Corneliu Bjola

Artificial Intelligence and Diplomatic Crisis Management: Addressing the ‘Fog of War’ Problem

DigDiploROx Working Paper No 6 (July 2022)

Abstract:

Artificial Intelligence promises to revolutionize the way in which international crises are anticipated, understood, and managed. Specifically, AI systems could provide assistance to diplomats and decision-makers in times of crisis by helping them make sense of what is happening (descriptive analytics), chart possible trends or patterns of evolution of the crisis (predictive analytics) and assess the validity of the response strategies (prescriptive analytics). What is less known, however, is how these models could work in practice and the conditions that AI models need to meet in order to deliver results. Drawing on the case of the international crisis generated by the Russian war of aggression in Ukraine, the study advances a framework for applying AI to crisis management and discusses the opportunities and challenges of integrating AI in diplomatic decision making.

I. Introduction

The term “artificial intelligence” (AI) 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” (McCarthy, 2011). While the quest for Artificial Intelligence has travelled through multiple “seasons of hope and despair” in the past decades (Bostrom, 2014, pp. 6–11), there is a growing consensus that the current stage of AI development is qualitatively different. Owing to the fast-paced development of complex machine and deep learning algorithms, AI applications have now reached the point at which they can learn on their own using statistical models and neural-like networks without being explicitly programmed (Collins, 2021). AI disruption could therefore have a strong impact on crisis management, especially since digital platforms have emerged as critical tools for assisting decision-makers manage crises in the digital age. They already help embassies and MFAs make sense of the nature and gravity of the events in real-time, streamline the decision-making process, manage public expectations, and facilitate crisis termination (Bjola & Coplen, 2022). At the same time, they need to be used with great care as factual inaccuracies, coordination gaps, mismatched disclosure levels, and poor signalling practices could easily derail digital efforts of crisis management (Bjola, 2017).

As discussed in more detail elsewhere,1 AI systems could aid diplomats in times of crisis by helping them make sense of what is happening (descriptive analytics), identify possible trajectories of the evolution of the crisis (predictive analytics), and prescribe possible response strategies (prescriptive analytics). AI has been already hailed as a possible solution for forecasting geopolitical events (Morstatter et al., 2019), predicting outbursts of violence and probing their causes (Guo et al., 2018) or for improving strategic intelligence assessments regarding the use of coercive and non-coercive tactics in complex social circumstances (Frank, 2017). The main challenge for AI is the semi-structured nature of the decisions to be taken. 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 only work if humans retain some level of control over the process. As a SIPRI study points out, AI systems may spectacularly fail when confronted with tasks or environments that differ

---

1 This section draws on a more comprehensive examination of AI applications to diplomacy that can be found in (Bjola, 2020, p. 28).
slightly to those they were trained for. AI algorithms are also opaque, which often makes difficult for humans to explain how they work and whether they mask inbuilt biases that could lead to problematic—if not dangerous—behaviours (Boulain, 2019).

Building on this literature, this paper seeks to advance the debate about the opportunities that AI can generate for diplomatic decision making in times of crisis by theorising about the challenges that diplomats face in times of crisis and developing a prototype model for understanding how unfolding crises can be monitored, analysed, and responded in real time. To this end, the paper will first explain the uncertainty challenge facing decision makers in times of crisis, then introduce the AI prototype model that may help address the said challenge and conclude with a short discussion of the advantages and limitations of the model.

II. The “Fog of War” Problem

How do Ministries of Foreign Affairs (MFA) handle uncertainty in times of crisis? We know from the classical literature on crisis management (Allison, 1971; Janis, 1972; Jervis, 2017) that uncertainty is a critical challenge that decision makers experience in times of crisis. The issue is, of course, hardly new. In his *magnus opus* "On War" (1984), Clausewitz actually proposed two terms for describing the problem of uncertainty: the “fog of war” and “friction.” The first term, *the fog of war*, refers to the diminished level of accuracy and reliability of the information exchanged in times of war and the difficulties encountered by political and military leaders when seeking to compensate for this limitation and maximize the value of the data used for taking decisions. According to Clausewitz, "three quarters of the factors on which action in war is based are wrapped in a fog of greater or lesser uncertainty” (p. 101). For example, the series of incidents taking place in Transnistria, the breakaway territory in Moldova bordering on Ukraine and controlled by Russia, have raised fears that the Ukraine conflict may be spreading (Peter, 2022). The lack of accurate information about the intention and capability of the parties involved is a good illustration of the “fog of war” problem.

*Friction*, on the other hand, refers to the interaction of chance and action and can be caused by many factors, including enemy forces, friendly actions, or the environment. For Clausewitz, friction differentiates "real war from war on paper," those surprising things that happen during wartime that make even the “simplest thing difficult.”(p. 119). One may think that the surprising impact of new weapons (e.g., drones), the arrival of a natural disaster or pandemic, or unforeseen political events may fall in this category. The two terms, the fog of war and friction, offer us different perspectives on how to reflect on the problem of uncertainty in times of crisis and encourages us to pay closer attention to the distinction between what is relatively controllable (given the quality and amount of available information) and what is less manageable (chance or unexpected events, which are harder to predict). In Clausewitz terms we might be able to handle the fog of war by making it less “foggy”, but it would be difficult if not impossible to avoid friction as the future is hardly predictable regardless of how much high-quality information we may manage to acquire.

It is important to note, at this point, that the goal of this paper is not to examine how military commanders or MoDs handle uncertainty, but how diplomats and MFAs cope with it. The distinction is important. MoDs are primarily interested in winning military campaigns, and they use lethal forces to achieve that. The military needs accurate and reliable information because
it seeks to maximise the level of damage and casualty that they can inflict upon the enemy, and to minimize both onto themselves. MFAs, on the other hand, are interested in building coalitions to minimize the overall costs of the conflict (economic, military, political, reputational) and they use diplomatic instruments to achieve that (bilateral and multilateral engagement, strategic communication, international law). Different goals, different means, and by extension, different approaches to managing the “fog of war”.

That being said, how does the issue of the "fog of war" apply to international crises from a diplomatic perspective? The answer revolves around the idea of signals that MFAs send and receive from one another. More specifically, MFAs are interested in understanding how other governments position themselves on key aspects informing and shaping the collective management of the crisis (e.g., international sanctions, military assistance, UN resolutions, peace negotiations), how robust their commitment to these positions is (any weak links?), and under what conditions their positions are likely to change. To this end, MFAs rely on their extensive networks of embassies and specialized departments to gather and analyse relevant information to assist them in their decision making. The capacity to collect and read signals is definitely important, but MFAs’ ability to reduce the uncertainty induced by the “fog of war” also depends on how well the signals are communicated by other parties and how free from interference they circulate through the network of formal and informal channels of communication that parties used in times of crisis.

According to signalling theory (ST), some of the signals that parties send to each other in times of conflict are easier to decipher. To project their resolve, intentions, and/or capabilities, parties may try to indicate that they are prepared to incur higher costs (ex-ante and ex-post) in order to reach their objectives (Gartzke et al., 2017). For example, as its military aggression against Ukraine has started to falter, Russia has insisted that it will be able to achieve its political objectives regardless of how high the military and economic costs the war may prove to be. At the same time, one should also bear in mind that parties do not always have a clear and consistent idea of the signals they would like to broadcast, and these signals may constantly evolve in line with the trajectory of the crisis (see, for example, Germany’s conflicting positions about supporting delivery of weapons to Ukraine). Parties may also try to send signals not to demonstrate resolve but to confuse others about their intentions (see, for instance, Russian officials’ statements before the start of the war in Ukraine falsely claiming that no invasion was planned). In addition, the receiver may have reason to doubt the signal received or may not have the capacity to read it properly. In short, the “fog of war” is a dynamic process influenced by a combination of factors pertaining to the clarity of the signals sent, the credibility of the message and the messenger, the suitability of the communication channels used for the exchange, as well the ability of the receiver to decipher, interpret and react to the message received.

III. AI Modelling and Crisis Management

The argument advanced in this paper is that AI can help MFA cope with the "fog of war" by adjusting the impact of the factors that contribute to reducing vs increasing uncertainty in times
of crisis. Drawing on the typology used in data analytics to distinguish between descriptive, predictive and prescriptive models (Lepenioti et al., 2020), the paper advances a conceptual model for integrating AI into crisis decision-making based on three components as shown in Fig 1:

**Fig 1: Data Analytics: The Descriptive - Predictive - Prescriptive Model**

![Diagram showing the relationship between Descriptive, Predictive, and Prescriptive analytics]

The first component, **descriptive analytics**, involves contextual mapping and the extraction of relevant information that can provide an accurate picture of the nature of the problem. The key question this component seeks to answer is what is happening? In the context of a crisis, MFAs are interested in detecting patterns that may indicate a potential challenge or opportunity for managing the crisis. Recalling the case of the war in Ukraine, questions that MFAs may ask could refer to how the positions of the parties involved in the conflict and of their key supporters evolve in real time? What aspects do they prioritize? How well these positions align or diverge from each other? The second component, **predictive analytics**, is about forecasting possible courses of action and their possible implications by testing and validating certain assumptions about the nature and the cause of the problem (what will happen?). How the positions of the parties involved in the crisis may evolve in view of the changing circumstances? Will country X likely support the EU ban on Russian oil and gas? If so, under what conditions? The last component, **prescriptive analytics**, encourages decision makers to integrate the information gathered in the previous steps and use the result to determine the best course of action to be taken (what should be done?). What implications the course A vs course B of action will have for the MFA’s relations with others? Shall country X take the lead of international efforts aiming to lift the Russian blockade of Ukraine grain in the Black Sea? How may such a decision affect the diplomatic unity among EU or NATO members?

All three components can be processed, of course, with no AI assistance. In fact, MFAs should be able to conduct such analysis in times of crises, and they have doing so on a regular basis, using in-house and commissioned expertise. What AI can presumably add to this is real-time insight and a more accurate evaluation of the substance and credibility of the signals that parties exchange with each other. AI may not be able to completely dissolve “the fog of war”, but they may be able to provide sufficient or actionable confidence in the value of the information used for taking decisions in times of crisis. To do this, an AI model need to take into consideration the factors that can blur crisis signalling and reduce the level of uncertainty that they induce as
much as possible. As indicated in Fig 2, AI modelling starts with a process of aggregation of the data gathered by the MFA and its network of embassies from static (e.g., macro-economic indicators, socio-demographic data) and dynamic sources (e.g., social media feeds, official statements, newspapers stories). The dataset so generated would then be split into two subsets (usually 70% training, 30% testing) to be used for training and testing models created with AI algorithms. After running and fine-tuning competing models of topic, social network and engagement analysis, an optimal AI model would be then selected to offer insight to assist decision-making. The model should be able to indicate the set of themes, the network of influencers, and the format of engagement that could most effectively capture the signals communicated by the relevant actors in the conflict. The framework may also include an assessment of the feasibility of integrating other AI models (marked with * in the diagram) from partnering countries or international organisations in an effort to further reduce the uncertainty induced by the “fog of War”. The insight gained from data analysis could be then converted into a plan of action to inform official reactions and policy responses to the crisis. The process continues with another round of data collection that feeds directly into data analysis, allowing decision makers to trace and react to novel developments in real time during the crisis.

**Fig 2: AI-Based Crisis Management Model**

While the model presented in Fig 2 equally applies to any of the three analytical components discussed above, it should be noted that the complexity of AI modelling and by extension its analytical value for crisis decision-making considerably varies between the descriptive,

---

2 For a more detailed discussion of the conditions for designing AI models for diplomacy, see (Bjola, 2020, pp. 34–41).
predictive, and prescriptive formats. The main difference lies with the quality of the data required to power the machine learning (ML) techniques of each component as well as with the degree of sophistication of these techniques. The data necessary for tracing and analysing the evolution of a crisis is more readily available and can be processed using relatively conventional ML algorithms. This is so because descriptive analytics rely on decisions that have been already taken and on actions that have been already implemented. The situation arguably becomes more complicated once the AI system is asked to predict possible courses of action and to assess the viability of the response strategies as the information required to generate such responses is based on decisions not yet taken and actions that are yet to be implemented. It is therefore important that discussions about the application of AI to crisis management pay close attention to the descriptive, predictive, and prescriptive sequence, so that the knowledge developed in each case can properly inform the development of AI solutions in the other cases. For this reason, the following section will focus on understanding the conditions of application of AI to the first component (descriptive analytics), with the hope that the lessons learned from this stage could be subsequently applied and expanded for developing AI solutions to support predictive and prescriptive analyses of crisis management as well.

IV. AI Modelling and the War in Ukraine

The Russian invasion of Ukraine represents the case study used in this paper for designing and testing an AI prototype to assist decision making in times of crisis. The objective of the prototype is to provide a preliminary evaluation of the capacity of AI systems to reduce the risk of the “fog of war” that diplomats may experience in times of crisis by improving the accuracy of the signals they receive from other parties involved in the conflict, as well as the time of reaction to these signals. To this end, the analysis will draw on a dataset containing Tweets extracted in real-time from 28 accounts representing the ministries of foreign affairs of the belligerent parties (Ukraine, Russia), as well as the countries closest to the conflict (the three Baltic states, Poland, Moldova, Romania, Bulgaria, Turkey, Finland). The dataset also includes Tweets posted by other international actors with a sensible stake in the conflict (United States, UK, China, Germany, France, Italy, Sweden, Norway, Canada, Japan, Australia, Taiwan, Korea, Israel as well as the EU, NATO and the UN). Tweets have been extracted from the Twitter API on July 14, 2022, and then processed in real time on the basis of an AI model developed by the author using the Orange data visualization, machine learning and data mining toolkit developed by the University of Ljubljana.

For the first, descriptive stage of the analysis, the AI model combines the following techniques: data extraction from Twitter API (max 75 tweets per account) followed by pre-process textual tokenization, filtering and normalization; topic modelling of underlying themes in the dataset based on clusters of words found in each tweet and their respective frequency; multidimensional scaling (MDS) of the distance between the positions of each tweet relative to the dominant topics; network analysis of the frequency words in tweets; and multi-class sentiment analysis of the set of emotions framing each tweet. The data extraction phase has generated a corpus of 3985 tweets in total, which has been subsequently reduced to 729 tweets after the removal of messages not mentioning Ukraine.
The frequency distribution of the 729 tweets by author is presented in Fig 3, which unsurprisingly shows the Ukrainian Foreign Minister, Dmytro Kuleba, as the most active communicator during this period (17.15%). He is followed by the President of the European Council, Charles Michel (9.47%), the NATO Secretary General, Jens Stoltenberg (9.33%), the Lithuanian Minister of Foreign Affairs, Gabrielius Landsbergis (5.35%), and the UK Foreign Office (4.53%). Interestingly, the US State Secretary, Antony Blinken, has made fewer interventions on Twitter during this period (3.16%), probably because of the overlapping visit of President Biden in the Middle East, slightly below the number of messages posted by the President of the European Commission, Ursula von der Leyen (3.98%), and that of the German Minister of Foreign Affairs, Annalena Baerbock (3.70%). The Latent Semantic Indexing algorithm used for topic modelling has revealed five coherent themes in the data corpus. The dominant topic is defined by keywords such as “Ukraine, support, we, Russia, war, EU, discuss”, suggesting the presence of a pro-active, solidarity-oriented narrative of international actors with Ukraine.
As Fig 4 shows, the emerging narrative is reasonably robust (the yellowish the colour, the more coherent the narrative) with Kuleba and Landsbergis promoting it most actively, followed by Charles Michel and Jens Stoltenberg. At the same time, the graph suggests that messages are relatively spread out with no clear “attractors” to facilitate their coagulation. This implies that the emerging narrative is likely to remain in a rather fluid and unstable configuration. This observation is confirmed by the graph in Fig 5, which offers the results of a network analysis of the most connected words in the dominant topic. The strongest and shorter ties in the narratives are between nodes labelled “support”, “Ukraine”, “EU”, and “aggression”. From a communicational perspective, the presence of these ties suggests the EU and international support for Ukraine remains strong after five months of war, but in rather generic, broad terms. Interestingly, the tie between the nodes of “food” and “security” appears to be strengthening, but it seems to remain outside the core area of discussion, at least for the time being.
That being said, statistical analysis of the list of words with lower p-values reveals a more nuanced picture of the positions of the various actors after five months of war. Lower p-values (<0.01) indicate a higher likelihood that the words in the list are significant for the selected authors. As Table 1 shows, the EU signals, for instance, through the messages of its two Presidents, Ursula von der Leyen and Charles Michel, that is committed to supporting the long-term reconstruction of Ukraine, but also to demonstrating solidarity with other countries that might be threatened by Russia, such as Moldova. The NATO Secretary General, Jens Stoltenberg, as well as the US State Secretary, Antony Blinken, insist that the Russian aggression should lead to stronger efforts of military preparation, collective deterrence, and coordinated support for Ukraine. Finally, the UN Secretary General, António Guterres, calls attention to the severe humanitarian costs of the war, not only for Ukraine and the region, but for the international community at large.
Table 1: List of words highly relevant for individual messages (p-value in brackets)

<table>
<thead>
<tr>
<th>Ursula von der Leyen</th>
<th>Charles Michel</th>
<th>Jens Stoltenberg</th>
<th>Antony Blinken</th>
<th>António Guterres</th>
</tr>
</thead>
<tbody>
<tr>
<td>Long (1.7e-08)</td>
<td>Solidarity (6.8e-04)</td>
<td>Support (2.4e-12)</td>
<td>Ukraine (5.4e-06)</td>
<td>Energy (5.0e-04)</td>
</tr>
<tr>
<td>Reconstruct (4.4e-06)</td>
<td>Moldova (1.9e-03)</td>
<td>Defence (1.3e-11)</td>
<td>Insecurity (1.1e-05)</td>
<td>Immediate (3.7e-03)</td>
</tr>
<tr>
<td>Ukraine (1.4e-05)</td>
<td>moment (4.3e-03)</td>
<td>Allies (1.9e-11)</td>
<td>Coordinate (6.5e-05)</td>
<td>end (5.9e-03)</td>
</tr>
<tr>
<td>reform (1.7e-05)</td>
<td>sanctions (5.8e-03)</td>
<td>prepare (4.9e-09)</td>
<td>Russia (3.8e-04)</td>
<td>action (6.7e-03)</td>
</tr>
<tr>
<td>invest (1.7e-05)</td>
<td>EU (6.2e-03)</td>
<td>presid (9.9e-09)</td>
<td>brutal (8.3e-04)</td>
<td>besiege (8.5e-03)</td>
</tr>
<tr>
<td>take (1.8e-04)</td>
<td>Marshal (7.1e-03)</td>
<td>meet (6.9e-08)</td>
<td>g20 (1.7e-03)</td>
<td>catastrophe (8.5e-03)</td>
</tr>
<tr>
<td>lead (2.2e-04)</td>
<td>Now (9.7e-03)</td>
<td>contribute (2.2e-07)</td>
<td>arm (2.6e-03)</td>
<td>delusion (8.5e-03)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>deter (2.2e-07)</td>
<td>American (3.5e-03)</td>
<td>fossil (8.5e-03)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>leader (2.2e-07)</td>
<td></td>
<td>fuel (8.5e-03)</td>
</tr>
</tbody>
</table>

Finally, sentiment analysis helps us capture the emotional framing of the messages posted on social media by the main actors in our sample. As graph in Fig 6 indicates, participants experience a range of emotions when communicating about Ukraine. Sadness (depression) and anger are clearly the dominant emotions in the dataset. This is actually to be expected given the context of the war and the constant flow of news regarding the atrocities committed by the Russian army, the loss of civilian lives, and the destruction of Ukrainian cities. These sentiments are likely to continue to dominate the way in which messages related to Ukraine will be exchanged online by MFAs and diplomats. At the same time, it is important to observe how the balance between “fatigue” and “vigour” may evolve over time. Traces of “fatigue” currently appear to increase in intensity, but “vigour” is also present, especially in messages posted by the representatives of Estonia, Slovakia, NATO, the EU, and Ukraine.

Fig 6: Sentiment analysis by authors
V. Conclusion

The main objective of the paper has been to explore, from a diplomatic perspective, the added value and feasibility of using AI solutions for managing international crises. It has been thus argued that AI can help MFAs cope with the "fog of war" by adjusting the impact of the factors that contribute to reducing vs increasing uncertainty in times of crisis. Due to space and technical constraints, the paper has only focused on exploring the contribution that AI can make to decision crisis management from the angle of descriptive analytics. To this end, the paper has sought to identify the relevant factors and patterns that can help diplomats make sense of unfolding crises in real time. An AI prototype has been built for this purpose using as a case study the international crisis generated by the Russian invasion of Ukraine. The model allows diplomats to trace in real time what international actors are most active and confident in terms of signalling, how these signals coalesce or diverge from each other, and to what extent these signals are consistent and predictable. At the same time, the model draws on a specific type of data (tweets) and uses conventional techniques, which are applied to a small data set. The performance of the AI model needs therefore to be improved by using a wider range of data (social media, newspapers stories, official statements) and more robust ML techniques. To increase confidence in the model and facilitate adoption, the results of the AI prototype also need to be compared, in terms of accuracy and speed, with those obtained from experiments conducted with a group of experts seeking to address and solve the same type of tasks. Despite the inherent constraints of the study, the expectation is that the lessons learned from this study could be subsequently applied and expanded for developing AI solutions to support predictive and prescriptive analyses of crisis management as well.
Bibliography:


