Diplomacy in the Age of Artificial Intelligence

Corneliu Bjola | Head of the Oxford Digital Diplomacy Research Group, University of Oxford (#DigDiploROx) | @CBjola 🐦

Theme

The key question on the mind of policymakers now is whether Artificial Intelligence would be able to deliver on its promises instead of entering another season of scepticism and stagnation.

Summary

The quest for Artificial Intelligence (AI) has travelled through multiple "seasons of hope and despair" since the 1950s. The introduction of neural networks and deep learning in late 1990s has generated a new wave of interest in AI and growing optimism in the possibility of applying it to a wide range of activities, including diplomacy. The key question on the mind of policymakers now is whether AI would be able to deliver on its promises instead of entering another season of scepticism and stagnation. This paper evaluates the potential of IA to provide reliable assistance in areas of diplomatic interest such as in consular services, crisis management, public diplomacy and international negotiations, as well as the ratio between costs and contributions of AI applications to diplomatic work.

Analysis

The term "artificial intelligence" was first coined by an American computer scientist, John McCarthy in 1956, who defined AI as "the science and engineering of making intelligent machines, especially intelligent computer programs".[1] In basic terms, AI refers to the activity by which computers process large volumes of data using highly sophisticated algorithms to simulate human reasoning and/or behaviour.[2] Russell & Norvig use these

---

[1] John McCarthy (2011): 'What Is AI? / Basic Questions', author's website accessed 22 May 2019, http://jmc.stanford.edu/artificial-intelligence/what-is-ai/index.html.

[2] In simple terms, behaviour refers to the way in which people act to a situation in response to certain internal or external stimuli. The classical theory of human behaviour, the Belief-Desire-Intention (BDI) model argues that individual behaviour is best explained by the way in which agents develop intentions (desires that the agent has committed to pursue) out of a broader range of desires (states of affairs they would like to bring about), which in turn are derived from a set of beliefs (information the agent has about the world). The way in which intentions are formed remain a matter of dispute between different schools of thought, with a traditional view emphasizing the role of rational reasoning (the rational dimension) , while others stressing the importance of internal mental processes (the cognitive dimension), or the social context in which this occurs (the social dimension). See Michael E. Bratman (1999): *Intention, Plans, and Practical Reason*, Cambridge: Cambrigde University Press.

(cont.)

two dimensions (reasoning and behaviour) to group AI definitions according to the emphasis they place on thinking vs acting humanly. [3]

Another approach to defining AI is by zooming in on the two constitutive components of the concept. Nils J. Nilsson defines, for instance, artificial intelligence as the "activity devoted to making machines intelligent" while "intelligence is that quality that enables an entity to function appropriately and with foresight in its environment". [4] Echoing Nilsson's view, the European Commission's High-Level Group on AI provides a more comprehensive understanding of the term:

"Artificial intelligence (AI) systems are software (and possibly also hardware) systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve the given goal".[5]

While the concept of artificial intelligence continues to evolve, one could argue that the ambition to push forward the frontier of machine intelligence is the main anchor that holds the concept together. As the authors of the report on "Artificial Intelligence and Life in 2030" point out, we should not expect AI to "deliver" a life-changing product, but rather to continue to generate incremental improvements in its quest to achieve and possibly surpass human standards of reasoning and behaviour. In so doing, AI also sets in motion the so-called "AI effect": as AI brings a new technology into the common fold, people become accustomed to this technology, it stops being considered AI, and newer technology emerges.[6]

In the same way that cars differ in terms of their quality and performance, AI programs also significantly vary along a broad spectrum ranging from rudimentary to super-intelligent forms. In consular and diplomatic affairs, the left side of this spectrum is already visible. At the lower end of the complexity scale, chat-bots now assist with visa applications, legal aid for refugees, and consular registrations.[7] More sophisticated algorithms are being developed by MFAs to either advance the spread of positive narratives or inhibit online disinformation and propaganda. [8] However, all these applications, regardless of their degree of technical sophistication, fall in the category of

---

[3] Stuart Russell and Peter Norvig (2010): *Artificial Intelligence A Modern Approach,* Third Ed., Pearson, p. 2.

[4] Nils J Nilsson (2010): *The Quest for Artificial Intelligence: A History of Ideas and Achievements,* Cambridge: Cambridge University Press, p. 13.

[5] High-Level Expert Group on Artificial Intelligence (2019): 'A Definition of Artificial Intelligence: Main Capabilities and Scientific Disciplines', European Commission, p. 6, https://ec.europa.eu/digital-single-market/en/news/definition-artificial-intelligence-main-capabilities-and-scientific-disciplines.

[6] Peter Stone et al. (2016): 'Artificial Intelligence and Life in 2030', Report of the 2015-2016 Study Panel, Stanford University, Stanford, CA. (September), https://ai100.stanford.edu/sites/g/files/sbiybj9861/f/ai_100_report_0831fnl.pdf.

[7] Elena Cresci (2017): 'Chatbot That Overturned 160,000 Parking Fines Now Helping Refugees Claim Asylum', The Guardian, 6/III/2017.

[8] Simon Cocking (2016): 'Using Algorithms to Achieve Digital Diplomacy', *Irish Tech News*, 19/IX/2016.

'*narrow*' or '*weak*' *AI*, as they are programmed to perform a single task. They extract and process information from a specific dataset to provide guidance on legal matters and consular services. The 'narrow' designation for such AI applications comes from fact that they cannot perform tasks outside the information confines delineated by their dataset.

By contrast, *general* AI refers to machines that exhibit human abilities ranging from problem-solving and creativity to taking decisions under conditions of uncertainty and thinking abstractly. They are thus able to perform intellectual activities like a human being, without any external help. Most importantly, *strong* AI would require some form of self-awareness or consciousness in order to be able to fully operate. If so, strong AI may reach a point in which it will be able not only to mimic the human brain but to surpass the cognitive performance of humans in all domains of interest. This is what Nick Bostrom calls *superintelligence*, an AI system that can do all that a human intellect can do, but faster ('speed superintelligence'), or that it can aggregate a large number of smaller intelligences ('collective superintelligence') or that it is at least as fast as a human mind but vastly qualitatively smarter ('quality superintelligence').[9]

That being said, strong AI, let alone superintelligence, remain merely theoretical constructs at this time, as all applications developed thus far, including those that have attracted media attention such as Amazon's Alexa or Tesla's self-driving prototypes fall safely in the category of narrow AI. However, this may change soon, especially if quantum computing technology will make significant progress. Results from a large survey of machine learning researchers on their beliefs about progress in AI is relatively optimistic. Researchers predict AI will outperform humans in many activities in the next ten years, such as translating languages (by 2024), writing high-school essays (by 2026), driving a truck (by 2027), working in retail (by 2031), writing a bestselling book (by 2049), and even working as a surgeon (by 2053). Furthermore, they believe there is a 50% chance of AI outperforming humans in all tasks in 45 years and of automating all human jobs in 120 years.[10]

AI and Diplomacy

Riding the waves of growing interest about AI in IR and security studies,[11] the debate about the role of AI in diplomacy is also gaining momentum, although academic discussions are progressing rather slowly, without a clear analytical focus. As the authors

[9] Nick Bostrom (2014), *Superintelligence : Paths, Dangers, Strategies*, First Ed., Oxford: Oxford University Press, pp. 63–69 and 6-11.

[10] Katja Grace et al. (2018): 'Viewpoint: When Will AI Exceed Human Performance? Evidence from AI Experts', *Journal of Artificial Intelligence Research* 62, July 31, pp. 729–54.

[11] Daniel W Drezner (2019): 'Technological Change and International Relations', *International Relations*, 20/III/2019. ; Stoney Trent and Scott Lathrop (2019): 'A Primer on Artificial Intelligence for Military Leaders' *Small Wars Journal*, 2019, https://smallwarsjournal.com/jrnl/art/primer-artificial-intelligence-military-leaders; Edward Geist and Andrew Lohn (2018): *How Might Artificial Intelligence Affect the Risk of Nuclear War?* RAND Corporation.; Greg Allen and Taniel Chan (2017): 'Artificial Intelligence and National Security', Belfer Center, https://www.belfercenter.org/sites/default/files/files/publication/AI%20NatSec%20-%20final.pdf; Miles Brundage et al. (2018): 'The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation', 20/II/2018, https://arxiv.org/ftp/arxiv/papers/1802/1802.07228.pdf.

(cont.)

of a recent report on AI opportunities for the conduct of diplomacy point out, discussions about AI in the context of foreign policy and diplomacy often lack clarity in terminology. They suggest that a better understanding of the relationship between AI and diplomacy could come from building on the distinction between AI as a diplomatic topic, AI as a diplomatic tool, and AI as a factor that shapes the environment in which diplomacy is practised. As a topic for diplomacy, AI is relevant for a broader policy agenda ranging from economy, business, and security, all the way to democracy, human rights, and ethics. As a tool for diplomacy, AI looks at how it can support the functions of diplomacy and the day-to-day tasks of diplomats. As a factor that impacts the environment in which diplomacy is practised, AI could well turn out to be the defining technology of our time and as such it has the potential to reshape the foundation of the international order.[12]

Taking note of the fact that developments in AI are so dynamic and the implications so wide-ranging, another report prepared by a German think tank calls on Ministries of Foreign Affairs (MFAs) to immediately begin planning strategies that can respond effectively to the influence of AI in international affairs. Economic disruption, security & autonomous weapons, and democracy & ethics are the three areas they identify as priorities at the intersection of AI and foreign policy. Although they believe that transformational changes to diplomatic institutions will eventually be needed to meet the challenges ahead, they favour, in the short term, an incremental approach to AI that builds on the successes (and learns from the failures) of "cyber-foreign policy", which, in many countries, has been already internalised in the culture of the relevant institutions, including of the MFAs.[13] In the same vein, the authors of a report prepared for the Centre for a New American Security see great potential for AI in national security-related areas, including diplomacy. For example, AI can help improve communication between governments and foreign publics by lowering language barriers between countries, enhance the security of diplomatic missions via image recognition and information sorting technologies, and support international humanitarian operations by monitoring elections, assisting in peacekeeping operations, and ensuring that financial aid disbursements are not misused through anomaly detection.[14]

From an AI perspective, *consular services* could be a low-hanging fruit for AI integration in diplomacy as decisions are amenable to digitisation, the analytical contribution is reasonable relevant and the technology favours collaboration between users and the machine. Consular services rely on highly structured decisions, as they largely involve recurring and routinised operations based on clear and stable procedures, which do not need to be treated as new each time a decision has to be made (except for crisis situations, which are discussed further below). From a knowledge perspective, AI-assisted consular services may embody declarative (know-what) and procedural knowledge (know-how) to automate routinised operations and scaffold human cognition by reducing cognitive effort. This can be done by using data mining and data discovery

---

[12] DiploFoundation (2019): 'Mapping the Challenges and Opportunities of Artificial Intelligence for the Conduct of Diplomacy', Geneva, https://www.diplomacy.edu/sites/default/files/AI-diplo-report.pdf.

[13] Ben Scott, Stefan Heumann, and Philippe Lorenz (2018): 'Artificial Intelligence and Foreign Policy', Stiftung Neue Veranwortung, https://www.stiftung-nv.de/sites/default/files/ai_foreign_policy.pdf.

[14] Michael C Horowitz et al. (2018): 'Artificial Intelligence and International Security', Center for New American Security (CNAS).

techniques to organize the data and make it possible to identify patterns and relationships that would be difficult to observe otherwise (e.g., variation of demand for services by location, time, and audience profile).

Case study #1: AI as Digital Consul Assistant

The consulate of country X has been facing uneven demand for emergency passports, visa requests and business certifications in the past five years. The situation has led to a growing backlog, significant loss of public reputation and a tense relationship between the consulate and the MFA. An AI system trained with data from the past five years uses descriptive analytics to identify patterns in the applications and concludes that August, May and December are the most likely months to witness an increase of the demand in the three categories next year. AI predictions are confirmed for August and May but not for December. AI recalibrates its advice using updated data and the new predictions help consular officers manage requests more effectively. As the MFA confidence in the AI system grows, the digital assistant is then introduced to other consulates experiencing similar problems.

Digital platforms could also emerge as indispensable tools for *managing diplomatic crises* in the digital age and for good reasons. They can help embassies and MFAs make sense of the nature and gravity of the events in real-time, streamline the decision-making process, manage the public's expectations, and facilitate crisis termination. At the same time, they need to be used with great care as factual inaccuracies, coordination gaps, mismatched disclosure level, and poor symbolic signalling could easily derail digital efforts of crisis management.[15] AI systems could provide great assistance to diplomats in times of crisis by helping them make sense of what it is happening (descriptive analytics) and identify possible trends (predictive analytics). The main challenge for AI is the semi-structured nature of the decisions to be taken. While many MFAs have pre-designed plans to activate in case of a crisis, it is safe to assume that reality often defies the best crafted plans. Given the high level of uncertainty in which crisis decision-making operates and the inevitable scrutiny and demand of accountability to occur if something goes wrong, AI integration can work only if humans retain control over the process. As a recent SIPRI study pointed out, AI systems may fail spectacularly when confronted with tasks or environments that differ slightly to those they were trained for. Their algorithms are also opaque, which makes difficult for humans to explain how they work and whether they include bias that could lead to problematic –if not dangerous– behaviours.[16]

As data is turning into the "new oil", one would expect that the influence of digital technologies on *public diplomacy* to maximise interest in learning how to make oneself better heard, listened and followed by the relevant audiences. As the volume of data-driven interactions continue to grow at an exponential rate, one can make oneself heard

---

[15] Corneliu Bjola (2017): 'How Should Governments Respond to Disasters in the Digital Age?', The Royal United Services Institute (RUSI).

[16] Vincent Boulanin (2019): 'The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk', SIPRI.

(cont.)

by professionally learning how to separate 'signals' from the background 'noise' and by pro-actively adjusting her message to ensure maximal visibility in the online space, in real time. Making oneself listened would require, by extension, a better understanding of the cognitive frames and emotional undertones that enable audiences to meaningfully connect with a particular message. Making oneself followed would involve micro-level connections with the audience based on individual interests and preferences.[17]

At the operational level of the digital diplomat decisions are expected to take a structured form as the way to meaningfully communicate with the audience would rely on continuously tested principles of digital outreach with a likely focus on visual enhancement, emotional framing, and algorithmic-driven engagement. AI could assist these efforts by providing reliable diagnostics of the scope conditions for impact via network, cluster and semantic analyses. Prescriptive analytics could also offer insight into the comparative value-added of alternative approaches to digital engagement (e.g., which method proves more impactful in terms of making oneself heard, listened and followed). On the downside, the knowledge so generated would likely stimulate a competitive relationship between the AI system and digital diplomats as most of the work done by the later could be gradually automated. However, such a development might be welcome by budget-strapped MFAs and embassies seeking to maintain their influence and make the best of their limited resources by harnessing the power of technological innovation.

Case study #2: AI as Digital PD Assistant

The embassy of country X in Madrid would like to conduct a public diplomacy campaign in support of one of the following policy priorities: increasing the level of educational exchanges of Spanish students in the home country, showcasing the strength of the military relationship between country X and the Spain and boosting Spanish investments in the home country. As it has only £25,000 in the budget for the campaign, it needs to know which version can demonstrate better return on investment. Using social media data, an AI system will first seek to listen and determine the level of interest and reception (positive, negative, neutral) of the public in the three topics. The next step will be to use diagnostic analytics to explain the possible drivers of interest in each topic (message, format, influencers) and the likelihood of the public reacting to the embassy's campaign. The last step will be to run simulations to evaluate which campaign will be able to have the strongest impact given the way in which the public positions itself on each topic and the factors that may help increase or decrease public interest in them.

Given the growing technical complexity and resource-intensive nature of *international negotiations* it is hardly surprisingly that AI has already started to disrupt this field. The Cognitive Trade Advisor (CTA) developed by IBM aims to assist trade negotiators dealing with rules of origin (criteria used to identify the origin /nationality of a product) by

[17] Bjola, Cassidy, and Manor (2019): 'Public Diplomacy in the Digital Age', *The Hague Journal of Diplomacy* 14, April, p. 87.

(cont.)

answering queries related to existing trade agreements, custom duties corresponding to different categories of rules of origin, and even to the negotiating profiles of the party of interest.[18] CTA uses descriptive analytics to provide timely and reliable insight into technically complex issues that would otherwise require days or possibly weeks for an experienced team to sort out. It does not replace the negotiator in making decisions, nor does it conduct negotiations by itself, or at least not yet. It simply assists the negotiator in figuring out the best negotiating strategy by reducing critical information gaps, provided that the integrity of the AI system has not been compromised by hostile parties. The competitive advantage that such a system could offer negotiators cannot be ignored, although caveats remain for cases in which negotiations would involve semi-structured decisions such as climate negotiations or the Digital Geneva Convention to protect cyberspace. The problem for such cases lies with the lower degree of data veracity (confidence in the data) when dealing with matters that can easily become subject to interpretation and contestation, hence the need for stronger human expertise and judgement to assess the value of competing courses of action in line with the definition of national interests as agreed upon by foreign policy makers.

Conclusions

As Bostrom has shown, the quest for Artificial intelligence has travelled through multiple "seasons of hope and despair". The early attempts in the 1950s at the Dartmouth College sought to provide a proof of concept for AI by demonstrating that machines were able to perform complicated logical tests. Following a period of stagnation, another burst of innovative thinking took place in early 1970s, which showed that logical reasoning could be integrated with perception and used to control physical activity. However, difficulties in scaling up AI findings soon led to an "AI winter" of declining funding and increased scepticism. A new springtime arrived with the launch of the Fifth Generation Computer System Project by Japan in early 1980s, which led to the proliferation of expert systems as new tools of AI-supported decision-making. After another period of relative stagnation, the introduction of neural networks and deep learning in late 1990s has generated a new wave of interest in AI and growing optimism in the possibility of applying it to a wide range of activities, including diplomacy. The key question on the mind of policymakers now is whether AI would be able to deliver on its promises instead of entering another season of scepticism and stagnation. If AI would be able to demonstrate value in a consistent manner by providing reliable assistance in areas of diplomatic interest such as in consular services, crisis management, public diplomacy and international negotiations, as suggested above, then the future of AI in diplomacy should look bright. If, on the other hand, the ratio between costs and contributions of AI applications to diplomatic work would stay high, then the appetite for AI integration would likely decline.

---

[18] Maximiliano Ribeiro Aquino Santos (2018), 'Cognitive Trading Using Watson', *IBM blog*, 12/XII/2018, https://www.ibm.com/cloud/blog/mraquinobr-ibm-co.m