be a substitute for genuine capacity building, as the developed world and installed relief and development organizations might use the cheap technological approach as a shortcut for more solid cooperation, technical assistance, and knowledge transfer.[586] However, if the aim is to diminish some operational asymmetries, AI must be embedded in a larger framework of technical empowerment, trainers and training, and inclusion.

The cases examined illustrate the concrete ways in which AI can assist diplomatic missions thus far: improving document management, assisting in access to consular services, assisting public

[582] Ivi, pp. 2-6.
[583] Ivi.
[584] Ivi, pp. 5-6.
[585] Ivi.
[586] Ivi, pp. 7-9.

communication, and aiding delegations with limited resources in navigating complex multilateral processes. But, given that AI is advancing in fields involving political communication, strategic analysis, and negotiating meeting preparation, and that diplomats' work is expanding to these domains, it is impossible to assess augmented diplomacy purely based on technical efficiency. It should also be evaluated in terms of accountability, trust, human control, information quality, and democratic legitimacy.[587] The first risk is associated with uncertainty management. Diplomacy relies on factors more than just rational and predictable; however, factors such as incomplete information, ambiguous signals, divergent interests, and dramatic changes. The prospect of generating these types of responses, which are fluid, unique, and can appear authoritative, poses a significant epistemic and aleatory uncertainty challenge for public diplomacy, notes Di Martino and Ford, as can an incorrect summary, an overly confident prediction or an ambiguous statement in diplomacy, where a request is underway to believe or trust in the answers.[588] Thus, AI-generated content does not constitute conclusions and serves as a guide rather than a definitive fact or message that needs to be fact-checked, contextualized, and interpreted with human wisdom.

The second set of threats is related to public communication, the dependence on AI systems, and negotiation. In foreign policy, covering much of which consists of informal interactions, exchanges, and communications with different players, trust, institutional memory, and interpretation of texts, this risk of reliance is especially critical, given the delicate nature of political exchanges, especially when negotiating complex matters.[589] In diplomacy, one example is the BBNJ chatbot that uses synthetic text and interpretation can help build public dialogue around intricate legal texts, but also slip in dominant perspectives and obscure political tension that are never captured in the official words themselves.[590] While AI can be helpful in negotiating preparation, planning of a scenario, and document analysis, it should never be used as a replacement for diplomatic opinion. Research on negotiations between AI agents also highlights the importance of relational factors like co-operation, trust, and rapport-building, as well as new

[587] Volker Stanzel e Daniel Voelsen, *Diplomacy and Artificial Intelligence: Reflections on Practical Assistance for Diplomatic Negotiations*, SWP Research Paper 1 (Berlin: Stiftung Wissenschaft und Politik, 2022), 6–8, 31.
[588] Luigi Di Martino e Heather Ford, "Navigating Uncertainty: Public Diplomacy vs. AI," *Place Branding and Public Diplomacy* 21 (2025): 350–354,
[589] Rhys Crilley e Robert A. Saunders, "The Resonant Slop Machine: Public Diplomacy and Strategic Narratives in the Age of Artificial Intelligence," *Place Branding and Public Diplomacy* (2025): 1–2, 8.
[590] Matt Ziegler, Sarah Lothian, Brian O'Neill, Richard Anderson e Yoshitaka Ota, "AI Language Models Could Both Help and Harm Equity in Marine Policymaking," *npj Ocean Sustainability* 4, no. 32 (2025): 2–6

ethical and procedural arguments regarding the use of AI-specific tactics such as prompt injection or reasoning obfuscation.[591]

To ensure the success of augmented diplomacy in the future, it is important to have AI literacy, internal governance, and the presence of humans. While not every diplomat needs to be a computer scientist, it is important to recognize that all diplomats need to have an understanding of the boundaries and limitations of the AI systems they interact with, be able to identify their biases, check their sources, analyze their results, and differentiate between information created by AI, verified knowledge, and political judgment. Diplomatic administrations should also craft foreign policy data strategies, and establish principles for improper use of the system, such as data security, data protection, human oversight and institutional accountability, should not be "AI instead of diplomacy," but should be "AI in support of diplomacy".[592] While AI can be a useful tool for advancing diplomacy, its role cannot replace the need for diplomacy to be prudent, legitimate, or just.[593] AI can help accelerate diplomacy, provide more information, and make it more accessible, but it can't make diplomacy more prudent, legitimate, or just.[594] These qualities remain subject to human judgment, institutional accountability, and the recognition and creation of trust among political actors.[595]

4.5 Sustainability Challenges in AI Governance

The study highlighted in the earlier sections has revealed the legal, geopolitical, diplomatic, and multilateral challenges surrounding AI's governance. But AI governance isn't complete without considering the substance of AI. AI can be called a digital, non-physical, and data-based technology. In fact, it relies on highly complex physical systems, data centers, servers, semiconductors, GPUs, power grids, cooling systems, as well as hardware recycling, data centers, etc., and water-cycle, etc., and on global chains of production and disposal. Sustainability of AI

[591] Michelle Vaccaro, Michael Caoson, Harang Ju, Sinan Aral e Jared R. Curhan, "Advancing AI Negotiations: New Theory and Evidence from a Large-Scale Autonomous Negotiations Competition," 2025, 1–12, 15–16.
[592] Stanzel e Voelsen, *Diplomacy and Artificial Intelligence*, 6–8, 31; Di Martino e Ford, "Navigating Uncertainty," 353.
[593] Volker Stanzel and Daniel Voelsen, *Diplomacy and Artificial Intelligence: Reflections on Practical Assistance for Diplomatic Negotiations*, SWP Research Paper 1 (Berlin: Stiftung Wissenschaft und Politik, 2022), 31
[594] Katharina E. Höne et al., *Mapping the Challenges and Opportunities of Artificial Intelligence for the Conduct of Diplomacy* (Geneva: DiploFoundation, 2019), 25–29.
[595] Luigi Di Martino and Heather Ford, "Navigating Uncertainty: Public Diplomacy vs. AI," *Place Branding and Public Diplomacy* 21 (2025): 350–354, especially 352–353; Stanzel and Voelsen, *Diplomacy and Artificial Intelligence*, 31.

cannot, therefore, be boiled down to just efficiency of algorithms, or simply decreasing algorithmic bias, but needs to be linked to the use of natural resources and to eco-economic costs being spread throughout the whole lifespan of the technology.

This is particularly relevant as it overturns one of the most pervasive myths in digital transformation: that digital somehow reduces the need for material or polluting or even intrusive traditional industrial infrastructure. UNCTAD has indicated that digitalization is largely called a virtual and intangible object; this representation may, however, hide its tangible environmental effect.[596] The environmental costs of digitalization encompass the extraction and utilization of raw materials, energy and water, greenhouse gas emissions, pollution, and the creation of electronic waste, semiconductors, computers, and other electronic goods, which is especially pertinent when it comes to artificial intelligence, as most AI systems, in particular large-scale generative models, depend increasingly on the consumption of increasingly sophisticated computing power and infrastructures.[597] Thus, artificial intelligence is inseparable from energy, becomes the problem. The International Energy Agency (IEA) clearly states that there is no AI without energy, and specifically without electricity; meanwhile, managing energy networks more efficiently, performing optimization or innovation, and developing AI in these sectors is about changing energy, at least as much as it is about changing the world.[598] The duality creates special challenges for environmental governance of AI; on the other hand, AI also consumes energy and creates new demand, particularly with the use of data centers and computing systems.[599] It can, on the other hand, make a significant contribution to the ecological transition by helping to enhance forecasting, efficiency, electricity grid management, predictive maintenance, and innovations for clean technologies.[600] The governance dilemma is exactly how to strike a balance between these two factors, to control the environmental impacts of artificial intelligence, while also not neglecting its climate potential.[601]

---

[596] UNCTAD, *Digital Economy Report 2024: Shaping an Environmentally Sustainable and Inclusive Digital Future — Overview* (Geneva: United Nations, 2024), pp. vi-vii, https://unctad.org/publication/digital-economy-report-2024

[597] Ivi.

[598] International Energy Agency, *Energy and AI* (Paris: IEA, 2025), pp. 3-4, disponibile all'indirizzo: https://www.iea.org/reports/energy-and-ai

[599] Ivi.

[600] David Rolnick et al., "Tackling Climate Change with Machine Learning," arXiv:1906.05433, 2019, 1–4, 5–11.

[601] OECD, *Measuring the Environmental Impacts of Artificial Intelligence Compute and Applications: The AI Footprint*, OECD Digital Economy Papers, no. 341 (Paris: OECD Publishing, 2022), **5–6**; UNFCCC, "#AI4ClimateAction" https://unfccc.int/ttclear/templates/ttclear/pages/ttclear/pages/artificial_intelligence

Before diving into this balance, let's go back to data centers. According to the IEA, in 2024, data centers used some 415 TWH of electricity, or roughly 1.5% of global electricity demand. The electricity grid is globally one of the largest consumers of electricity, and considering the local impacts of this consumption would be important, since IEA data shows that, by area, data centers are concentrated and can be significant percentage shares of electricity consumption in certain local markets.[602] This is important to highlight because the sustainability of AI needs to be measured not only in global percentages but also at the local level in territorial pressures on the electricity grid, water, and land, and local communities. Not only is the issue more salient, but also more relevant as the demand is projected to increase on the high scale. In the IEA's 'base case' scenario, energy use in data centers could grow to about 945 TWh by 2030, which is roughly the same as Japan's total electricity use today.[603] The analysis by the IEA also reveals that while data centers will remain a small fraction of global electricity consumption, they will expand at a much faster rate than other sectors and could play an important role in emerging economies where electricity demand has been flat or is rising slowly.[604] By this separation of the consequences of activity on a global and local scale, two equally opposing errors are avoided.[605] One way to do this would be to minimize the issue, pointing to the fact that data centers only consume a small share of the electricity used around the world.[606] The second would be to overestimate it, referring to AI as the main and sole contributor to the climate crisis.[607] A more demanding reading reveals, however, that AI is driving up electricity demand alongside other pressures on the energy system, namely, electrification of transport, industrial growth, air conditioning, urbanization, and the growth of 'emerging economies'.[608] The extent of the energy usage is certainly a question of quantity but of quality as well, perhaps how, when, where, and with what kind of energy and with what effects on local networks?[609]

The lessons learned from Ireland are particularly informative in this regard. Irish Central Statistics Office data shows metered electricity usage in data centers rising by 10% from 6,335 GWh in 2023

[602] IEA, *Energy and AI*, p. 14.
[603] Ivi.
[604] Ivi, pp. 14-15.
[605] International Energy Agency, *Energy and AI* (Paris: IEA, 2025), 50, 65. https://iea.blob.core.windows.net/assets/de9dea13-b07d-42c5-a398-d1b3ae17d866/EnergyandAI.pdf
[606] Ivi.
[607] Ivi.
[608] International Energy Agency, *Energy and AI*, 51, 65.
[609] International Energy Agency, *Energy and AI*, 50, 65, 216

to 6,969 GWh in 2024.[610] The case illustrates how there is not an equal distribution of environmental and infrastructural burden by data centers, with the share of consumption from data centers increasing from 5% in 2015 to 22% in 2024.[611] While it may seem small-scale at the global level, in certain countries or regions, it can become a key driver of energy planning and grid management, or a heavy part of the political debate on the sustainability of digital development.[612] Demonstrating that AI governance needs to be positioned alongside energy governance in Ireland.[613] While trying to control the transparency, security, or social risks of AI models, we must also consider whether the infrastructure for their development and operation is compatible with the goals of energy and climate.[614] As data centers consume a larger portion of national electricity usage, decisions on where best to locate these applications, their connection to the grid, energy efficiency, and where to get energy from renewables, become public policy issues. [615] Put another way, AI scale is not just a private technical choice, as it has systemic impacts on power systems and States' climate policies. [616] A second aspect of AI sustainability is its impact on networks. This statement moves the discussion to several infrastructural aspects, as electrification strains electricity networks already in many regions, and without addressing the risks, approximately 20% of planned data center projects are expected to be delayed in 2030 because of congestion, connection queues, and difficulties in supply chains and purchasing critical technologies like transformers or turbines.[617] The electric grid will need to be resilient to this demand from AI, both in its energy and emissions consumption and its capacity to handle this load. The need for grid resiliency will be especially important for land-use planning, energy security, and coordinating between public and private investments in the state.[618]

---

[610] Central Statistics Office Ireland, *Data Centres Metered Electricity Consumption 2024 — Key Findings*, 10 giugno 2025, https://www.cso.ie/en/releasesandpublications/ep/p-dcmec/datacentresmeteredelectricityconsumption2024/keyfindings/

[611] Ivi.

[612] International Energy Agency, *Energy and AI* (Paris: IEA, 2025), 50, 65

[613] Central Statistics Office Ireland, "Data Centres Metered Electricity Consumption 2024: Key Findings," June 10, 2025.

[614] European Commission, "Energy Performance of Data Centres." https://energy.ec.europa.eu/topics/energy-efficiency/energy-efficiency-targets-directive-and-rules/energy-efficiency-directive/energy-performance-data-centres_en?utm

[615] IEA, *Energy and AI*, p. 15.

[616] European Commission, "Commission Adopts EU-wide Scheme for Rating Sustainability of Data Centres," March 15, 2024 https://energy.ec.europa.eu/news/commission-adopts-eu-wide-scheme-rating-sustainability-data-centres-2024-03-15_en?prefLang=lv&utm International Energy Agency, *Energy and AI*, pp, 50, 65, 216

[617] International Energy Agency, *Energy and AI*, p. 50

[618] Ivi, pp. 65, 216

The type of computing infrastructure is another issue. AI-driven accelerated servers' consumption will be faster than conventional servers' consumption, which will, in turn, accelerate at a higher rate than ever before, according to IEA.[619] This is important as it reveals that the data centers that use accelerated servers can be responsible for nearly half of global data center electricity consumption growth. AI is reshaping data center interiors with newer power densities and unique hardware requirements. As AI grows, data centers are becoming increasingly complex, with higher power densities and specialized hardware. Evaluate the sustainability of AI; therefore, should be done throughout the whole system lifecycle, and not only at its usage. UNEP (United Nations Environment Program) envisions an end-to-end strategy, where environmental impacts of AI can be evaluated across the software life cycle and the environmental impacts across the hardware life cycle.[620] This is an approach suitable to governance – one which does not only consider one phase of the software cycle (e.g. training) without accounting for other phases (e.g. data collection, software development, deployment, inference, maintenance, decommissioning) or one in which only a single phase of the hardware cycle (e.g. chip production) is considered without addressing the rest (e.g. data center operations, extract mining, hardware manufacture, maintenance, disposal of e-waste).[621] Especially important is the final phase of inference. The environmental impact of AI has been the subject of discussion in the past and mostly revolves around the training process of the big models, as large models involve large amounts of data and computational power. With the advent of generative models and large language models, however, more equality or even more requirements are needed for inference than for training, particularly if many or even billions of users use such a model.[622] This point is vital for governance, as a model can be trained from one shot and yet repeatedly queried on an ongoing basis. The sustainability of AI is therefore not strictly related to the energy consumed during the model's creation but also to the number of interactions between the model and end-users on a day-to-day basis.[623]

The OECD also does this in its report "Measuring the Environmental Impacts of Artificial Intelligence Computing and Applications", which highlights the importance of quantifying two aspects of the environmental impacts of AI: computing and applications. The direct impacts relate

[619] Ivi, p. 64.
[620] United Nations Environment Programme, *Artificial Intelligence (AI) End-to-End: The Environmental Impact of the Full AI Life Cycle Needs to Be Comprehensively Assessed*, Issues Note, 21 settembre 2024, p. 2.
[621] UNEP, *Artificial Intelligence (AI) End-to-End*, pp. 2-3.
[622] Ivi.
[623] Ivi.

to the lifecycle of computational resources, mostly water use, energy, greenhouse gas, and raw materials requirements, including their production, transportation, operational use, and end-of-life processes. On the other hand, indirect impacts are developments that deal with the application of AI in different domains: positive impacts, such as smart grid or digital twin; negative impacts, such as fostering unsustainable consumption patterns.[624] The separation of direct and indirect effects is important as it ensures that the relationship between AI and the environment isn't oversimplified. AI is about more than using it to consume energy; it can also be used for environmental optimization. But few can say that AI can play a role in the ecological transition, and that is as far as sustainability goes. Requires measures to assess the environmental impacts of the applications of AI technologies over the costs of the infrastructure that must be created and mobilized to develop and implement the applications. An AI governance framework should accordingly build on a rigorous logic of evaluation that considers both the costs of AI and the impacts of using AI to lower other environmental costs.

A third aspect concerns water; its use in data centers can be direct, indirect, or related to cooling systems. While current scientific evidence regarding the water needs of AI lacks a cohesive body of work, UNEP reports that some sources estimated that the water demands for AI worldwide could amount to 4.2–6.6 billion cubic meters by 2027.[625] The need to use large quantities of ultrapure water in semiconductor manufacturing, the competition between industrial and human uses for water resources in hot and water-stressed areas, is an environmental justice issue, especially given that the cost of AI can be felt at the local level as it strains already limited resources, particularly water.[626]

The topic of water also concerns the reader, who is not concerned about the sustainability of AI, which does not reduce its carbon footprint.[627] A data center run with renewable power can still have a high water usage environmental impact when situated in a dry climate region.[628] Likewise, a more efficient design might involve the use of more water-based, mineral-based, and chemical-

[624] OECD, *Measuring the Environmental Impacts of Artificial Intelligence Compute and Applications: The AI Footprint*, OECD Digital Economy Papers, no. 341 (Paris: OECD Publishing, 2022), pp. 5-6.
[625] UNEP, *Artificial Intelligence (AI) End-to-End*, p. 3.
[626] Ivi.
[627] OECD, *Measuring the Environmental Impacts of Artificial Intelligence Compute and Applications: The AI Footprint*, OECD Digital Economy Papers, no. 341 (Paris: OECD Publishing, 2022), 5–6
[628] United Nations Environment Programme, *Artificial Intelligence (AI) End-to-End*, 3.

based production processes in making the chips.[629] Sustainable governance can't be measured by a single metric; it can be measured by several indicators, which include energy, emissions, water, materials, waste, supply chain, and territorial impact.[630]

There is also a fourth: critical minerals, semiconductors, and e-waste.[631] A fourth aspect concerns critical minerals, semiconductors, and e-waste. UNEP reminds the public that minerals and metals are needed for data centers, GPUs, and digital infrastructure industries, and that the increased technology and tools around AI drive the demand for chips. [632] The same note emphasizes how the mix of materials needed for digitalization is, in many cases, roughly equivalent to those needed for the energy transition. This competition between different sustainability missions is of fundamental importance to the chapter because the challenges required for digitalization largely coincide with those required for the energy transition. Sustainability of AI, not just the use of models, but the industrial ecosystem that enables it. UNCTAD has noted, however, that demand for digital and low-carbon technologies will have a significant impact, with the production of minerals such as graphite, lithium, and cobalt expected to grow significantly.[633] The issue of the environment of AI, therefore, overlaps with the geopolitics of raw materials and with existing disparities at the global level that already emerged from point 4.3.

E-waste is also a problem. As AI becomes more common, so will the amount of e-waste, or what UNEP calls “e-waste rations”, which is one of the fastest-growing waste streams in the world today, though it is uncertain how much of e-waste can be attributed to data centers or to AI chips. UNCTAD, for its part, documents that the volume of waste thrown was up 30% from 2010 to 2022, to the tune of 10.5 million tons, and that formal (hybrid) waste collection is much lower in developing countries than in developed ones.[634] The data show that the principles of the circular economy cannot be separated from the concept of AI sustainability.[635] The issue is not just about operation and shows a trend toward replacing old systems and disposing of them, as well as

[629] Ivi, p. 2-3.
[630] United Nations Environment Programme, *Artificial Intelligence (AI) End-to-End*, 2–4; OECD, *Measuring the Environmental Impacts of Artificial Intelligence Compute and Applications*, 5–6
[631] United Nations Environment Programme, *Artificial Intelligence (AI) End-to-End*, 2–3; UNCTAD, *Digital Economy Report 2024*.
[632] UNEP, 2024, p. 3
[633] UNCTAD, 2024, vi-vii
[634] UNEP, *Artificial Intelligence (AI) End-to-End*, p. 3.
[635] UNCTAD, *Digital Economy Report 2024 — Overview*, pp. vi-vii.

frequent hardware changes, and more powerful accelerators and data centers will be required when leveraging AI.[636]

Thus, environmental governance for artificial intelligence should take into consideration the requirements for sustainable design, recycling, reuse, material traceability, and extended producer responsibility. If not, the likelihood is that innovation can lead to technological obsolescence, and environmental costs can be pushed to the "undead" parts of the value chain. Another difficulty will be measurement. UNEP also notes that existing research on the environmental cost of AI is fragmented and that there's a shortage of granular information and data to fairly evaluate the environmental footprint of AI systems. It particularly recommends that the direct impacts are prioritized first on the most measurable and alarming areas: the use of minerals, energy, and water, the emission of greenhouse gases, pollution, and e-waste.[637] Without measures, comparable data, and specific indicators relating to AI, OECD also points to the lack of public policies aimed at sustainability.[638]

Without a measurement, there is an accountability dilemma. Lacking common metrics, it is hard to compare models, to judge corporate practices, to identify the responsibility, and to help consumers, governments, and international organizations choose a model.[639] While companies can state their sustainability commitments or make renewable energy investments, it is hard to tell what the real-world impact of such AI systems is without a common standard for reporting.[640] Likewise, the integration of energy and environmental aspects is not comprehensive, and the implementation of national AI strategies can be recommended by the states without the energy and environmental aspects being fully considered.[641]

Additionally, the process of sustainable AI governance calls for transparency, shared metrics, and verification mechanisms, as well as the institutional capacity to verify.[642]

---

[636] Ivi.

[637] UNEP, *Artificial Intelligence (AI) End-to-End*, pp. 3-4.

[638] OECD, *Measuring the Environmental Impacts of Artificial Intelligence Compute and Applications*, pp. 5-6.

[639] OECD, *Measuring the Environmental Impacts of Artificial Intelligence Compute and Applications: The AI Footprint*, OECD Digital Economy Papers, no. 341 (Paris: OECD Publishing, 2022), 18–21.

[640] Nicolas Alder, Kai Ebert, Ralf Herbrich, and Philipp Hacker, "AI, Climate, and Transparency: Operationalizing and Improving the AI Act," arXiv:2409.07471, 2024, 1–2.

[641] OECD, *Measuring the Environmental Impacts of Artificial Intelligence Compute and Applications*, 6–7.

[642] OECD, *Measuring the Environmental Impacts of Artificial Intelligence Compute and Applications*, 18–21; Alder et al., "AI, Climate, and Transparency," 2–4.

Making the environmental impact of digital infrastructures visible, measurable, and verifiable is as important as technical innovation and efficiency gains in AI sustainability as institutions spread its adoption.[643] In addition to technical innovation and greater AI model efficiency, it is essential to bring the environmental use of AI to the forefront and make it measurable and verifiable, as its institutions spread it.[644] The overarching question here isn't just how to decrease the energy usage and carbon emissions of the AI system; it's how to develop a regulatory dynamic that enables transparency with respect to energy use and emissions, data center location, indirect impacts, and the hardware life cycle.[645] The European Union (EU) is the most advanced instance of a regulatory solution that merges multiple tools: standards on energy efficiency for data centers, reporting requirements, sustainability principles, critical raw materials regulation, and early steps of transparency in the environmental sector related to the use of AI within the framework of the AI Act.[646] But that's not the whole story. It mainly focuses on the "infrastructure" level, and to a lesser degree, the "AI Model" level.[647] The outcome is an approach that recognizes the environmental impact of AI but doesn't adequately reflect the most challenging aspects: the inference process, the use of water directly linked to AI systems, emissions from AI use, and access to environmental data by the public.[648] One major component of the European approach is the Energy Efficiency Directive that regulates data centers.[649] As data centers are increasingly a key component of digital services, such as streaming, cloud, and artificial intelligence (AI) services, the European Commission has recognized that data centers represent vital and rapidly expanding infrastructure that has an undeniable environmental and climate impact, even in terms of cooler consumption and CO_2 emissions if the energy consumed by the data centers is not decarbonized.[650] For this

[643] United Nations Environment Programme, *Artificial Intelligence (AI) End-to-End: The Environmental Impact of the Full AI Lifecycle Needs to Be Comprehensively Assessed* (Nairobi: UNEP, 2024), 2–4.
[644] UNEP, *Artificial Intelligence (AI) End-to-End*, 2–4; OECD, *Measuring the Environmental Impacts of Artificial Intelligence Compute and Applications*, 5–6.
[645] UNEP, *Artificial Intelligence (AI) End-to-End*, 2–4; OECD, *Measuring the Environmental Impacts of Artificial Intelligence Compute and Applications*, 5–6.
[646] European Commission, "Energy Performance of Data Centres," accessed June 26, 2026; European Commission, "Commission Adopts EU-wide Scheme for Rating Sustainability of Data Centres," March 15, 2024; European Commission, "European Critical Raw Materials Act," accessed June 26, 2026; Alder et al., "AI, Climate, and Transparency," 1–3.
[647] Alder et al., "AI, Climate, and Transparency," 2–4.
[648] Ivi.
[649] Directive (EU) 2023/1791 of the European Parliament and of the Council of 13 September 2023 on energy efficiency, art. 12; European Commission, "Energy Performance of Data Centres," accessed June 26, 2026.
[650] European Commission, *Energy Performance of Data Centres*, https://energy.ec.europa.eu/topics/energy-efficiency/energy-efficiency-targets-directive-and-rules/energy-efficiency-directive/energy-performance-data-centres_en

reason, an obligation to monitor and report energy performance has been introduced in the Energy Efficiency Directive.[651] The Commission has created a European database to store and present relevant data about the energy consumption/energy performance, and water footprint of data centers that consume a substantial amount of energy. Delegated Regulation (EU) 2025/1364 is important because it introduces an obligation for operators to share information and key performance indicators in a more structured manner.[652] Energy and water don't remain optional but are essential to sustainability: To verify the real-life performance of digital infrastructure or to benchmark other corporate working practices, transparent measurement data is required.[653]

The Commission also in March adopted the first part of an initiative to evaluate data center sustainability in Europe. A key metric from this scheme will be whether EU operators are submitting necessary KPIs to a European database by September 15, 2024, and every subsequent 15th of the month. The purpose of the tool is to boost transparency and drive designs and efficiency gains that could be achieved to lower the energy and water footprint, encourage the use of renewable energy, improve efficiency in the grid, and encourage waste heat recovery in nearby facilities or heating networks, according to the Commission.[654]

It allows for an interpretation of an important aspect of environmental governance in Europe; it is not only about collecting the data, but also about directing the action of operators with indicators (and notes), through ratings, reputational pressures, and comparability.[655] Sustainability ratings can be an indirect policy tool, rather than banning data centers or coal plans, they will simply increase the visibility of the gap between efficient and inefficient data center infrastructure and coal plans.[656] It is, however, precisely because the system is built upon reporting and transparency, that the effectiveness of the system relies on the quality of the information, the clarity of the

[651] Directive (EU) 2023/1791, art. 12; Commission Delegated Regulation (EU) 2024/1364 of 14 March 2024 establishing the first phase of a common Union rating scheme for data centres.

[652] Ivi; Commission Delegated Regulation (EU) 2024/1364 of 14 March 2024 on the first phase of the establishment of a common Union rating scheme for data centres https://eur-lex.europa.eu/eli/reg_del/2024/1364/oj

[653] Commission Delegated Regulation (EU) 2024/1364; European Commission, "Energy Performance of Data Centres," accessed June 26, 2026.

[654] European Commission, *Commission Adopts EU-Wide Scheme for Rating Sustainability of Data Centres*, 15 marzo 2024, https://energy.ec.europa.eu/news/commission-adopts-eu-wide-scheme-rating-sustainability-data-centres-2024-03-15_en

[655] Commission Delegated Regulation (EU) 2024/1364 of 14 March 2024 establishing the first phase of a common Union rating scheme for data centres, recitals 11–15 and Annex II; European Commission, "Energy Performance of Data Centres," accessed June 26, 2026.

[656] Commission Delegated Regulation (EU) 2024/1364, recitals 11 and 13; European Commission, "Commission Adopts EU-wide Scheme for Rating Sustainability of Data Centres," March 15, 2024.

indicators, the timeliness of the information, and the public authorities, researchers, civil society, and local communities have access to the relevant information.[657] One way of looking at it is that the regulation of data centers is necessary but not sufficient.[658] It focuses on the infrastructural aspects that support a lot of AI systems, but sometimes lacks clarity on the distinction between AI compute workloads and other digital workloads. One of the most difficult hurdles in measuring the environmental footprint of AI is the difficulty in identifying AI-specific computing usage of AI systems from other forms of general purpose, scientific, and mathematical computing.[659] This is important as a data center can contain highly variable workloads such as cloud storage, streaming, enterprise applications, blockchain, AI model training and inference, and scientific workloads.[67] Environmental governance, however, ought to enable measurement to extend beyond data center governance and achieve success in connecting energy use to its types.

The other pillar is the AI Act. Although the primary focus of the analysis of the European regulation was on structural elements based on risk, fundamental rights, and the safety of AI systems, it also contains some obligations on environmental transparency. The number of floating-point operations, training time, and other training details constitute examples of the computational resources used to train the model, as described in Annex XI of the AI Act. The number of floating-point operations, training time, and any relevant training details are examples of computational resources used to train the model in accordance with Annex XI of the AI Act.[660] This is an important obligation as it lays the first step of a type of transparency at the model level. Annex XI differs from data center regulations in that it directly ties energy usage to the AI model itself. That's a significant move towards more specific governance of AI's impact on the environment.[661]

Authorities could then start to create an information bank to gauge the environmental footprint of the biggest models, if providers of general-purpose AI models must log computational resources

[657] Commission Delegated Regulation (EU) 2024/1364, recitals 11–13 and Annex II; Nicolas Alder, Kai Ebert, Ralf Herbrich, and Philipp Hacker, "AI, Climate, and Transparency: Operationalizing and Improving the AI Act," arXiv:2409.07471, 2024, 2–4.

[658] Alder et al., "AI, Climate, and Transparency," 2–4; Kai Ebert, Boris Gamazaychikov, Philipp Hacker, and Sasha Luccioni, "The Global Landscape of Environmental AI Regulation: From the Cost of Reasoning to a Right to Green AI," arXiv:2603.00068, 2026, 2–5.

[659] OECD, *Measuring the Environmental Impacts of Artificial Intelligence Compute and Applications: The AI Footprint*, OECD Digital Economy Papers, no. 341 (Paris: OECD Publishing, 2022), pp. 5-6.

[660] AI Act Service Desk, *Annex XI: Technical Documentation Referred to in Article 53(1), Point (a) — Technical Documentation for Providers of General-Purpose AI Models*, https://ai-act-service-desk.ec.europa.eu/en/ai-act/annex-11

[661] Ivi.

and energy use. But the framework is incomplete. Both the obligations imposed under the AI Act and the gaps in the Act are far from negligible, as far as the climate is concerned, observes Alder, Ebert, Herbrich, and Hacker.[662] In Article 11 and Annex IV of the guidance, there is a requirement to document the computation resources used in system development, training, testing, and validation, which does not impose a specific requirement for reporting the use of energy.[663] While Article 53 and Annex XI mandate disclosure of energy usage for general-purpose AI systems, they leave the energy use of the systems during their inference stage largely untouched.[664]

How this difference is accounted for between training and inference is pivotal. The most visible (and sometimes the most spectacular) impact of AI on the environment is its training: large models are trained on massive amounts of data. Inference can, however, be much more meaningful when viewed as cumulative when there are millions of users of a model. This criticism echoes part of what was already mentioned in Part I, whereby if AI environmental governance is to apply only to the training phase of a model, then there is a major gap, as production usage can consume even more energy than the training stage itself.[665] The authors suggest a possible corrective understanding. In Article 53, taken in combination with Annexes XI and XII, it could be argued that this allows information about requirements for a general-purpose AI model to integrate into downstream systems, which allows for a subsequent provider to make an estimate of the energy use for inferences. The interpretation of the authors, however, is still not certain, and they themselves write that it would be better to have the necessary clarification to come from the AI Office or from the Commission itself or from delegated acts in the future.[666] It has become a legal, as well as technical, issue, then, because even if the legislator mandates obligations for dissemination, it is imperative to create very clear parameters and definitions of what is to be accessed, by whom, through such methodologies, and with such levels of public visibility. Another constraint is with respect to open-source models. Although it may seem like open-source general-purpose AI models are free of some transparency obligations (unless they present a systemic risk),

---

662 [69] Nicolas Alder, Kai Ebert, Ralf Herbrich e Philipp Hacker, "AI, Climate, and Transparency: Operationalizing and Improving the AI Act," arXiv:2409.07471, 2024, pp. 1-2.

663 Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024, Artificial Intelligence Act, art. 11 and Annex IV, point 2(c).

664 Regulation (EU) 2024/1689, art. 53(1)(a) and Annex XI, Section 1, points 2(d)–(e); Nicolas Alder, Kai Ebert, Ralf Herbrich, and Philipp Hacker, "AI, Climate, and Transparency: Operationalizing and Improving the AI Act," arXiv:2409.07471, 2024, 2–4.

665 Alder et al., 2024, p. 2

666 Ivi, pp. 2-4

this does not seem to be convincing from a climate point of view: "However, making the weights, architecture, or code of a model public does not necessarily imply a lower energy footprint or environmental impact for the model.[667] Technical transparency and environmental transparency are different.[668] An open model can ultimately have a weighty energy impact despite being open and a proprietary model can unveil some information without feeling like the impact is like that of an open model.

The third constraint is relating to water consumption and indirect emissions. The AI Act includes provisions for reporting information on energy use but doesn't sufficiently cover water use or other non-critical effects of AI uses. The regulation doesn't address water usage, an important consideration for data centers, or indirect emissions from AI applications, for example in the field of oil and gas exploration.[669] What is notable is that this disparity highlights how the environmental footprint of AI goes beyond just the electricity usage of the model, as there is a risk of widespread functionality creepages. What's important to emphasize is that this gap shows how the impact of AI doesn't stop at the electricity consumed by the model alone; it's essential to address this risk of widespread functionality creep as well. Some AI technologies can have significant indirect environmental impacts, promoting faster growth in environmentally challenging activities or consumption patterns. Public access to data is the fourth limitation. Authorities have access to the data for implementing the AI Act, but the users of the data downstream from the authority, in other words, those who could implement it themselves, or the wider public, do not have access to it in large numbers.[670] This method reduces researchers, NGOs, analysts, consumers, and local communities' capacity to monitor, limits the threats to reputation and the market impact that could otherwise result from climate transparency.[671] The authors maintain that such disclosures need to be disclosed, at least in the broadest terms, as this transparency will lead to accountability, competition, and democratic control over them. This is the core issue in the debate of governance. Transparency is not only technical, but a prerequisite for political and social accountability. But when the data is only available to the authorities, they can no longer judge the sustainability of a given model or infrastructure based on the environmental data. In the same way, the verifiability

[667] Ivi.
[668] Ivi.
[669] Ivi.
[670] Ivi.
[671] Ivi.

of corporate statements remains unclear, the participation of local communities in decision-making is limited, and the choice regarding the service with a lower environmental footprint will not be possible for consumers. Thus, the approach of relying solely on confidential disclosure to create an AI environmental governance framework may fail to provide administrative transparency and public accountability. Another critical point to consider is methodology. Energy can be obtained using several different approaches: through the data center, through the total server, and per GPU.[672] Data at the data center level can reflect total consumption, which includes all cooling costs, as well as overheads, but might not identify variation between model or architecture-specific costs. A simpler measurement, taken at the GPU level, will underestimate the total consumption as it only cares about one part of the server.[673] Hence, the authors propose using a cumulative measurement of consumption on a server and report a PUE value for the data center.[674]

The proposal is valuable because it shows how one can achieve this balance between technical rigor and administrative feasibility in environmental governance of AI systems. Too widely applicable a methodology prevents ' '. Too detailed may be hard to implement or may be inadequate in estimating the overall effect. Measuring at the server-cluster level tries to achieve a balance between these two goals: correlation of energy consumption with the model used or workload run, and a holistic understanding of the efficiency of the infrastructure. In the future, method selection will play a key role in future AI sustainability policies' legitimacy. This is a recent topic that can be found in literature that is not just focusing on Europe. As global environmental regulation of AI is somewhat fragmented and inadequate, says Ebert, Gamazaychikov, Hacker, and Luccioni.[675] Their comparative analysis finds that in most legal systems, regulation is centered around infrastructure level (the data centers) instead of the type of AI model used; the focus is mainly on training, while few systems outside the European Union impose any obligation to disclose AI-related energy use at all.[676] This creates a regulatory landscape of 'no single answer' and makes it difficult to compare the environmental footprint of the different models, with the potential to add new forms of regulatory arbitrage. The idea of regulatory arbitrage could well be important here. Companies may be tempted to place data centers, training runs, or inference services in places

[672] Ivi, pp. 3-4.
[673] Ivi.
[674] Ivi.
[675] Kai Ebert, Boris Gamazaychikov, Philipp Hacker e Sasha Luccioni, “The Global Landscape of Environmental AI Regulation: From the Cost of Reasoning to a Right to Green AI,” 2026, pp. 1-5.
[676] Ivi.

where the environmental transparency laws are less onerous. Then the impact of environmental regulation may come in the form of redistribution. At the international level, they suggest that mandatory transparency, including inference consumption, user rights to exclude unwanted generative integrations, and the ability to choose environmentally optimized models should be developed. computational location to avoid placing computationally heavy models in particularly sensitive areas, as well as to ensure that models are not operated in overly polluted facilities.[677]
This proposal makes it possible to link point 4.5 to point 4.3. To govern AI sustainability purely at national/regional level would be impractical since the digital infrastructures required are global, models can be trained in one country and used in another, chip supply chains can have multiple legs spanning countries and continents, and the impacts of AI can be externalized to other territories with weaker governance capacity. Thus, it is necessary to coordinate environmental governance at the international level, establish standards, and share responsibilities. Otherwise, if such resources are not present, sustainability could remain as a soft marketing branding for some businesses and as a pandering political agenda for some states, instead of being measurable and impactful on business industrial operations. But the European Union has another relevant instrument: the Critical Raw Materials Act. While it doesn't have a primary focus on AI, it is essential to consider how essential the technological ecosystem it relies on is for sustainability. The critical and strategic raw materials identified by the regulation (CRM) play a key role in the green and digital transition as well as in the defense and space sector, and 2030 target values are defined to boost European capacities in the value chain, which include 10% extraction, 40% processing, and 25% recycling, depending on these conditions.[678] This is a crucial but indirect tie-in with AI. Data centers, semiconductors, servers, and digital infrastructures underpin artificial intelligence systems and are essential for trade, key business infrastructure, and connectivity. An AI environmental governance geared toward electricity consumption alone, therefore, would miss a crucial part of the problem: The sustainability of the supply of hardware.[679] The Critical Raw Materials Act offers an opportunity to include aspects of supply security, product recycling, the circular economy, and reliance on fewer suppliers of critical materials into the considerations.[680]

[677] Ivi, pp. 1-2.
[678] Ivi.
[679] Ivi.
[680] Ivi.

AI governance also relates to European industrial geopolitics and claims technologically sustainable sovereignty in this context.[681]

This is especially true of the circular dimension. The Critical Raw Materials Act's goal of promoting a more sustainable and circular use of raw materials is addressed by promoting the recovery of these materials from extractive waste and the recovery of items containing critical raw materials, as well as by mitigating the negative effects on employment rights, human rights and environmental sustainability, including the recognition of certification schemes to advance the sustainability of raw materials entering the European market.[682]

Sustainability of models cannot be divorced from the sustainability of the materials and infrastructure that make models possible.[683] However, the Critical Raw Materials Act documents a tension which is typical of European governance: balancing sustainability, industrial resilience and competitiveness.[684] It makes the transition to environmental standards, recyclability and diversification of the supply chain, but it will also speed up the building of the extraction and industrial and processing capacity needed for the digital and green transition.[685] This is inevitable: decarbonization and digitalization consume additional infrastructure, chips, batteries, networks and materials.[686] The challenge is not to stop using materials, it is to manage them, minimize dependencies, material externalities and injustices along the global value chain.[687] In short, the application of AI in Europe on sustainability is a multi-layered approach, but still not yet completed. [688]The Energy Efficiency Directive and the Delegated Regulation on data centers aim to establish a reporting system and indicators, and to make a European database available at an infrastructure level.[689] The AI Act requires disclosure of some information about computational

---

[681] Ivi.
[682] Ivi.
[683] United Nations Environment Programme, *Artificial Intelligence (AI) End-to-End: The Environmental Impact of the Full AI Lifecycle Needs to Be Comprehensively Assessed* (Nairobi: UNEP, 2024), 2–4; OECD, *Measuring the Environmental Impacts of Artificial Intelligence Compute and Applications: The AI Footprint*, OECD Digital Economy Papers, no. 341 (Paris: OECD Publishing, 2022), 5–6.
[684] European Commission, "European Critical Raw Materials Act," accessed June 26, 2026.
[685] European Commission, "European Critical Raw Materials Act"; European Commission, "Critical Raw Materials Act," Internal Market, Industry, Entrepreneurship and SMEs.
[686] European Commission, "European Critical Raw Materials Act"; OECD, *Measuring the Environmental Impacts of Artificial Intelligence Compute and Applications*, 5–6.
[687] UNEP, *Artificial Intelligence (AI) End-to-End*, 2–4; UNCTAD, *Digital Economy Report 2024: Shaping an Environmentally Sustainable and Inclusive Digital Future* (Geneva: UNCTAD, 2024).
[688] Nicolas Alder, Kai Ebert, Ralf Herbrich, and Philipp Hacker, "AI, Climate, and Transparency: Operationalizing and Improving the AI Act," arXiv:2409.07471, 2024, 1–4.
[689] Directive (EU) 2023/1791 of the European Parliament and of the Council of 13 September 2023 on energy efficiency, art. 12; Commission Delegated Regulation (EU) 2024/1364 of 14 March 2024 establishing the first phase

resources and energy of general-purpose AI models at the model level.[690] The Critical Raw Materials Act tackles the challenge on the industrial side: Resilience of raw materials and supply chain.[691] There are also huge holes to be filled: indirect emissions, public disclosure, water, inference, methodological standards, AI applications with negative environmental impacts, e-waste, and compute location.[692]

The dimensions of sustainability of AI cannot only be examined in terms of technical dimensions, reporting requirements, energy efficiency tools, but also through environmental justice. As illustrated in the prior sections, AI depends upon material structures, such as data centers, electrical grid networks, water, critical minerals, semiconductors, international supply chains, and hardware disposal mechanisms. It is not evenly spread out around the world but rather located in geographies. It requires the use of local resources. It can generate geographies of conflict between digital development, environmental protection, energy security, and community rights. Therefore, territorial allocation of costs and benefits of digital transition must also be integrated into the environmental governance of AI: transparency must not only be specific to the models, but also to their costs and benefits.

The case studies presented by Latin America show in concrete terms how global digital infrastructure can induce local environmental conflicts, especially when data centers are in areas of water stress and/or climate vulnerability. The environmental proposal was critiqued by some of the residents and officials in Chile because of the project's potential impact on the fragile Central Santiago aquifer, which was already under pressure.[693] However, Google soon went on to change its original project and to start a new process, based on technology for air-cooling the data centers.[694] This case shows that the sustainability of data centers cannot be measured in an abstract

---

of a common Union rating scheme for data centres, recitals 8–10, art. 1 and Annex II. https://eur-lex.europa.eu/eli/reg_del/2024/1364/oj?locale=eng&utm

[690] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024, Artificial Intelligence Act, art. 53(1)(a) and Annex XI, Section 1, points 2(d)–(e). https://arxiv.org/abs/2409.07471?utm

[691] uropean Commission, "European Critical Raw Materials Act"; European Commission, "Critical Raw Materials Act," Internal Market, Industry, Entrepreneurship and SMEs.

[692] Alder et al., "AI, Climate, and Transparency," 2–4; OECD, *Measuring the Environmental Impacts of Artificial Intelligence Compute and Applications*, 18–21; UNEP, *Artificial Intelligence (AI) End-to-End*, 2–4.

[693] Reuters, "Chile Partially Pulls Google Data Center Permit, Seeks Tougher Environmental Checks," February 27, 2024. https://www.reuters.com/world/americas/chile-partially-pulls-google-data-center-permit-seeks-tougher-environmental-2024-02-27/

[694] Reuters, "Google Takes Chile Data Center Plans Back to Square One on Environmental Concerns," September 17, 2024. https://www.reuters.com/technology/google-takes-chile-data-center-plans-back-square-one-environmental-concerns-2024-09-17/

sense, or at normal temperatures; it must be measured against the specific climate and geographical context in which they are located.

As the Uruguay case shows, it's the same issue on the opposite side of the coin. The lesson to be learned, then, is that an air-cooled data center doesn't necessarily have no environmental impact but it may lead to an increase in CO_2 emissions or hazardous waste, or place increased demand on the nation's power supply.[695] This showcases how AI sustainability is not only defined by the service's general level of efficiency but also by the local environmental justice brought out by these technologies: AI services are global, but environmental costs are mostly localized.[696] The added factor of critical raw materials also applies here. The Critical Raw Materials Act has arisen within the European Union to enhance the security of supply of strategic materials required for its green and digital transition and to set targets for the extraction, processing, recycling, and import diversification by 2030.[697] But it is not only a matter of European strategic autonomy. Not only do resources needed to produce chips, servers, and datacenter power plants require global value chains, but they can also cause environmental and social effects in other countries. Therefore, AI Governance must address certain material traceability, environmental and social procurement criteria as well as reinforce recycling operations and be linked to the requirements of data center efficiency. AI's environmental impact can be significant, but when there are no tools present, it becomes a hidden element of the production chain.

More broadly, indirect effects and potential rebound effects need to be considered for AI sustainability. The OECD classifies the direct environmental effects of AI as those arising when AI systems and equipment are developed, used, or abandoned, and the indirect effects as positive or negative impacts caused by AI applications in other sectors where AI is not directly being used.[698] These distinctions show that AI can help achieve sustainable goals but can also boost consumption and harmful activities if the AI application of unsustainable processes is cheaper, faster or more efficient. Digital and material consumption can be increased by better

---

[695] David Feliba, "In Latin America, Data Center Plans Fuel Water Worries," *Context / Thomson Reuters Foundation*, September 26, 2023. https://www.context.news/big-tech/in-latin-america-data-center-plans-fuel-water-worries?utm

[696] Grace Livingstone, "Anger Mounts over Environmental Cost of Google Datacentre in Uruguay," *The Guardian*, August 1, 2024. https://www.theguardian.com/global-development/article/2024/aug/01/uruguay-anger-environmental-cost-google-datacentre-carbon-emissions-toxic-waste-water

[697] European Commission, "European Critical Raw Materials Act."

[698] OECD, *Measuring the Environmental Impacts of Artificial Intelligence Compute and Applications: The AI Footprint*, OECD Digital Economy Papers, no. 341 (Paris: OECD Publishing, 2022), 5–6.

recommendation systems, the distribution cost of goods that have a weak environmental footprint can be decreased by logistics optimizations, and optimization in the exploration of fossils by AI can enhance the efficiency of sectors that should be continuously diminished. Governance needs to therefore differentiate between "AI for climate action" and "AI for carbon-intensive optimization." This is reflected in the literature of machine learning and climate change. This is supported in the literature of machine learning and climate change. This crucial observation is important to prevent technological pessimism and naive optimism, as machine learning is not a "silver bullet"; there are many other technologies available that are useful for emission reductions and climate change adaptation in the sectors mentioned above, but have not yet been uniformly scaled up due to a lack of suitable policy, economic, and social choices.[699] To help the ecological transition, AI needs to be adopted in public policies, sectoral know-how, and group decisions to sufficiently lower emissions.[700]

It is important to note this because it connects AI to climate capacity-building. The most vulnerable countries could use AI tools to monitor climate risks, manage water resources, build improved early warning systems, build for adaptation, and optimize energy use.[701] But to make this possible, the only thing that is not enough is the provision of tools from technology. Capacity, access to data, digital infrastructure, training, participatory governance, and protection of data are all needed. However, AI technologies for climate action should be based on empowerment, not dependency, which aligns with a more transformative view of artificial intelligence. In that spirit, Artificial Intelligence for climate action is primarily about empowerment, not dependency. From this perspective, sustainability becomes a litmus test for global AI governance. International legal governance for artificial intelligence is decentralized and can be found in a variety of multilateral efforts, soft law, emerging treaties, and variegated regional approaches, as previously discussed in section 4.3. The environmental dimension is also likely to introduce an additional layer of fragmentation: certain states can be home to stringent environmental requirements relating to, amongst others, the use of data centers, emissions, water, and supply chains; others may offer more lax environmental requirements to bring in investment. In the absence of coordination, companies could move available resources, such as training, inference, or infrastructure, to less-regulated

[699] David Rolnick et al., "Tackling Climate Change with Machine Learning," arXiv:1906.05433, 2019, 1–4, 5–11.
[700] Rolnick et al., "Tackling Climate Change with Machine Learning," 4.
[701] UNFCCC, "#AI4ClimateAction."
https://unfccc.int/ttclear/templates/ttclear/pages/ttclear/pages/artificial_intelligence

countries. As a result, there is a need to arrive at minimum international standards on environmental disclosure, energy usage, water usage, emissions, E-waste, and supply chain traceability.

Sustainability will become a more prominent topic in the governance of AI. The growth of AI systems will demand physical facilities, energy, water, critical raw materials, and global supply chains that will have visible effects on local places and communities. Ethical, legal, geopolitical, or security issues must, therefore, not be the only dimension considered on which to evaluate AI, but also material dimensions, since it is where technologies can be developed and used. However, AI's environmental footprint and sustainability impact will only be realized when this AI infrastructure is brought under ecological accountability measures, and when AI supports countries in managing climate change, rather than making them more technologically dependent and vulnerable to external events; when it helps countries to have a better efficiency while reducing the effects of rebound; and when it helps countries better predict how much demand will change over time without uncontrolled growth. This is where true AI governance comes in: not only is it essential to consider how to deal with the dangers after AI has been mastered, but also which type of technological growth and development is desired.

Conclusion

The thesis has explored the impact of AI on modern diplomacy, showing that it is not just a new technical instrument, but a greater shift in how states collect information, develop their strategic knowledge, communicate, negotiate and set their governance agenda. One of the major findings of the analysis is that artificial intelligence should not only be seen as a means of diplomacy but also as a diplomacy itself. It can assist diplomatic activities via data analysis, document management, translation, crisis monitoring, consular services, public communication, on the other hand. However, it has emerged as a key topic on the global agenda with implications for security, fundamental rights, economic competition, regulation of digital markets, environmental sustainability and the distribution of power between states. The first chapter has shown that the current transformation that is enabled by Artificial Intelligence is not happening in isolation but is a wider transformation of diplomacy that is digital. The changes that the internet, social media, and digital information are making to modern diplomacy are already happening. They were further accelerated with the help of the pandemic caused by the novel coronavirus. It has shown that the diplomatic processes could function well online, yet with all the advantages of face-to-face communication. In this context, AI is an additional development, changing not only how diplomatic communication takes place, but also the thoughts and decisions that precede it. Moreover, according to the thesis, the use of AI should be considered as an auxiliary technology, not a substitute for people’s decisions. Full automation in diplomacy is a very unrealistic and dangerous concept, politically speaking. The efficient processing of information is not all of diplomacy, but it also requires interpretation, prudence, accountability, the ability to understand context, to evaluate actors' intentions and to build trust. Thus, AI can only foster the improvement of diplomatic efforts when implemented in a model of "augmented diplomacy" that takes technology as an enabler rather than a replacement for the human component but does not remove it from the equation. In the second chapter, we showed that AI is now a key element of state-to-state strategic competition. In the second chapter, we showed that AI has become an important part of state-to-state strategic competition. The power of international actors is no longer just military, economic or territorial, but also based on their ability to collect information, manage digital

infrastructure, create algorithmic models, control information flow, and influence technological standards. AI plays a role in reshaping power dynamics in the international order. States and tech giants who hold on to data, cloud services, computing resources, semiconductors and advanced models have a tremendous strategic advantage. On the other hand, actors lacking in technological resources have the danger of relying on infrastructure and systems created in other places. The AI divide is not only diplomatic, economic, and regulatory, but also technical. In the military field and hybrid warfare this competition assumes a particularly sensitive aspect. In addition, AI will assist in improving intelligence, surveillance, cybersecurity, targeting, logistics, command and control, yet it may lead to quick decision-making and escalation. At the same time, AI can be used for creating misinformation, deep fakes, public opinion manipulation, and other influence operations. It means that it is evident that international security does not remain only on the battlefield anymore since it is also moving into the information realm, where perceptions, trust, and political consensus are formed.

The ethical and political consequences of using AI in diplomacy were examined in the third chapter. These include, among other problems, algorithmic bias, absence of transparency, absence of accountability and respect for basic rights and democratic legitimacy. The use of AI in institutions and diplomacy may bring about the possibility of faster decision-making and analysis of complicated documents; however, it may lead to discrimination, untraceable errors and loss of control by humans.

The problem is not just technical, it's political and democratic. In a situation where an automated system has an impact on a public decision, it should be possible to explain, understand and know who is responsible for the decision and what is in place to dispute it. Disinformation and deepfakes are a good example of the relationship between AI, rights, and democracy in this context. Fake news generated or amplified by artificial intelligence can negatively impact the quality of public discourse, influence perceptions of foreign events and create a loss of faith in institutions. In the case of diplomacy, the task is to manage the information space, security, credibility and legitimacy. One of the key challenges in digital diplomacy and AI governance will thus be the capacity to counter with information manipulation without infringing on freedom of expression. The fourth chapter examined governance reactions to AI, including the European Union, which is one of the most advanced attempts to govern AI based on a risk-based, fundamental rights protecting and actor-s accountable approach. The AI Act is part of the EU's "trust, transparency and protection of

the individual" digital governance model, in addition to the GDPR and other European digital laws. As seen in this comparison however, there is no single global model for the regulation of AI. The Chinese model seems more centralized and more oriented towards state control, while the US model seems more fragmented and market oriented. This diversity reflects the fact that AI governance is also a game of regulatory competition. The thesis has also outlined the challenges of global governance of AI.

The UN, G7, UNESCO, and the Council of Europe are involved in efforts towards agreeing on common principles and dialogue platforms; however, the international system remains divided. Principles such as transparency, safety, reliability, being human-centered, sustainable, and upholding human rights are commonly accepted by the international community; however, there are often lacking enforcement measures, authorities to make demands globally, and tools to control transnational actors in the private sector. AI diplomacy is at an intermediary stage: While there is a common language, there is not a global regime yet.

An additional significant outcome is in the field of sustainability. While usually considered an intangible technology, it may be said that AI is a tangible technology, since it relies on tangible elements such as infrastructure, data centers, power grids, cooling facilities, water, semiconductors, rare earth metals, and global supply chains. Therefore, the issues of rights, security, and innovation are not enough for AI governance; it's a matter of energy, emissions, water, e-waste, and territorial impact, too. AI sustainability issues aren't just about the efficiency of the model; they're also about the sustainability of the materials used to build it and its associated infrastructure. Introducing AI technology is not only about the efficiency of the technology itself, but also about the sustainability of the materials that enable the AI to operate and the infrastructure that it is part of. Based on this analysis, the thesis concludes that AI is not a threat to diplomacy, but it is a reformer of its conditions, instruments, and issues. It might assist in developing the capability of the diplomatic entities in terms of comprehending the complexity, risk prediction, information control, and communication improvement. It might exacerbate the issue of inequality, lack of transparency, dependency on technology, vulnerabilities of democratic processes, and environmental consequences. It is not whether artificial intelligence should be applied to the realm of diplomacy, but how.

AI must be human-centered, transparent, responsibly managed within institutions, and subjected to critical evaluation in accordance with democratic diplomacy's principles. The future needs

diplomats who are not IT experts, but who are also techno-literate, developing a new kind of technological literacy that involves understanding the limitations of AI systems, detecting biases and hallucinations, verifying sources, safeguarding sensitive data, and distinguishing between analytical support and political decision-making. Meanwhile, states need to develop their capacities within, draw up clear rules, and encourage forms of international cooperation that do not allow AI to be a tool of domination or exclusion. Finally, artificial intelligence is one of the most defining challenges in 21st-century diplomacy as it brings together innovation, power, ethics, and governance. It provides a new vista of more informed, quicker, and more effective diplomacy for dealing with the complexities of the world. But these opportunities can only be achieved when AI is treated as a political and social technology, not a neutral one.

Institutions will not decide the future of diplomacy by means of the automation of decision-making, but by their capacity to combine technological innovation with responsibility, rights, sustainability, and trust. While AI can help with diplomacy, it can't replace the need for human judgment, political responsibility, and relationship-building between international actors.

Bibliography

- Afandi, Ahmad Ali, and Badrus Sholeh. “The Construction of Artificial Intelligence as an International Security Issue in the United Nations Agenda 2025.” *Jurnal Transformasi Global* 12, no. 2 (2025): 168–185. file:///C:/Users/grant/Downloads/7.+The+Construction+of+Artificial+Intelligence+as+an+International+Security+Issue+A+Constructivist+in+the+United+Nations+Agenda+2025.docx%20(1).pdf
- Ahmed, Mustak. “Artificial Intelligence, Nuclear Warfare and Towards Laws: A Study on Limitations of Existing International Laws and Treaties.” *Preprints.org* 30 (2025). https://www.researchgate.net/profile/Mustak-Ahmed/publication/393172654_Artificial_Intelligence_Nuclear_Warfare_and_Towards_Laws_A_Study_on_Limitations_of_Existing_International_Laws_and_Treaties/links/6862cd6c39c35835120593b8/Artificial-Intelligence-Nuclear-Warfare-and-Towards-Laws-A-Study-on-Limitations-of-Existing-International-Laws-and-Treaties.pdf
- Bjerde, Anna, et al. *Artificial Intelligence and International Affairs: Disruption Anticipated*. London: Chatham House, 2020. https://www.chathamhouse.org/sites/default/files/publications/research/2018-06-14-artificial-intelligence-international-affairs-cummings-roff-cukier-parakilas-bryce.pdf
- Bjola, Corneliu. “Digital Diplomacy as World Disclosure: The Case of WikiLeaks.” *International Studies Perspectives* 16, no. 3 (2015): 311–330. https://ora.ox.ac.uk/objects/uuid:4b13205a-1002-49c3-aeb0-1ed0351043b9/files/m80b2ede8cee9d8344f2e9608d7f6bf9f
- Bjola, Corneliu, and Marcus Holmes, eds. *Digital Diplomacy: Theory and Practice*. Abingdon: Routledge, 2015.
- Bjola, Corneliu, and Markus Kornprobst. *Understanding International Diplomacy: Theory, Practice and Ethics*. 2nd ed. London: Routledge, 2018.
- Bjola, Corneliu, and Ilan Manor, eds. *The Oxford Handbook of Digital Diplomacy*. Oxford: Oxford University Press.
- Brasioli, Diego, Laura Guercio, Giovanna Gnerre Landini, and Andrea de Giorgio, eds. *The Routledge Handbook of Artificial Intelligence and International Relations*. Abingdon: Routledge, 2025. file:///C:/Users/grant/Downloads/7.+The+Construction+of+Artificial+Intelligence+as+an+International+Security+Issue+A+Constructivist+in+the+United+Nations+Agenda+2025.docx.pdf
- Council of Europe. *Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law*. Strasbourg: Council of Europe, 2023.
- Dev, Rishik. “AI World Order: How Artificial Intelligence Is Reshaping Global Authority and Legitimacy.” 2025. https://preprints.apsanet.org/engage/apsa/article-details/69332c8f65a54c2d4a513184
- DiploFoundation. “Digital Diplomacy in 2025: Geopolitics, New Topics & Tools.” https://www.diplomacy.edu/topics/digital-diplomacy/
- DiploFoundation. “Technology and Diplomacy.” https://www.diplomacy.edu/technology-and-diplomacy/

- European Commission. *General Data Protection Regulation (GDPR)*. Brussels: European Commission, 2018. https://commission.europa.eu/law/law-topic/data-protection_en
- European Commission. *White Paper on Artificial Intelligence: A European Approach to Excellence and Trust*. Brussels: European Commission, 2020. https://commission.europa.eu/document/download/d2ec4039-c5be-423a-81ef-b9e44e79825b_en?filename=commission-white-paper-artificial-intelligence-feb2020_en.pdf
- European Commission. *Proposal for a Regulation Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act)*. Brussels: European Commission, 2021.
- European External Action Service (EEAS). *Digital Diplomacy and AI Governance Frameworks*. Brussels: EEAS, 2023.
- European Parliament. *Artificial Intelligence Act: Final Text*. Brussels: European Parliament, 2024.
- European Parliament Research Service. *Artificial Intelligence and Diplomacy in the EU Context*. Brussels: EPRS, 2022.
- Floridi, Luciano, Josh Cowls, Monica Beltrametti, Raja Chatila, Patrice Chazerand, Virginia Dignum, Christoph Luetge, et al. "AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations." *Minds and Machines* 28, no. 4 (2018): 689–707. https://link.springer.com/content/pdf/10.1007/s11023-018-9482-5.pdf
- Garcia, Denise. "The Global Diplomacy of Governing Military Artificial Intelligence." *Ethics and Information Technology* 27, no. 4 (2025): 55. https://link.springer.com/content/pdf/10.1007/s10676-025-09863-0.pdf
- Hamidouche, Karim. "Artificial Intelligence: A New Tool for Diplomats." In *Artificial Intelligence and Digital Diplomacy*, edited by Fatima Roumate. Cham: Springer, 2021.
- High-Level Expert Group on Artificial Intelligence. *Ethics Guidelines for Trustworthy AI*. Brussels: European Commission, 2019.
- Horowitz, Michael C. "Artificial Intelligence, International Competition, and the Balance of Power." *Texas National Security Review* 1, no. 3 (2018): 37–57. https://muse.jhu.edu/pub/15/article/964758/pdf
- Klimburg, Alexander. *The Darkening Web: The War for Cyberspace*. New York: Penguin Press, 2017. https://www.idsa.in/system/files/jds/jds_11_4_2017_the-darkening-we-the-war-for-cyberspace.pdf?utm
- Manor, Ilan. *The Digitalization of Public Diplomacy*. Cham: Palgrave Macmillan, 2019. https://www.researchgate.net/profile/Ilan-Manor/publication/330377221_The_Digitalization_of_Public_Diplomacy/links/5e29af5d299bf152167858e8/The-Digitalization-of-Public-Diplomacy.pdf
- Miailhe, Nicholas, and Yolanda Lannquist. "Global Governance of Artificial Intelligence." In *Handbook of Artificial Intelligence and Robotic Process Automation: Policy and Government Applications*, 23–30, 2020. https://books.google.it/books?hl=en&lr=&id=hokIEAAAQBAJ&oi=fnd&pg=PA23&dq=World+Economic+Forum.+(2020).+Global+Governance+of+Artificial+Intelligence.&ots=PbgkqXMSXM&sig=YNNeC82NdN4O-

4p4sWj6UvPuhtg&redir_esc=y#v=onepage&q=World%20Economic%20Forum.%20(2020).%20Global%20Governance%20of%20Artificial%20Intelligence.&f=false

- Morandín-Ahuerma, Fabio. "Ten UNESCO Recommendations on the Ethics of Artificial Intelligence." *OSF Preprints* (2023). file:///C:/Users/grant/Downloads/(3)%20TEN%20UNESCO%20RECOMMENDATIONS.pdf
- OECD. *OECD Principles on Artificial Intelligence*. Paris: OECD, 2019. https://oecd.ai/en/ai-principles
- Roumate, Fatima, ed. *Artificial Intelligence and Digital Diplomacy*. Cham: Springer Nature, 2021.
- Seib, Philip. *Real-Time Diplomacy*. New York: Palgrave Macmillan, 2012.
- Tallinn Paper No. 14. *Cyber Diplomacy: From Concept to Practice*. https://ccdcoe.org/uploads/2024/06/Tallinn_Papers_Cyber_Diplomacy_From_Concept_to_Practice_Christou.pdf
- UNCTAD. *Digital Economy Report 2021*. Geneva: United Nations Conference on Trade and Development, 2021. https://unctad.org/publication/digital-economy-report-2021
- UNESCO. *Recommendation on the Ethics of Artificial Intelligence*. Paris: UNESCO, 2021. https://www.unesco.org/en/artificial-intelligence/recommendation-ethics
- UNESCO. "Recommendation on the Ethics of Artificial Intelligence." 2021. https://bsu.buap.mx/b2m
- Ziani, K. "Security and Reliability in Digital Diplomacy: Review of Selected Literature." *Humanities and Social Sciences* 32, no. 3 (2025): 139–152. https://journals.prz.edu.pl/hss/article/download/2064/1537/
- Agenda Digitale. "Digital Omnibus: come cambia il GDPR per AI e ricerca." https://www.agendadigitale.eu/sicurezza/privacy/digital-omnibus-come-cambia-il-gdpr-per-ai-e-ricerca/. Last accessed June 18, 2026.
- AI Act Service Desk. "Annex XI: Technical Documentation Referred to in Article 53(1), Point (a) — Technical Documentation for Providers of General-Purpose AI Models." https://ai-act-service-desk.ec.europa.eu/en/ai-act/annex-11. Last accessed June 18, 2026.
- AI Act Service Desk. "Article 5: Prohibited AI Practices." https://ai-act-service-desk.ec.europa.eu/en/ai-act/article-5. Last accessed June 18, 2026.
- AI Act Service Desk. "What Systems Are Prohibited under Article 5 AI Act?" https://ai-act-service-desk.ec.europa.eu/en/ai-act/faq/what-systems-are-prohibited-under-article-5-ai-act-eg-social-scoring-emotion-recognition. Last accessed June 18, 2026.
- Alder, Nicolas, Kai Ebert, Ralf Herbrich, and Philipp Hacker. "AI, Climate, and Transparency: Operationalizing and Improving the AI Act." arXiv:2409.07471, 2024. https://arxiv.org/abs/2409.07471.
- Bommasani, Rishi, Drew A. Hudson, Ehsan Adeli, Russ Altman, Simran Arora, Sydney von Arx, Michael S. Bernstein, et al. *On the Opportunities and Risks of Foundation Models*. Stanford University, 2021. https://arxiv.org/abs/2108.07258.
- Bradford, Anu. *The Brussels Effect: How the European Union Rules the World*. Oxford: Oxford University Press, 2020.

- Central Statistics Office Ireland. “Data Centres Metered Electricity Consumption 2024 — Key Findings.” June 10, 2025. https://www.cso.ie/en/releasesandpublications/ep/p-dcmec/datacentresmeteredelectricityconsumption2024/keyfindings/. Last accessed June 18, 2026.
- China Daily/Xinhua. “346 Generative AI Services Filed with Cyberspace Administration of China.” April 9, 2025. https://www.chinadaily.com.cn/a/202504/09/WS67f5e234a3104d9fd381e55e.html. Last accessed June 18, 2026.
- Chun, Jon, Christian Schroeder de Witt, and Katherine Elkins. *Comparative Global AI Regulation: Policy Perspectives from the EU, China, and the US*. 2024. https://arxiv.org/abs/2410.21279.
- Comparative AI. “Interim Measures for the Management of Generative AI Services.” 2023. https://comparativeai.org/en/rules/china/generative-ai-interim-measures/. Last accessed June 18, 2026.
- Congressional Research Service. *Regulation of TikTok under the Protecting Americans from Foreign Adversary Controlled Applications Act: Analysis of Selected Legal Issues*. LSB11127. Updated May 9, 2024. https://crsreports.congress.gov/product/pdf/LSB/LSB11127. Last accessed June 18, 2026.
- Council of Europe. *Council of Europe Opens First Ever Global Treaty on AI for Signature*. September 5, 2024. https://www.coe.int/en/web/portal/-/council-of-europe-opens-first-ever-global-treaty-on-ai-for-signature. Last accessed June 18, 2026.
- Council of Europe. *Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law*. 2024. https://www.coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence. Last accessed June 18, 2026.
- Crilley, Rhys, and Robert A. Saunders. “The Resonant Slop Machine: Public Diplomacy and Strategic Narratives in the Age of Artificial Intelligence.” *Place Branding and Public Diplomacy*, 2025. https://doi.org/10.1057/s41254-025-00419-z.
- Cyberspace Administration of China. “Provisions on the Administration of Algorithmic Recommendations for Internet Information Services.” 2022. https://www.cac.gov.cn/2022-01/04/c_1642894606364259.htm. Last accessed June 18, 2026.
- DigiChina. “Translation: Internet Information Service Algorithmic Recommendation Management Provisions — Effective March 1, 2022.” Stanford University. https://digichina.stanford.edu/work/translation-internet-information-service-algorithmic-recommendation-management-provisions-effective-march-1-2022/. Last accessed June 18, 2026.
- Di Martino, Luigi, and Heather Ford. “Navigating Uncertainty: Public Diplomacy vs. AI.” *Place Branding and Public Diplomacy* 21 (2025): 350–354. https://doi.org/10.1057/s41254-024-00330-z.
- DiploFoundation. *Mapping the Challenges and Opportunities of Artificial Intelligence for the Conduct of Diplomacy*. Geneva: DiploFoundation, 2019. https://www.diplomacy.edu/AI-diplo-report. Last accessed June 18, 2026.
- Diritto Bancario. “Digital Omnibus e AI Act: verso un sistema di governance integrato?” https://www.dirittobancario.it/art/digital-omnibus-e-ai-act-verso-un-sistema-di-governance-integrato/. Last accessed June 18, 2026.

- DSGVO-Gesetz. *Datenschutz-Grundverordnung*. https://dsgvo-gesetz.de/. Last accessed June 18, 2026.
- Ebert, Kai, Boris Gamazaychikov, Philipp Hacker, and Sasha Luccioni. "The Global Landscape of Environmental AI Regulation: From the Cost of Reasoning to a Right to Green AI." 2026. https://arxiv.org/abs/2603.00068.
- European Commission. *Approval of the Content of the Draft Communication from the Commission — Guidelines on the Scope of the Obligations for General-Purpose AI Models Established by Regulation (EU) 2024/1689*. Brussels, July 18, 2025. C(2025) 5045 final.
- European Commission. *Artificial Intelligence Act*. https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai. Last accessed June 18, 2026.
- European Commission. *Commission Adopts EU-Wide Scheme for Rating Sustainability of Data Centres*. March 15, 2024. https://energy.ec.europa.eu/news/commission-adopts-eu-wide-scheme-rating-sustainability-data-centres-2024-03-15_en. Last accessed June 18, 2026.
- European Commission. *Commission Delegated Regulation (EU) 2024/1364 of 14 March 2024 on the First Phase of the Establishment of a Common Union Rating Scheme for Data Centres*. https://eur-lex.europa.eu/eli/reg_del/2024/1364/oj. Last accessed June 18, 2026.
- European Commission. *Commission Guidelines on the Scope of the Obligations for Providers of General-Purpose AI Models Established by Regulation (EU) 2024/1689 (AI Act)*. Brussels, November 19, 2025. C(2025) 7719 final.
- European Commission. *Commission Opens Formal Proceedings against TikTok under the Digital Services Act*. February 19, 2024. https://digital-strategy.ec.europa.eu/en/news/commission-opens-formal-proceedings-against-tiktok-under-digital-services-act. Last accessed June 18, 2026.
- European Commission. *Commission Preliminarily Finds TikTok and Meta in Breach of Their Transparency Obligations under the Digital Services Act*. October 24, 2025. https://digital-strategy.ec.europa.eu/en/news/commission-preliminarily-finds-tiktok-and-meta-breach-their-transparency-obligations-under-digital. Last accessed June 18, 2026.
- European Commission. *Commission Preliminarily Finds TikTok's Addictive Design in Breach of the Digital Services Act*. February 6, 2026. https://digital-strategy.ec.europa.eu/en/news/commission-preliminarily-finds-tiktoks-addictive-design-breach-digital-services-act. Last accessed June 18, 2026.
- European Commission. *Data Protection*. https://commission.europa.eu/law/law-topic/data-protection_en. Last accessed June 18, 2026.
- European Commission. *Energy Performance of Data Centres*. https://energy.ec.europa.eu/topics/energy-efficiency/energy-efficiency-targets-directive-and-rules/energy-efficiency-directive/energy-performance-data-centres_en. Last accessed June 18, 2026.
- European Commission. *European Critical Raw Materials Act*. https://commission.europa.eu/topics/competitiveness/green-deal-industrial-plan/european-critical-raw-materials-act_en. Last accessed June 18, 2026.
- European Commission. *Hiroshima Process International Code of Conduct for Organizations Developing Advanced AI Systems*. October 30, 2023. https://digital-

strategy.ec.europa.eu/en/library/hiroshima-process-international-code-conduct-advanced-ai-systems. Last accessed June 18, 2026.

- European Commission. *Hiroshima Process International Guiding Principles for Organizations Developing Advanced AI Systems*. October 30, 2023. https://digital-strategy.ec.europa.eu/en/library/hiroshima-process-international-guiding-principles-advanced-ai-system. Last accessed June 18, 2026.
- European Data Protection Board. *Facial Recognition: Italian SA Fines Clearview AI EUR 20 Million*. 2022. https://www.edpb.europa.eu/news/national-news/2022/facial-recognition-italian-sa-fines-clearview-ai-eur-20-million_en. Last accessed June 18, 2026.
- European Data Protection Board. *Report of the Work Undertaken by the ChatGPT Taskforce*. May 23, 2024. https://www.edpb.europa.eu/our-work-tools/our-documents/other/report-work-undertaken-chatgpt-taskforce_en. Last accessed June 18, 2026.
- European External Action Service. *Digital Diplomacy: Secure and Competitive Europe*. https://www.eeas.europa.eu/eeas/digital-diplomacy-secure-and-competitive-europe_en. Last accessed June 18, 2026.
- European Parliament. *Artificial Intelligence Act: MEPs Adopt Landmark Law*. 2024. https://www.europarl.europa.eu/news/en/press-room/20240308IPR19015/artificial-intelligence-act-meps-adopt-landmark-law. Last accessed June 18, 2026.
- European Parliament. *EU AI Act: First Regulation on Artificial Intelligence*. 2023. https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence. Last accessed June 18, 2026.
- European Parliament and Council of the European Union. *Regulation (EU) 2016/679 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data (General Data Protection Regulation)*. *Official Journal of the European Union*, 2016. https://eur-lex.europa.eu/eli/reg/2016/679/oj. Last accessed June 18, 2026.
- European Parliament and Council of the European Union. *Regulation (EU) 2024/1689 Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act)*. *Official Journal of the European Union*, 2024. http://data.europa.eu/eli/reg/2024/1689/oj. Last accessed June 18, 2026.
- Federal Trade Commission. *FTC Authorizes Compulsory Process for AI-related Products and Services*. November 21, 2023. https://www.ftc.gov/news-events/news/press-releases/2023/11/ftc-authorizes-compulsory-process-ai-related-products-services. Last accessed June 18, 2026.
- Federal Trade Commission. *FTC Warns about Misuses of Biometric Information and Harm to Consumers*. May 18, 2023. https://www.ftc.gov/news-events/news/press-releases/2023/05/ftc-warns-about-misuses-biometric-information-harm-consumers. Last accessed June 18, 2026.
- Federal Trade Commission. *Generative AI Raises Competition Concerns*. June 29, 2023. https://search.ftc.gov/policy/advocacy-research/tech-at-ftc/2023/06/generative-ai-raises-competition-concerns. Last accessed June 18, 2026.
- Feliba, David. "In Latin America, Data Center Plans Fuel Water Worries." Reuters/Thomson Reuters Foundation, September 26, 2023. https://www.reuters.com/article/technology/feature-in-latin-america-data-center-plans-fuel-water-worries-idUSL8N3AU1PY/. Last accessed June 18, 2026.

- Garante per la Protezione dei Dati Personali. *Intelligenza artificiale: il Garante blocca ChatGPT. Raccolta illecita di dati personali. Assenza di sistemi per la verifica dell'età dei minori*. March 31, 2023. https://www.garanteprivacy.it/home/docweb/-/docweb-display/docweb/9870847. Last accessed June 18, 2026.
- Garante per la Protezione dei Dati Personali. *Provvedimento del 30 marzo 2023, n. 112. Limitazione provvisoria del trattamento dei dati personali nei confronti di OpenAI L.L.C. (ChatGPT)*. Doc. web n. 9870832. https://www.garanteprivacy.it/home/docweb/-/docweb-display/docweb/9870832. Last accessed June 18, 2026.
- Garante per la Protezione dei Dati Personali. *Provvedimento nei confronti di Clearview AI Inc.* 2022. Doc. web n. 9751362. https://www.garanteprivacy.it/home/docweb/-/docweb-display/docweb/9751362. Last accessed June 18, 2026.
- Hellenic Republic Ministry of Foreign Affairs. *Chatbot — Virtual Assistant of the Consular Authority*. https://www.mfa.gr/en/services/digital-tools/chatbot/. Last accessed June 18, 2026.
- Hine, Emmie. "Governing Silicon Valley and Shenzhen: Assessing a New Era of Artificial Intelligence Governance in the United States and China." *Digital Society* 3, no. 50 (2024). https://doi.org/10.1007/s44206-024-00138-7.
- International Committee of the Red Cross. *Autonomous Weapons*. https://www.icrc.org/en/law-and-policy/autonomous-weapons. Last accessed June 18, 2026.
- International Committee of the Red Cross. *Autonomous Weapon Systems and International Humanitarian Law: Selected Issues*. https://www.icrc.org/en/article/autonomous-weapon-systems-and-international-humanitarian-law-selected-issues. Last accessed June 18, 2026.
- International Committee of the Red Cross. *International Committee of the Red Cross Position on Autonomous Weapon Systems: ICRC Position and Background Paper*. *International Review of the Red Cross* 102, no. 915 (2021): 1335–1349. https://www.icrc.org/en/document/icrc-position-autonomous-weapon-systems. Last accessed June 18, 2026.
- International Committee of the Red Cross. *Submission on Autonomous Weapon Systems to the United Nations Secretary-General*. 2024. https://www.icrc.org/en/document/autonomous-weapons-icrc-submits-recommendations-un-secretary-general. Last accessed June 18, 2026.
- International Energy Agency. *Energy and AI*. Paris: IEA, 2025. https://www.iea.org/reports/energy-and-ai. Last accessed June 18, 2026.
- International Energy Agency. *Energy Demand from AI*. In *Energy and AI*. 2025. https://www.iea.org/reports/energy-and-ai/energy-demand-from-ai. Last accessed June 18, 2026.
- Jung, Won Kyung, and Hun Yeong Kwon. "Privacy and Data Protection Regulations for AI Using Publicly Available Data: Clearview AI Case." In *ICEGOV '24: Proceedings of the 17th International Conference on Theory and Practice of Electronic Governance*, 48–55. New York: Association for Computing Machinery, 2024. https://doi.org/10.1145/3680127.3680200.
- Lee, Jyh-An, and Peng Zhou. "FRT Regulation in China." In *The Cambridge Handbook of Facial Recognition in the Modern State*. Cambridge: Cambridge University Press, 2023. https://doi.org/10.1017/9781009321211.020.

- Livingstone, Grace. “Anger Mounts over Environmental Cost of Google Datacentre in Uruguay.” *The Guardian*, August 1, 2024. https://www.theguardian.com/global-development/article/2024/aug/01/uruguay-anger-environmental-cost-google-datacentre-carbon-emissions-toxic-waste-water. Last accessed June 18, 2026.
- Maslej, Nestor, Loredana Fattorini, Raymond Perrault, Vanessa Parli, Anka Reuel, Erik Brynjolfsson, John Etchemendy, et al. *The AI Index 2025 Annual Report*. Stanford, CA: Stanford Institute for Human-Centered Artificial Intelligence, 2025. https://doi.org/10.48550/arXiv.2504.07139.
- Ministry of Foreign Affairs of Ukraine. *The Ministry of Foreign Affairs of Ukraine Has Appointed a Digital Person to Inform on Consular Issues*. May 1, 2024. https://mfa.gov.ua/en/news/mzs-ukrayini-priznachilo-cifrovu-osobu-dlya-informuvannya-shchodo-konsulskih-pitan. Last accessed June 18, 2026.
- National Institute of Standards and Technology. *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*. NIST AI 100-1. January 2023. https://doi.org/10.6028/NIST.AI.100-1.
- National Institute of Standards and Technology. *Face Recognition Vendor Test (FRVT)*. https://www.nist.gov/programs-projects/face-recognition-vendor-test-frvt. Last accessed June 18, 2026.
- OECD. *Measuring the Environmental Impacts of Artificial Intelligence Compute and Applications: The AI Footprint*. OECD Digital Economy Papers, no. 341. Paris: OECD Publishing, 2022.
- OECD. *OECD Launches Pilot to Monitor Application of G7 Code of Conduct on Advanced AI Development*. July 22, 2024. https://www.oecd.org/en/about/news/press-releases/2024/07/oecd-launches-pilot-to-monitor-application-of-g7-code-of-conduct-on-advanced-ai-development.html. Last accessed June 18, 2026.
- Okolo, Chinasa T., and Mubarak Raji. *The Global Majority in International AI Governance*. 2026. https://arxiv.org/abs/2601.17191.
- Pavlidis, Georgios. “The EU AI Act and the Rights-Based Approach to Technological Governance.” *Review of European and Comparative Law*, 2026. https://doi.org/10.31743/recl.19283.
- Reuters. “Chile Partially Pulls Google Data Center Permit, Seeks Tougher Environmental Checks.” February 27, 2024. https://www.reuters.com/world/americas/chile-partially-pulls-google-data-center-permit-seeks-tougher-environmental-2024-02-27/. Last accessed June 18, 2026.
- Reuters. “China Says Facial Recognition Should Not Be Forced on Individuals.” March 21, 2025. https://www.reuters.com/technology/china-says-facial-recognition-should-not-be-forced-individuals-2025-03-21/. Last accessed June 18, 2026.
- Reuters. “Google Takes Chile Data Center Plans Back to Square One on Environmental Concerns.” September 17, 2024. https://www.reuters.com/technology/google-takes-chile-data-center-plans-back-square-one-environmental-concerns-2024-09-17/. Last accessed June 18, 2026.
- Reuters. “US, Britain, EU to Sign First International AI Treaty.” September 5, 2024. https://www.reuters.com/technology/artificial-intelligence/us-britain-eu-sign-agreement-ai-standards-ft-reports-2024-09-05/. Last accessed June 18, 2026.

- Reviglio, Urbano, and Giulio Santoni. “Governing Platform Recommender Systems in Europe: Insights from China.” *Global Jurist* 23, no. 2 (2023): 151–181. https://doi.org/10.1515/gj-2023-0013.
- Rolnick, David, Priya L. Donti, Lynn H. Kaack, Kelly Kochanski, Alexandre Lacoste, Kris Sankaran, Andrew Slavin Ross, et al. “Tackling Climate Change with Machine Learning.” arXiv:1906.05433, 2019. https://arxiv.org/abs/1906.05433.
- Ruschemeier, Hannah, and Jascha Bareis. “Searching for Harmonised Rules: Understanding the Paradigms, Provisions and Pressing Issues in the Final EU AI Act.” In *Digital Decade: How the EU Shapes Digitalisation Research*, edited by Gsenger and Sekwenz. Baden-Baden: Nomos, 2025.
- Say, Mark. “Foreign Secretary Points to Future of AI in Diplomacy.” UKAuthority, July 14, 2025. https://www.ukauthority.com/articles/foreign-secretary-points-to-future-of-ai-in-diplomacy. Last accessed June 18, 2026.
- Smuha, Nathalie A. “From a ‘Race to AI’ to a ‘Race to AI Regulation’: Regulatory Competition for Artificial Intelligence.” *Law, Innovation and Technology* 13, no. 1 (2021): 57–84. https://doi.org/10.1080/17579961.2021.1898300.
- Stanzel, Volker, and Daniel Voelsen. *Diplomacy and Artificial Intelligence: Reflections on Practical Assistance for Diplomatic Negotiations*. SWP Research Paper 1. Berlin: Stiftung Wissenschaft und Politik, 2022. https://doi.org/10.18449/2022RP01.
- Stoyanova, Vladislava. “Council of Europe Framework Convention on Artificial Intelligence: Context, Regulatory Approach and Scope of Obligations.” *European Journal of Risk Regulation* 17 (2026): 536–565. https://doi.org/10.1017/err.2025.10070.
- The Guardian. “Ukraine Unveils AI-Generated Foreign Ministry Spokesperson.” May 3, 2024. https://www.theguardian.com/technology/article/2024/may/03/ukraine-ai-foreign-ministry-spokesperson. Last accessed June 18, 2026.
- The White House. *White House Unveils America’s AI Action Plan*. July 23, 2025. https://www.whitehouse.gov/releases/2025/07/white-house-unveils-americas-ai-action-plan/. Last accessed June 18, 2026.
- UK Government. *Seoul Declaration for Safe, Innovative and Inclusive AI by Participants Attending the Leaders’ Session: AI Seoul Summit, 21 May 2024*. 2024. https://www.gov.uk/government/publications/seoul-declaration-for-safe-innovative-and-inclusive-ai-ai-seoul-summit-2024/seoul-declaration-for-safe-innovative-and-inclusive-ai-by-participants-attending-the-leaders-session-ai-seoul-summit-21-may-2024. Last accessed June 18, 2026.
- UK Government. *Seoul Statement of Intent toward International Cooperation on AI Safety Science, AI Seoul Summit 2024*. 2024. https://www.gov.uk/government/publications/seoul-declaration-for-safe-innovative-and-inclusive-ai-ai-seoul-summit-2024/seoul-statement-of-intent-toward-international-cooperation-on-ai-safety-science-ai-seoul-summit-2024-annex. Last accessed June 18, 2026.
- UK Government. *The Bletchley Declaration by Countries Attending the AI Safety Summit, 1–2 November 2023*. 2023. https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023. Last accessed June 18, 2026.

- UNCTAD. *Digital Economy Report 2024: Shaping an Environmentally Sustainable and Inclusive Digital Future — Overview*. Geneva: United Nations, 2024. https://unctad.org/publication/digital-economy-report-2024. Last accessed June 18, 2026.
- UNCTAD. *Technology and Innovation Report 2025: Inclusive Artificial Intelligence for Development*. Geneva: United Nations, 2025. https://unctad.org/publication/technology-and-innovation-report-2025. Last accessed June 18, 2026.
- UNESCO. *Global AI Ethics and Governance Observatory*. https://www.unesco.org/ethics-ai/es. Last accessed June 18, 2026.
- UNESCO. *Recommendation on the Ethics of Artificial Intelligence*. 2021. https://www.unesco.org/en/artificial-intelligence/recommendation-ethics. Last accessed June 18, 2026.
- UNFCCC. *#AI4ClimateAction*. https://unfccc.int/ttclear/artificial_intelligence. Last accessed June 18, 2026.
- United Nations. *AI Panel and Dialogue*. https://www.un.org/global-digital-compact/en/ai. Last accessed June 18, 2026.
- United Nations. *Global Digital Compact*. 2024. https://www.un.org/pact-for-the-future/en/annex-i-global-digital-compact. Last accessed June 18, 2026.
- United Nations Environment Programme. *Artificial Intelligence (AI) End-to-End: The Environmental Impact of the Full AI Life Cycle Needs to Be Comprehensively Assessed*. Issues Note, September 21, 2024.
- United Nations General Assembly. *Resolution A/RES/78/265, Seizing the Opportunities of Safe, Secure and Trustworthy Artificial Intelligence Systems for Sustainable Development*. March 21, 2024. https://digitallibrary.un.org/record/4040897. Last accessed June 18, 2026.
- United Nations General Assembly. *Resolution A/RES/79/325, Terms of Reference and Modalities for the Establishment and Functioning of the Independent International Scientific Panel on Artificial Intelligence and the Global Dialogue on Artificial Intelligence Governance*. August 26, 2025. https://www.un.org/global-digital-compact/en/ai. Last accessed June 18, 2026.
- United Nations Secretary-General's High-level Advisory Body on Artificial Intelligence. *Governing AI for Humanity: Final Report*. September 2024. https://www.un.org/en/ai-advisory-body. Last accessed June 18, 2026.
- U.S. Congress. *H.R. 815 — 118th Congress: Making Emergency Supplemental Appropriations for the Fiscal Year Ending September 30, 2024, and for Other Purposes*. 2024. https://www.congress.gov/bill/118th-congress/house-bill/815. Last accessed June 18, 2026.
- U.S. Government Accountability Office. *Facial Recognition Technology: Federal Law Enforcement Agency Efforts Related to Civil Rights and Training*. GAO-24-107372. March 8, 2024. https://www.gao.gov/products/gao-24-107372. Last accessed June 18, 2026.
- Vaccaro, Michelle, Michael Caosun, Harang Ju, Sinan Aral, and Jared R. Curhan. *Advancing AI Negotiations: A Large-Scale Autonomous Negotiation Competition*. 2025. https://doi.org/10.2139/ssrn.5175035.
- Wachter, Sandra, Brent Mittelstadt, and Chris Russell. "Why Fairness Cannot Be Automated: Bridging the Gap between EU Non-Discrimination Law and AI." *Computer Law & Security Review* 41 (2021): 105567. https://doi.org/10.1016/j.clsr.2021.105567.

- Westminster Brief. *Foreign, Commonwealth and Development Office: Artificial Intelligence — Written Question 73676*. Answered September 16, 2025. https://westminsterbrief.co.uk/archive/pq/73676. Last accessed June 18, 2026.
- Yang, Fan, and Yuhan Yao. "A New Regulatory Framework for Algorithm-Powered Recommendation Services in China." *Nature Machine Intelligence* 4 (2022): 802–803. https://www.nature.com/articles/s42256-022-00546-9. Last accessed June 18, 2026.
- Ziegler, Matt, Sarah Lothian, Brian O'Neill, Richard Anderson, and Yoshitaka Ota. "AI Language Models Could Both Help and Harm Equity in Marine Policymaking." *npj Ocean Sustainability* 4, no. 32 (2025). https://doi.org/10.1038/s44183-025-00132-7.
- Bellingcat. *MH17: The Open Source Evidence: A Bellingcat Investigation*. 2015.
- Bellingcat. *MH17: The Open Source Investigation Three Years Later*. 2017.
- Block, Ludo. "The Long History of OSINT." *Journal of Intelligence History* 23, no. 2 (2024): 95–109. https://doi.org/10.1080/16161262.2023.2224091.
- Böhm, Isabelle, and Samuel Lolagar. "Open Source Intelligence: Introduction, Legal, and Ethical Considerations." *International Cybersecurity Law Review* 2 (2021): 317–337. https://doi.org/10.1365/s43439-021-00042-7.
- Cascavilla, Giuseppe, Filipe Beato, Andrea Burattin, Mauro Conti, and Luigi Vincenzo Mancini. "OSSINT — Open Source Social Network Intelligence: An Efficient and Effective Way to Uncover 'Private' Information in OSN Profiles." arXiv:1611.06737, 2016. https://arxiv.org/abs/1611.06737. Last accessed June 10, 2026.
- Dylan, Huw, and Thomas J. Maguire. "Secret Intelligence and Public Diplomacy in the Ukraine War." *Survival* 64, no. 4 (2022): 33–74. https://doi.org/10.1080/00396338.2022.2103257.
- Esri. "Ukraine: Open-Source Data Aided Response and Documents Damages and Atrocities." June 14, 2022. https://www.esri.com/about/newsroom/blog/ukraine-open-source-intelligence. Last accessed June 10, 2026.
- EUvsDisinfo. "About." https://euvsdisinfo.eu/about/. Last accessed June 10, 2026.
- EUvsDisinfo. "To Challenge Russia's Ongoing Disinformation Campaigns: Eight Years of EUvsDisinfo." https://euvsdisinfo.eu/to-challenge-russias-ongoing-disinformation-campaigns-eight-years-of-euvsdisinfo/. Last accessed June 10, 2026.
- European External Action Service. "Information Integrity and Countering Foreign Information Manipulation & Interference (FIMI)." https://www.eeas.europa.eu/eeas/information-integrity-and-countering-foreign-information-manipulation-interference-fimi_en. Last accessed June 10, 2026.
- Higgins, Eliot. "The Open Source Hunt for Syria's Favourite Sarin Bomb." *Bellingcat*, April 21, 2020. https://www.bellingcat.com/news/2020/04/21/the-open-source-hunt-for-syrias-favourite-sarin-bomb/. Last accessed June 10, 2026.
- Karalis, Magdalene. "Open-Source Intelligence in Ukraine: Asset or Liability?" *Chatham House*, December 16, 2022. https://www.chathamhouse.org/2022/12/open-source-intelligence-ukraine-asset-or-liability. Last accessed June 10, 2026.
- Leite, João A., Olesya Razuvayevskaya, Kalina Bontcheva, and Carolina Scarton. "EUvsDisinfo: A Dataset for Multilingual Detection of Pro-Kremlin Disinformation in News Articles." In *Proceedings of the 33rd ACM International Conference on Information and Knowledge Management (CIKM '24)*, 2024. https://doi.org/10.1145/3627673.3679167.

- Niu, Johannes, Mila Stillman, and Anna Kruspe. “OSINT — or BULLSHINT? Exploring Open-Source Intelligence Tweets about the Russo-Ukrainian War.” 2025.
- Office of the Director of National Intelligence and Central Intelligence Agency. “ODNI and CIA Release the Intelligence Community OSINT Strategy for 2024–2026.” March 8, 2024. https://www.odni.gov/index.php/newsroom/press-releases/press-releases-2024/3784-odni-and-cia-release-the-intelligence-community-osint-strategy-for-2024-2026. Last accessed June 10, 2026.
- Organisation for the Prohibition of Chemical Weapons. “Investigation and Identification Team.” https://www.opcw.org/iit. Last accessed June 10, 2026.
- Organisation for the Prohibition of Chemical Weapons. “OPCW Releases First Report by Investigation and Identification Team.” April 8, 2020. https://www.opcw.org/media-centre/news/2020/04/opcw-releases-first-report-investigation-and-identification-team. Last accessed June 10, 2026.
- Anthropic. “Get Started with Claude for Government.” *Claude Help Center*. April 9, 2026. https://support.claude.com/en/articles/14503590-get-started-with-claude-for-government.
- Anthropic. “Claude in Amazon Bedrock: Approved for Use in FedRAMP High and DoD IL4/5 Workloads.” June 11, 2025. https://www.anthropic.com/news/claude-in-amazon-bedrock-fedramp-high.
- Anthropic. “Offering Expanded Claude Access across All Three Branches of the U.S. Government.” August 12, 2025. https://www.anthropic.com/news/offering-expanded-claude-access-across-all-three-branches-of-government.
- Anthropic. “Anthropic’s Responsible Scaling Policy.” September 19, 2023. https://www.anthropic.com/news/anthropics-responsible-scaling-policy.
- Anthropic. “Model System Cards.” https://www.anthropic.com/system-cards.
- Anthropic. “Detecting and Countering Misuse of AI: August 2025.” August 27, 2025. https://www.anthropic.com/news/detecting-countering-misuse-aug-2025.
- Anthropic. “Constitutional AI: Harmlessness from AI Feedback.” December 15, 2022. https://www.anthropic.com/research/constitutional-ai-harmlessness-from-ai-feedback.
- Priyanshu, Aman, Yash Maurya, and Zuofei Hong. “AI Governance and Accountability: An Analysis of Anthropic’s Claude.” arXiv:2407.01557, 2024. https://arxiv.org/abs/2407.01557.
- Appel, Ruth, Peter McCrory, Alex Tamkin, Miles McCain, Tyler Neylon, and Michael Stern. *The Anthropic Economic Index Report: Uneven Geographic and Enterprise AI Adoption*. Anthropic, September 15, 2025.
- Alon-Barkat, Saar, and Madalina Busuioc. “Human–AI Interactions in Public Sector Decision Making: ‘Automation Bias’ and ‘Selective Adherence’ to Algorithmic Advice.” *Journal of Public Administration Research and Theory* 32, no. 1 (2022): 1–17. https://doi.org/10.1093/jopart/muac007.
- Government Digital Service. “Algorithmic Transparency Recording Standard — Guidance for Public Sector Bodies.” GOV.UK. Updated May 8, 2025. https://www.gov.uk/government/publications/guidance-for-organisations-using-the-algorithmic-transparency-recording-standard/algorithmic-transparency-recording-standard-guidance-for-public-sector-bodies.

- Höne, Katharina E., et al. *Mapping the Challenges and Opportunities of Artificial Intelligence for the Conduct of Diplomacy*. Geneva: DiploFoundation, 2019.
- Immigration, Refugees and Citizenship Canada. "Advanced Analytics and Automation for Processing Applications." Government of Canada. Last modified November 18, 2024. https://www.canada.ca/en/immigration-refugees-citizenship/corporate/transparency/digital-transparency-advanced-data-analytics/processing-applications.html.
- Immigration, Refugees and Citizenship Canada. "Automation to Help Process Privately Sponsored Refugee Applications." Government of Canada. December 16, 2022. https://www.canada.ca/en/immigration-refugees-citizenship/news/notices/automation-private-refugees.html.
- Immigration, Refugees and Citizenship Canada. *Artificial Intelligence Strategy*. Ottawa: Government of Canada, 2026.
- Immigration, Refugees and Citizenship Canada. "How We Use Advanced Analytics, Automation and Other Technologies: Immigration, Refugees and Citizenship Canada." Government of Canada. Last modified February 20, 2026. https://www.canada.ca/en/immigration-refugees-citizenship/corporate/transparency/digital-transparency-advanced-data-analytics.html.
- Electronic Immigration Network. "Home Office Suspends Use of Digital Streaming Tool for Visa Applications after Legal Action by JCWI and Foxglove." August 9, 2020. https://www.ein.org.uk/news/home-office-suspends-use-digital-streaming-tool-visa-applications-after-legal-action-jcwi-and.
- Alon-Barkat, Saar, and Madalina Busuioc. "Human–AI Interactions in Public Sector Decision Making: 'Automation Bias' and 'Selective Adherence' to Algorithmic Advice." *Journal of Public Administration Research and Theory* 33, no. 1 (2023): 153–169. https://doi.org/10.1093/jopart/muac007.
- Holsti, Ole R. "Cognitive Process Approaches to Decision-Making: Foreign Policy Actors Viewed Psychologically." *American Behavioral Scientist* 20, no. 1 (1976): 11–32. https://doi.org/10.1177/000276427602000103.
- Parasuraman, Raja, and Victor Riley. "Humans and Automation: Use, Misuse, Disuse, Abuse." *Human Factors* 39, no. 2 (1997): 230–253. https://doi.org/10.1518/001872097778543886.
- Redd, Steven B. "Key Concepts in the Poliheuristic Theory of Foreign Policy Decision Making: A Comparative Examination Using Systemist Theory." *Social Sciences* 12, no. 8 (2023): 446. https://doi.org/10.3390/socsci12080446
- European External Action Service. "Information Integrity and Countering Foreign Information Manipulation & Interference (FIMI)." March 17, 2026. https://www.eeas.europa.eu/eeas/tackling-disinformation-foreign-information-manipulation-interference_en?s=2803.
- Manor, Ilan, and Rhys Crilley. "The Mediatisation of MFAs: Diplomacy in the New Media Ecology." *The Hague Journal of Diplomacy* 15, no. 1 (2020): 66–92. https://doi.org/10.1163/1871191X-15101051.
- Manor, Ilan, and Zhao Alexandre Huang. "Digitalization of Public Diplomacy: Concepts, Trends, and Challenges." *Communication and the Public* 7, no. 4 (2022): 167–175. https://doi.org/10.1177/20570473221138401.

- Bjola, Corneliu. “Digital Diplomacy as World Disclosure: The Case of the COVID-19 Pandemic.” *Place Branding and Public Diplomacy* 18 (2022): 22–25. https://doi.org/10.1057/s41254-021-00242-2.
- Bjola, Corneliu, and Michaela Coplen. “Virtual Venues and International Negotiations: Lessons from the COVID-19 Pandemic.” *International Negotiation* 27 (2022): 1–25. https://doi.org/10.1163/15718069-bja10060.
- Bramsen, Isabel, and Anine Hagemann. “The Missing Sense of Peace: Diplomatic Approachment and Virtualization during the COVID-19 Lockdown.” *International Affairs* 97, no. 2 (2021): 539–560. https://doi.org/10.1093/ia/iiaa229.
- Manor, Ilan, and James Pamment. “At a Crossroads: Examining Covid-19’s Impact on Public and Digital Diplomacy.” *Place Branding and Public Diplomacy* 18 (2022): 1–3. https://doi.org/10.1057/s41254-021-00249-9.
- Manor, Ilan, and Moran Yarchi. “From the Global to the Local and Back Again: MFAs’ Digital Communications During COVID-19.” *International Journal of Communication* 17 (2023): 860–881.
- United Nations Security Council. “Highlights of Security Council Practice 2020.” https://main.un.org/securitycouncil/en/content/highlights-2020.
- European External Action Service. “Information Integrity and Countering Foreign Information Manipulation & Interference (FIMI).” March 17, 2026. https://www.eeas.europa.eu/eeas/tackling-disinformation-foreign-information-manipulation-interference_en?s=2803.
- International Committee of the Red Cross. “ICRC Position on Autonomous Weapon Systems.” May 12, 2021. https://www.icrc.org/en/document/icrc-position-autonomous-weapon-systems.
- International Committee of the Red Cross and Stockholm International Peace Research Institute. “Limits on Autonomy in Weapon Systems.” June 2, 2020. https://www.icrc.org/en/document/limits-autonomous-weapons.
- Lee, Joyce. “Sixty Countries Endorse ‘Blueprint’ for AI Use in Military; China Opts Out.” *Reuters*, September 10, 2024. https://www.reuters.com/technology/artificial-intelligence/south-korea-summit-announces-blueprint-using-ai-military-2024-09-10/.
- Mazzucchi, Nicolas. “AI-Based Technologies in Hybrid Conflict: The Future of Influence Operations.” Hybrid CoE Paper 14, June 2022. https://www.hybridcoe.fi/publications/hybrid-coe-paper-14-ai-based-technologies-in-hybrid-conflict-the-future-of-influence-operations/.
- NATO. “NATO Releases Revised AI Strategy.” July 10, 2024. https://www.nato.int/en/news-and-events/articles/news/2024/07/10/nato-releases-revised-ai-strategy.
- NATO. “Summary of NATO’s Revised Artificial Intelligence (AI) Strategy.” July 10, 2024. https://www.nato.int/en/about-us/official-texts-and-resources/official-texts/2024/07/10/summary-of-natos-revised-artificial-intelligence-ai-strategy.
- NATO. “NATO’s Approach to Counter Information Threats.” October 18, 2024. https://www.nato.int/en/about-us/official-texts-and-resources/official-texts/2024/10/18/natos-approach-to-counter-information-threats.
- Thiele, Ralph. “Artificial Intelligence – A Key Enabler of Hybrid Warfare.” Hybrid CoE Working Paper 6, March 2020. https://www.hybridcoe.fi/publications/hybrid-coe-working-paper-6-artificial-intelligence-a-key-enabler-of-hybrid-warfare/.

- U.S. Department of State. “Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy.” 2023. https://2021-2025.state.gov/bureau-of-arms-control-deterrence-and-stability/political-declaration-on-responsible-military-use-of-artificial-intelligence-and-autonomy/.
- Council of Europe. *Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law*. Strasbourg: Council of Europe, 2024.
- European Commission, High-Level Expert Group on Artificial Intelligence. “Ethics Guidelines for Trustworthy AI.” 2019. https://knowledge4policy.ec.europa.eu/publication/ethics-guidelines-trustworthy-ai_en.
- Floridi, Luciano, Josh Cowls, Monica Beltrametti, Raja Chatila, Patrice Chazerand, Virginia Dignum, Christoph Luetge, et al. “AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations.” *Minds and Machines* 28, no. 4 (2018): 689–707. https://doi.org/10.1007/s11023-018-9482-5.
- Hagendorff, Thilo. “The Ethics of AI Ethics: An Evaluation of Guidelines.” *Minds and Machines* 30 (2020): 99–120.
- Jobin, Anna, Marcello Ienca, and Effy Vayena. “Artificial Intelligence: The Global Landscape of Ethics Guidelines.” *Nature Machine Intelligence* 1 (2019): 389–399.
- Leslie, David, Christopher Burr, Mhairi Aitken, Michael Katell, Morgan Briggs, and Cami Rincon. *Human Rights, Democracy, and the Rule of Law Assurance Framework for AI Systems: A Proposal Prepared for the Council of Europe’s Ad Hoc Committee on Artificial Intelligence*. The Alan Turing Institute, 2021. https://doi.org/10.5281/zenodo.5981676.
- OECD. “AI Principles.” https://www.oecd.org/en/topics/ai-principles.html.
- UNESCO. “Recommendation on the Ethics of Artificial Intelligence.” 2021. https://www.unesco.org/en/artificial-intelligence/recommendation-ethics
- Bond, Shannon. “Fake Viral Images of an Explosion at the Pentagon Were Probably Created by AI.” *NPR*, May 22, 2023. https://www.npr.org/2023/05/22/1177590231/fake-viral-images-of-an-explosion-at-the-pentagon-were-probably-created-by-ai.
- European External Action Service. “Information Integrity and Countering Foreign Information Manipulation & Interference (FIMI).” March 17, 2026. https://www.eeas.europa.eu/eeas/tackling-disinformation-foreign-information-manipulation-interference_en?s=2803.
- European External Action Service. “Inside the Infrastructure of Foreign Information Manipulation and Interference (FIMI) Operations.” March 19, 2025. https://www.eeas.europa.eu/eeas/inside-infrastructure-foreign-information-manipulation-and-interference-fimi-operations_en.
- Hatmaker, Taylor. “Meta Takes Down Deepfake of Ukraine’s President Zelensky Surrendering.” *TechCrunch*, March 16, 2022. https://techcrunch.com/2022/03/16/facebook-zelensky-deepfake/.
- Meaker, Morgan. “Slovakia’s Election Deepfakes Show AI Is a Danger to Democracy.” *Wired*, October 3, 2023. https://www.wired.com/story/slovakias-election-deepfakes-show-ai-is-a-danger-to-democracy/.
- NATO. “NATO’s Approach to Counter Information Threats.” October 18, 2024. https://www.nato.int/en/about-us/official-texts-and-resources/official-texts/2024/10/18/natos-approach-to-counter-information-threats.

- OpenAI. “Disrupting Deceptive Uses of AI by Covert Influence Operations.” May 30, 2024. https://openai.com/index/disrupting-deceptive-uses-of-ai-by-covert-influence-operations/.
- Shoaib, Mohamed R., Zefan Wang, Milad Taleby Ahvanooey, and Jun Zhao. “Deepfakes, Misinformation, and Disinformation in the Era of Frontier AI, Generative AI, and Large AI Models.” ICCA, 2023.
- UNESCO. “Guidelines for the Governance of Digital Platforms.” November 11, 2023. https://www.unesco.org/en/articles/guidelines-governance-digital-platforms?hub=751.