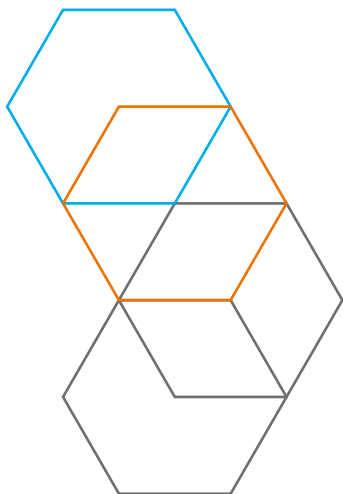
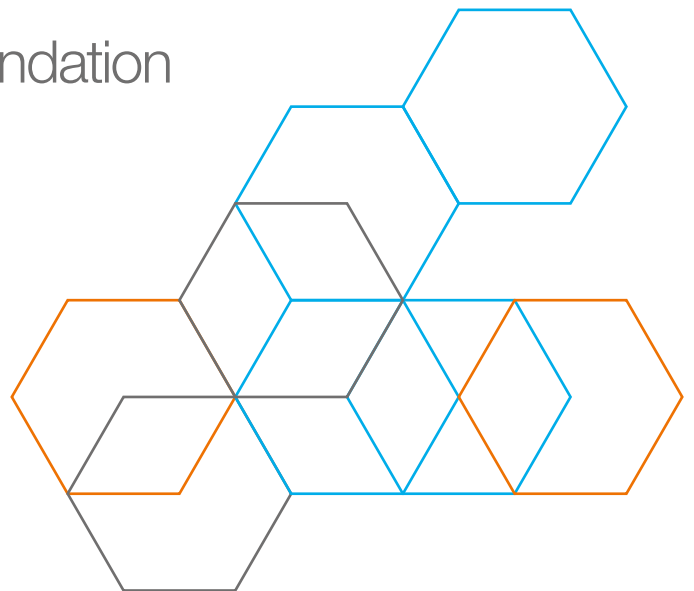


Mediation and artificial intelligence: Notes on the future of international conflict resolution



DiploFoundation



Layout and design: Viktor Mijatović, Aleksandar Nedeljkov
Copy editing: Mary Murphy
Author and researcher: Katharina E. Höne



Except where otherwise noted, this work is licensed under
<http://creativecommons.org/licenses/by-nc-nd/3.0/>

***di*PLO**
www.diplomacy.edu



DiploFoundation, 7bis Avenue de la Paix, 1201 Geneva, Switzerland

Publication date: November 2019

For further information, contact DiploFoundation at AI@diplomacy.edu

For more information on Diplo's AI Lab, go to <https://www.diplomacy.edu/AI>

To download the electronic copy of this report, click on https://www.diplomacy.edu/sites/default/files/Mediation_and_AI.pdf



Table of Contents



Acknowledgements.....	5
Executive summary.....	6
Introduction.....	7
Mediation and AI: The broad picture	8
Understanding AI and related concepts	8
A three-part typology to think through the relationship of AI and mediation.....	9
AI literacy for mediators.....	10
AI as a tool for mediators: Mapping the field.....	11
Potential applications and uses of AI	11
Key considerations and precautions when using AI tools	14
Implications for trust and buy-in	16
Conclusions and further recommendations	18



Acknowledgements



In March 2018, a group of organisations came together to form the #CyberMediation initiative to inform mediation practitioners about the impact of new information and communication technologies on mediation, including their benefits, challenges, and risks in relation to peacemaking; to develop synergies between the mediation community and the tech sector; and to identify areas of particular relevance and co-operation. Over the past 18 months, a number of meetings, both in situ and online, furthered this conversation among members of the #CyberMediation initiative. Ideas and insights generated through this exchange helped to inspire this report. Particular thanks goes to the Mediation Support Unit (United Nations Department of Political and Peacebuilding Affairs), the Centre for Humanitarian Dialogue, and swisspeace.¹

This report builds on previous work by DiploFoundation on the role of big data and artificial intelligence (AI) in

diplomacy.^{2,3} The work of colleagues and the many conversations about these topics have been important in pushing the ideas in this report forward. Members of Diplo's Data Team and Diplo's AI Lab⁴ have been indispensable intellectual sparring partners.

A special thanks goes to two interviewees for this report, Miloš Strugar (executive director of Conflux center, UN Senior Mediation Advisor) and Diego Osorio (expert in mediation, negotiations, and peace processes) for their time and support at various stages of this project.⁵

For their support in finalising this report, special thanks go to colleagues at Diplo – in particular Viktor Mijatović, Mina Mudrić, and Aleksandar Nedeljkov – and to Mary Murphy.

This report is published in the framework of a project supported by the Federal Department of Foreign Affairs of Switzerland. We gratefully acknowledge this support.



Executive summary



This report provides an overview of artificial intelligence (AI) in the context of mediation. Over the last few years AI has emerged as a hot topic with regard to its impact on our political, social, and economic lives. The impact of AI on international and diplomatic relations has also been widely acknowledged and discussed. This report focuses more specifically on the practice of mediation. It aims to inform mediation practitioners about the impact of AI on mediation, including its benefits, its challenges, and its risks in relation to peacemaking. It also hints at synergies between the mediation and technology communities and possible avenues of co-operation.

Broadly speaking, the report provides (a) a mapping of the relationship between AI and mediation, (b) examples of AI tools for mediation, (c) thoughts on key considerations and precautions when using AI tools, and (d) a discussion on the potential impact of AI tools on trust in the context of mediation.

In **mapping the relationship between AI and mediation**, the report offers a **three-part typology**: (a) AI as a topic for mediation, (b) AI as a tool for mediation, and (c) AI as something that impacts the environment in which mediation is practised.

Based on this, the report argues that even if mediators do not directly engage with AI tools, they need to have a **basic AI literacy**. AI has started to shape the environment of conflict resolution and might, in the not so distant future, play a role in the conflict itself.

The report identifies **three areas in which AI applications can usefully support the work of mediators and their teams**: (a) supporting knowledge management

and background research, (b) generating a good understanding of the conflict and the parties to the conflict, and (c) creating greater inclusivity of mediation processes.

The report also highlights **three areas of consideration and precaution** when using AI as a tool for mediation: (a) potential biases related to data sets and algorithms that need to be acknowledged, (b) the need for strong provisions regarding data privacy and security, and (c) the need for collaboration with the private sector on some of these tools.

Trust can play an important role on the road towards successful mediation. The report identifies **three ways in which the use of AI tools might impact trust in the context of mediation**: (a) if AI tools are used by mediators and their teams, it is important to undertake steps to build trust in these tools; (b) trust in AI tools or the lack thereof might also impact the trust enjoyed by mediators and their teams and vice versa; (c) AI tools might also be used to build a 'working trust' between the parties in conflict, in particular when it comes to monitoring and implementation.

Lastly, the report offers **four additional recommendations for taking work in the context of AI and mediation further**: (a) engaging in capacity building for AI in mediation that goes beyond AI literacy; (b) considering a re-shaping of mediation teams to include relevant topical and technical expertise in the area of AI; (c) developing best practices and guidelines for working with the private sector; and (d) engaging in foresight scenario planning and proof of concept regarding the use of existing and future AI tools for mediation



Introduction



This report provides an overview of AI in the context of mediation. Over the last years AI has emerged as a hot topic with regard to its impact on our political, social, and economic lives. The impact of AI on international and diplomatic relations has also been widely acknowledged and discussed. However, at this point in the discussion, we need to carefully navigate between hype around AI and a dystopian vision of the technology. It is equally important that the discussion on AI moves from general observations to specific examples and applications. This report takes these discussions forward in a productive way.

More specifically, this report focuses on the practice of mediation, understood as ‘the active search for a negotiated settlement to an international or intrastate conflict by an impartial third party’.⁶ It informs mediation practitioners about the impact of AI on mediation, including its benefits, its challenges, and its risks in relation to peacemaking. In doing so, this report builds on two previous DiploFoundation studies, *Data Diplomacy – Updating Diplomacy to the Big Data Era* and *Mapping the Challenges and Opportunities of Artificial Intelligence for the Conduct of Diplomacy*.^{7,8}

It is important to emphasise from the outset that mediation is profoundly interpersonal and highly dependent on the skills of the mediators and their teams. It hardly needs stressing that AI tools are not here to replace mediators or automate broad aspects of their work.

Having said this, it is still worth exploring new tools for mediation, if they offer the chance to support the journey towards conflict resolution. More broadly speaking, AI will impact the work of mediators by changing the environment in which conflict resolution is practised and regardless of whether mediators actively use AI tools for their work, this should be enough reason to learn more about the impact of AI on mediation.

The first part of this report provides a broad overview of AI and mediation by (a) introducing key concepts related to AI, (b) mapping the relationship between AI and mediation, and (c) stressing the importance of AI literacy for mediators. The second part focuses more specifically on AI tools for mediation. It sheds light on potential applications of AI, adds key considerations and precautions regarding the use of AI tools, and explores trust in the context of mediation and AI. By way of conclusion, the report adds four additional recommendations regarding the next steps in using AI tools for international conflict resolution.

This report presents the first study dedicated exclusively to AI and mediation. That is why we undertake a mapping exercise before zooming in on a discussion of AI tools and their implications. Many of the topics that this report touches on deserve closer attention and more detailed research, which hopefully will follow over the next few years.



Mediation and AI: The broad picture



In this section, we set the scene for understanding the relationship between AI and mediation by (a) providing an overview of AI and related concepts such as machine learning and big data, (b) offering a three-part typology to think through the relationship between AI and mediation, and (c) defining and discussing AI literacy for non-technical experts in the context of mediation.

To fully appreciate the impact of AI on mediation, it is important to start from a broader perspective. While this report focuses on AI as a tool for mediation in the following section, a broad appreciation of the relationship between AI and mediation furthers our understanding of the topic and helps to generate ideas for future research and practice in this area.

Understanding AI and related concepts

To understand and map the potential of AI for peacemaking, we need to start from a basic understanding of AI. To begin with, we can **define AI** as '[t]he theory and development of computer systems able to perform tasks normally requiring human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages'.⁹ **Areas of AI** include automated reasoning, robotics, computer vision, and natural language processing (NLP).¹⁰

The most prominent examples of AI discussed today have been achieved by machine learning. Machine learning is behind AlphaGo's ability to beat human board-games champions, Amazon's ability to make recommendations, PayPal's ability to recognise fraudulent activities, and Facebook's ability to translate posts on its site to other languages.¹¹ In simple terms, machine learning builds on vast amounts of data which are analysed using algorithms to discover patterns. Depending on the goals of analysis and the specific algorithms used, we can distinguish between supervised, unsupervised, and reinforcement learning.^{12,13}

In the case of **supervised learning**, there is a goal, for example understanding how specific advertising affects sales in stores and analysing data from past cases in which sales were particularly high or particularly low. Based on this data, the supervised model can be trained to understand the relationship between specific advertising and sales and is then able to predict similar situations accurately. Note that in supervised

learning, there is a clear understanding of the problem to be solved and its relationship to other factors. Supervised learning also depends on the availability of historical data that can be categorised clearly (Which type of advertising was used? Were sales high or low at that time?).^{14,15}

Unsupervised learning is useful when it is not yet clear what the goal is and how the available data relates to this problem. An unsupervised learning model is good at uncovering correlations and coming up with suggestions on how to group the data you have. In the case of sales, unsupervised learning might suggest other relevant factors for you to take into account, such as the impact of the geographic location of stores on sales and specific types of advertising. In unsupervised learning, the data is not yet categorised and the model is asked to find correlations between different factors.^{16,17}

Reinforcement learning models start from a specific goal to be accomplished; for example, becoming proficient at playing Go. However, the model is not explicitly told how to achieve this and is given data that is not categorised. Learning in this case occurs through trial and error. Given the specific goal to be achieved, the model adjusts its 'approach' and 'assumptions' in order to improve. AlphaGo Zero used reinforcement learning to learn how to play Go by playing against itself without human input.^{18,19} Reinforcement learning is also used to train AI that is proficient at more than one board game.



Machine learning builds on **big data**, which is often defined through reference to the volume, variety, and velocity of the data.²⁰ For the purpose of this report, it is useful to distinguish between various sources of big data in relation to human activity. Big data generated by **what people say** includes online news, social media, radio, and TV. Big data generated by **what people do** includes traffic movements, mobile communication, financial transactions, postal traffic, utility consumption, and emissions of various kinds. On the other hand, we can also consider different **sources of big data**. Digital data is generated automatically by digital services such as global positioning system (GPS) data from mobile phones. Online information includes data generated by activities on the Internet. Geospatial data includes data from satellites and remote sensors.

Before moving on to mapping the relationships between AI and mediation, three points are worth emphasising. First, AI is a moving target in the sense that the term is often used to describe the cutting-edge of technology development. This sometimes prevents us from perceiving and analysing the various ways in which the technology is already influencing our everyday lives.^{21,22} Some AI already seems ordinary to us, such as Google search or Google translate. However, this does not mean that they are less important than more complex AI systems. When it comes to AI as a tool for peacemaking, it is

important to explore the full breadth of AI applications. Tools that are no longer routinely referred to as AI, such as smart searches, need to be kept in mind.

Second, AI is often used as a **suitcase** or **umbrella term**. 'It's a concept that appears simple enough but is actually endlessly complex and packed – like a suitcase – with lots of other ideas, concepts, processes and problems.'²³ This serves as a useful reminder that discussions around AI and mediation are most useful when the term AI is unpacked by using specific examples referring to specific applications. In this sense, the discussion about the application of AI to mediation is most useful when specific examples are discussed.

Third, especially when it comes to peacemaking, it is worth emphasising that AI is a dual-use technology. In simple terms, any innovation in AI might be used for good or for ill. While AI can support peacemaking and humanitarian efforts, it can also be used to exacerbate conflict situations.²⁴ Deep fakes, which describe 'content [that] is doctored [using AI] to give uncannily realistic but false renderings of people', are a particular concern.²⁵ Debates around lethal autonomous weapons systems (LAWS) and other military applications of AI are another reminder of the dual-use nature of AI and the potential impact of AI on conflict situations.

A three-part typology to think through the relationship of AI and mediation

While this report focuses on the use of AI as a tool for mediation, it is important to create an understanding of the broader relationship between peacemaking and AI.^{26,27} To do this, we apply a three part-typology including (a) AI as a topic that mediation teams need to be aware of, (b) AI as a tool for mediation teams to support their work in various stages of mediation, and (c) AI as something that impacts the environment and context in which mediation takes place, particularly in the context of online misinformation.

Mediators need to be aware of **AI as a topic for negotiation** as it is likely to become part of or impact future peace negotiations. In a previous report, we illustrated how AI is becoming part of the topics that diplomats are negotiating.²⁸ We introduced the distinction between (a) AI triggering shifts in the ways in which existing topics are discussed and (b) AI bringing new topics to international agendas and specific negotiations. We are already

seeing shifts in debates around the economy, security, ethics, and human rights. At the same time, new topics such as debates related to LAWS are entering international agendas. The same dynamic will likely affect international conflict resolution.

In addition to AI as a topic of mediation, **AI can serve as a tool for mediators** and their teams. This is not to say that all future mediation will include the use of AI. Rather, certain tools and their context-specific application can support mediation and increase the effectiveness of the work of mediators. Awareness of the tools and the opportunities and challenges associated with them will be key. In the following section of this report we illustrate this in greater detail and discuss the application of specific AI tools in relation to mediation.

Lastly, mediators need to be aware of **shifts in the environment and context in which mediation is practised**



due to advances in AI. Two concerns stand out in particular: (a) the emergence of new kinds of conflicts around the application or development of AI, and (b) the use of AI in conflict situations. The current debate around advances in AI is sometimes framed as a competition, in particular as a competition between the USA and China. This 'race for AI' is sometimes likened to the Cold War. It is conceivable that as more and more countries race to benefit from advances in AI and develop various AI

applications, including military and security applications, AI will become a component of conflict situations and consequently an element that needs to be considered in mediation. Related to this is the potential use of AI in conflict situations. We have already mentioned LAWS. In addition, various surveillance applications of AI, for example facial recognition, will become elements of conflict situations. Mediators need to be aware of this potential of AI.

AI literacy for mediators

Regardless of whether mediators and their teams are using tools to support mediation, they 'have a responsibility to be literate about the technologies present in the mediation environment'.²⁹ In broader terms, the United Nations Secretary-General's Strategy on New Technologies argues that 'engagement with new technologies is necessary for preserving the values of the UN Charter and the implementation of existing UN mandates'.³⁰ Even if mediators and their teams do not rely on the use of AI tools to support their work, AI is likely to become a topic of mediation and a cause for shifts in the environment in which mediation is practised. This means that regardless of whether mediators plan on using AI for their work, they need to have an awareness of AI and its potential impact on the content and context of mediation.

This raises the question of what knowledge and skills mediators should have or develop in relation to AI. It is clear that not everyone can or should become a technical expert. Everyone involved in mediation who uses AI or is likely impacted by it should have a basic understanding of it. A good way of framing this is to use the concept of technological literacy and apply it to AI.

The idea of technological literacy is certainly not a new one.³¹ As a starting point, it can be defined as 'having knowledge and abilities to select and apply appropriate technologies in a given context'.³² Many definitions of technological and digital literacy focus on the **ability to use a given technology appropriately**. In addition, it is important to add a component that includes **critical thinking**, which in the first instance is the 'critical

assessment of the impact of digital technology on [...] society'.³³

When it comes to AI literacy, the two components of digital literacy also apply. On the one hand, there are the skills and knowledge needed to use technology appropriately for a given task and in a given context. On the other hand, there are the skills and knowledge needed to assess the implications of this use for the given conflict and for wider society. These skills are all the more important as we are only in the early stages of thinking through the relevance of AI for mediation. First, **the usefulness of specific tools** needs further exploration; challenges and opportunities in a given context need to be recognised. Second, the **implications of AI for the broader context of peacemaking** need to be thought through. For example, what role will LAWS play in future conflicts? To what extent will deep fakes develop into spoilers in peace processes and how can mediators respond effectively? Third, mediators need to be able to **navigate both hype and pessimism** when it comes to AI applications. They need to be careful not to oversell the possibilities of the technology while not missing out on new opportunities for peacemaking.

It is crucial to be able to navigate between hype and pessimism when it comes to AI and to recognise both opportunities and challenges as they arise in a given context. Regardless of the extent to which mediators and their teams wish to include AI as part of their tool box, this level of AI literacy for mediators will be crucial going forward.



AI as a tool for mediators: Mapping the field



Mediation is about people and interpersonal relationships. It is clear that the essence of mediation will not change due to the introduction of new technology in the field. However, AI tools can usefully assist mediators and their teams in various ways. In other words, the core assumption that mediation is about people does not contradict exploring the potential use of AI in the context of mediation. In fact, thinking of AI as a tool for mediation raises a number of interesting questions: What elements, if any, are we comfortable outsourcing to a machine? Are there any tasks that

could be automated? What does 'meaningful human control',³⁴ a term borrowed from the discussion on LAWS, mean in this context? In what cases and to what extent do we know or need to know how an AI system arrived at a certain result? In this section, we introduce a number of ways in which AI applications can usefully support the activities of mediators and their teams and also raise points of caution regarding the use of AI. In the final section, we discuss the potential impact of AI tools on trust in the context of mediation.

Potential applications and uses of AI

To better understand the potential of AI to serve as a tool for mediators, it is useful to build on the distinction between assisted, augmented, and automated intelligence.³⁵ **Assisted intelligence** supports the work of a human being, **augmented intelligence** allows humans to do something that they would otherwise not be able to accomplish, and **automated intelligence** describes those cases in which the entire task is performed by an AI. Generally speaking, 'the technology's greater power is in complementing and augmenting human capabilities.'³⁶ A 2018 report by PricewaterhouseCoopers argues that currently, the best applications for AI in business are (a) the automation of simple tasks and processes and (b) the analysis of unstructured data.³⁷ Automation receives lots of attention in general debates on AI, partially due to concerns about the loss of jobs. In diplomatic practice, we see attempts at automation in the area of consular affairs where a first contact with citizens is managed via chat bots.³⁸ In the area of conflict arbitration, we also see attempts at automation.³⁹ Neither of these examples carries over into mediation practice as it is understood here. While most applications discussed in the following fall under the category of assisted and augmented intelligence, there is a small potential for automating elements of mediation tasks through smart searches and automated summaries.

In the following, we look at the potential of AI applications for mediation in three areas: knowledge management, understanding conflict situations and the parties to the conflict, and creating inclusivity through interaction with the wider population.

Knowledge management for mediators and their teams

A vast amount of information about the process of mediation, guidelines, best practices, and lessons learned is available for mediators and their teams. The resources provided through *UN Peacemaker* are a great example of this – including key UN documents, mediation guidance, UN guidance for effective mediation, and the peace agreement database.⁴⁰ However, **information and key lessons are often not readily available** because traditional computerised search methods are not helpful when data is mostly available in unstructured form, meaning that it is not available in an organised form according to predefined categories.

Smart searches can make information and knowledge already accumulated more readily available and easier to search. Such searches would go beyond a simple keyword search in a repository and draw connections between various types of documents. Textual data



relevant for mediation – including treaties, mediational manuals, and notes from the field – are often available in an unstructured form and spread across a variety of locations. Using NLP, AI applications can help in creating better access to and analysis of this data.⁴¹ This would save time in the research and preparatory activities of mediators and make the lessons learned from previous mediation activities more readily available. In addition, such a system might draw connections between various documents that did not occur to human researchers. This ability of AI applications to **discover patterns and draw connections** might make it particularly useful for knowledge management tasks. Challenges of applying NLP include understanding context, dealing with large and diverse vocabularies, dealing with different meanings, and grasping word play and ambiguity.⁴²

An example of such an application of AI for knowledge management is the Cognitive Trade Advisor (CTA). Introduced at the 2018 World Trade Organization (WTO) Public Forum, the system helps with the analysis of rules of origin stipulations across a large number of treaties.⁴³ In addition to identifying rules of origin stipulations, the system gives additional information according to context, responds to more complex queries, and includes a virtual assistant to answer questions and provide further graphical illustration.⁴⁴ CTA was built to showcase that it is possible to apply AI to facilitate the research and preparation process of negotiators (including saving time and resources) and to tackle complexity.⁴⁵

From this perspective, it is worth considering to what extent a resource like *UN Peacemaker* can be turned into a smart repository. However, applications for this purpose need to be specifically trained – usually involving human experts who highlight key issues in the early training phase of a neural network using supervised learning.⁴⁶ This means that off-the-shelf solutions are not likely to yield good results and that time and resources need to be invested in training such tools. This also means that their application will be narrow and that the goals and scope of the project need to be well-defined in advance when working with a technical team to realise such an AI application.⁴⁷

Knowing about conflict situations and the parties in conflict

The key for any successful mediation is preparation and an excellent understanding of the conflict situation and the parties to the conflict. It is crucial to 'identify the actors and their interests, their position, what are their best alternatives, in order to come up with a strategy'.⁴⁸

AI applications can assist in creating this understanding, it is worth stressing from the outset, however, that these applications will only ever be in the realm of assisted or augmented intelligence. They do not work well without human input, human oversight, and the addition of contextualisation that can only be provided by human insight.

In terms of using new technologies to better understand a conflict situation and the parties to the conflict, **social media** is discussed prominently at the moment. Generally speaking, social media can be used by mediators and their teams to understand local attitudes and the wider attitudes towards the conflict.⁴⁹ **Sentiment analysis**, using widely available AI tools, is a prominent example of this. It includes a combination of NLP, computational linguistics, and text analytics to extract 'subjective information in source material'.⁵⁰ **Network analysis** can help to show prominent voices in the discussion and uncover connections between those voices. As part of a network analysis, a network of key participants in a debate on a specific topic is built which illustrates connections between participants. Key nodes in the system are assumed to be well-connected individuals, who can connect diverse groups and are important for a topic to gain traction in the discussion.⁵¹

The analysis of social media allows for finding information that would otherwise be excluded from view, identifying key people, monitoring key societal groups and social movements, discovering commonalities, asking the right questions, and pre-empting the escalation of conflict situations. In addition, if mediators and their teams engage in communication with the public, this type of analysis is helpful in adapting communication campaigns to the target group and in monitoring the reception and effectiveness of these messages. Of course, this focus on social media comes with the caveat that only those opinions represented on social media are captured. This is a bias that any such analysis needs to take into account and depending on the mediation context might prove to be a substantial limit to the effectiveness of social media analysis for mediation.

A step further, as far as AI applications are concerned, are **integrated systems** which bring different sources of information, including social media, together. Recently, the Department of Crisis Prevention, Stabilisation, Post-conflict Care and Humanitarian Aid of the German Federal Foreign Office launched an initiative to 'evaluate publicly available data on social, economic, and political developments' to increase the department's ability to detect crisis at an early stage.⁵² A similar tool,



Haze Gazer, was developed and tested by UN Global Pulse Jakarta in co-operation with the government of Indonesia. It was developed to help address the forest and peatland fires that happen in Indonesia every year and affect the environment and livelihoods in the Southeast Asian region. It is a web-based tool that combines three types of data: (a) open data in the form of fire hotspot information from satellites and baseline information on population density and distribution; (b) citizen-generated data from the national complaint system in Indonesia called LAPOR! as well as citizen journalism from videos uploaded to an online news channel; and (c) real-time big data from social media channels, online video channels, like YouTube, and online photo sharing platforms, such as Instagram.⁵³ The tool provides real-time information on the location of fires and haze, the location of the most vulnerable populations, the most affected regions, and the behaviour of affected populations.

It is conceivable that similar tools can support the work of mediators and their teams. More generally, getting a **broad picture of the conflict situation** – a kind of 3-D real-time picture⁵⁴ – facilitates the process of mediation through creating a better understanding of the situation. In addition, better agreements are possible, at least theoretically, through the inclusion of factors hitherto unconsidered, the possibility of finding additional facts that can be leveraged on the way towards an agreement, and the ability of mediators and their teams to provide a fuller picture to the parties in conflict.⁵⁵

Regardless of the AI tools being used, human insight – including values and perceptions – are indispensable when putting the findings and suggestions of AI tools into context. Further, human mediation experts must ask the right questions in the first place to help design more effective and better targeted tools. Having said that, findings and suggestions of AI tools might be very helpful in challenging stereotypes and prejudices and in pointing mediators to issues they have not yet considered.⁵⁶

Creating greater inclusivity

Inclusivity is the cornerstone of effective mediation. Inclusivity means that mediators take the needs of the broader society into account and involve a variety of stakeholders in the process in order to bring a variety of perspectives to the table.⁵⁷ This builds on the assumption that the parties to the conflict may not be representative of the wider public and that other voices, though not represented in the formal negotiations, need to be

heard in order to create broader legitimacy.⁵⁸ Herein lies the possibility to create a ‘democratisation of peace-making’ and go beyond a focus on elites in the mediation process.⁵⁹ Previous publications on cybermediation discuss a number of cases in which new technological tools, in particular social media or dedicated websites, have been used to foster inclusivity in mediation.^{60,61}

As we have seen, AI applications can play a role when it comes to the **automated analysis of vast amounts of unstructured textual data**. This is particularly useful when it comes to analysing the results of large-scale public consultations. The use of machine learning to analyse public opinions and sentiments in these cases has been explored in a mediation context by the UN Department of Political and Peacebuilding Affairs’ (DPPA’s) Middle East Division (MED).⁶² MED is also working on a tool to ‘evaluate the public’s receptivity to an aspect of a peace agreement’ via digital focus groups across large constituencies in various Arab dialects. Such a tool would allow ‘thousands of members of a concerned constituency in a country and its diaspora (e.g. refugees) to be consulted in real time’.⁶³ This information is not only important to further contextualise the peace process, it might also be crucial for shaping agreements and ultimately the success of a mediation effort.

In aiming for greater inclusivity, it is important to **go beyond social media and data generated through Internet use**. UN Global Pulse ran a project in Uganda that analysed radio conversations on public policy and state initiatives. More than half of Ugandan households rely on the radio as their primary source of information and use radio shows to call in and voice their opinions.⁶⁴ Instead of relying on input generated on social media, the project was able to bridge the digital divide and include voices that a narrow focus on textual or social media data would have missed. It did so by using text-to-speech AI applications to convert spoken words into text and including relevant conversations in a searchable database.

At the 2019 AI for Good Global Summit, IBM’s Project Debater showcased the ability to provide a written **synthesis from larger amounts of texts generated through public consultation**. The project’s Speech by Crowd service started from a ‘collection of free-text arguments from large audiences on debatable topics to generate meaningful narratives out of these numerous independent [crowd-sourced] contributions’.⁶⁵ Speech by Crowd is particularly noteworthy for its ability to synthesise text into summaries of key points. Given further recent



developments in the ability of AI applications to generate coherent texts given specific inputs, such as Open

AI's GPT-2,⁶⁶ it is reasonable to expect AI applications of this kind to play a role in the future of mediation efforts.

Key considerations and precautions when using AI tools

While it is important to explore new possibilities and tools for mediation, it is equally important to remain pragmatic about the feasibility of these tools. In our current times of AI hype, there might be a temptation to over-sell the technology. In the words of a mediator interviewed regarding new technologies and mediation, it is important to analyse workflows and adapt the technology accordingly.⁶⁷ This means that any introduction of new technology to the work of mediators and their teams **needs to fit the existing circumstances and context of the mediation** and the needs of the mediation team. Further, as the application of AI tools in the field of humanitarian, development, and legal sectors are further explored in the future, a more nuanced picture regarding the application of these and similar tools for the context of mediation will no doubt emerge. Having said that, three broad areas that need to be taken into consideration and that demand a cautious approach can be clearly identified: questions around bias and neutrality; questions regarding data security, privacy, and human rights; and questions around collaboration with the private sector. These areas will remain key concerns for the foreseeable future.

AI between bias and neutrality

Bias is often discussed in relation to mediators and their teams and guidelines give advice on how to avoid the impression of bias. '[I]f a mediation process is perceived to be biased, this can undermine meaningful progress to resolve the conflict.'⁶⁸ However, when we look at AI as a tool for mediation, the question of bias takes on a new and prominent meaning.

AI tools to support the work of mediators – by providing better knowledge management, a better understanding of the conflict situation and the parties in conflict, and greater inclusivity of the process – are not neutral. Mediators and their teams need to be **aware of potential sources of bias** and put effort into minimising sources of potential bias. All of the AI tools we have been discussing so far are examples of machine learning. As already mentioned, machine learning depends on vast amounts of data for its training. However, big data does not ensure neutrality.

Rather, when using AI applications, we also need to be concerned with **data quality**. Biases in the training data will lead to biases in the outcomes. Such biases can exacerbate disadvantages and discrimination along socioeconomic, racial, or gender lines. This is a particular concern, when far-reaching decisions are based on such biased results. While it is impossible to avoid all forms of bias, an important question is to ask whether the data captures the important aspects of the phenomenon under consideration.

To illustrate bias in machine learning systems, an early version of Google image recognition is often cited. The system was designed to recognise key elements in a photo, for example whether the image contained a human being. However, the image recognition was unable to identify people of colour, most likely due to a lack of diversity in the training data.⁶⁹ Another prominent example of worrying bias in an AI application is the chatbot 'Tay', which was designed by Microsoft to engage in casual conversations on Twitter. However, as Twitter users began tweeting sexist and racist remarks at the bot, it quickly picked up these positions.⁷⁰

Machines function on the basis of what humans tell them. If a system is fed with human biases (conscious or unconscious) the result will inevitably be biased. The lack of diversity and inclusion in the design of AI systems is therefore a key concern: instead of making our decisions more objective, they could reinforce discrimination and prejudices by giving them an appearance of objectivity.⁷¹

While some potential bias can be mitigated by carefully choosing data sets that include diversity, the **black box nature of machine learning algorithms** is much harder to address. The recent ethics guidelines for trustworthy AI published by the EU call for **transparency** of AI applications. This includes **explainability**, which 'requires that the decisions made by an AI system can be understood and traced by human beings'.⁷² There are limits to this approach, given the nature of machine learning algorithms. However, the importance of transparency as an ethical principle for using AI tools should also be applied to the context of mediation.



Data security and privacy

Given the dependency of AI applications on vast amounts of data, data security and data privacy are key issues when it comes to the use of AI tools. In the context of mediation, these concerns are even more heightened given the volatility of the situation and the potential hostility between parties. Lack of data security or lack of provisions to ensure privacy are issues that could damage the mediator's reputation and thereby jeopardise the process as a whole.

In terms of data protection, the **EU's General Data Protection Regulation (GDPR)** offers a very important starting point that might be applied even outside of the current application of the GDPR. The GDPR is a particularly good point of reference for data privacy standards when it comes to personal data. For example, an AI application 'which gathers or reuses personal data from various sources yet does not inform, explain, and/or obtain the consent of the data subject, could fail the purpose limitation principle'.⁷³ Similarly, **consent for personal data** to be used cannot take the form of a 'blank cheque covering any type of machine learning or AI technology'.⁷⁴ Data protection concerns are particularly important for vulnerable groups whose ability to consent might be impaired or diminished.

Regarding securing data from a technological perspective, it is useful to think in terms of **(a) securing the data location, (b) securing the data format, and (c) securing data by design**.⁷⁵

Broadly speaking, the following points are helpful to consider when it comes to **data security and ensuring privacy**⁷⁶:

- Keep sensitive and personal data to a minimum.
- Make the processing of personal data transparent to those whose data is used.
- Ensure the purpose of the use of personal data is legitimate and proportionate.
- Encrypt the data that is stored.
- Restrict access to data to only those directly involved.
- Destroy data when the purpose for which it was collected and held is no longer applicable [...]
- Keep sensitive data on secure internal servers, or, if available, reliable external hosts.

Collaborations with the private sector

Using AI as a tool for mediation almost inevitably requires working with the private sector. Very few

off-the-shelf AI applications, perhaps with the exception of certain tools for sentiment analysis on social media, will be suitable for the specific context of mediation. Collaboration will be needed and can take a number of forms. First, international organisations with substantial expertise in using data and AI in their specific context are great potential partners. In particular, organisations working in the humanitarian field might be useful partners for sharing data and insights. Among others, these include UN Global Pulse and the Centre for Humanitarian Data, which is managed by the United Nations Office for the Coordination of Humanitarian Affairs. Second, some form of collaboration with the private sector might also be required because some form of customisation or a collaboration needs to be negotiated with private sector companies. For example, the tools explored by MED depend on collaboration with private sector companies.

Looking more closely at potential collaboration with the private sector, it is first crucial to underline that working with the private sector should **never undermine the (perceived) impartiality of the process or the mediator and their team**. Second, questions such as data security, data storage, and data ownership, as already mentioned, need to be clearly and openly discussed when negotiating any form of collaboration. Third, mediation teams should choose their level of engagement with the private sector based on careful planning and needs assessments. In this regard, it is useful to distinguish between **three levels of engagement** – free user, customer, or partner⁷⁷ – and define which of these relationships is desirable and useful for the given context.

When negotiating the customer or partner relationship, it will be **crucial to ask the right kind of questions** before proceeding. The following questions offer a flavour of what these could entail⁷⁸:

- Is there a need to analyse historical records and identify patterns – an area of descriptive analytics?
- Is this about performing sentiment analysis on documents, social media communication, or even speeches?
- Is the important aspect the ability to analyse large amounts of documents according to specific criteria?
- Is the interest located in forecasting future events as part of predictive analytics?

Questions like these guide the project description, further research activities, negotiations with the private sector company, and implementation.⁷⁹



Implications for trust and buy-in

The success of mediation relies fundamentally on the 'trust built between a mediator, or a mediation team, and the parties, and the ability to generate and maintain buy-in to the process'.⁸⁰ Some go further and argue that 'the mediator acts as a crutch of trust'.⁸¹ While the parties to the conflict do not trust each other, they must be able to place trust in the mediators and their teams. Most commonly, it is argued that trust and buy-in depend on the interpersonal skills of the mediator and their teams.⁸²

The **use of new technologies can have implications for trust and buy-in** and therefore the success of conflict resolution. It is worth stressing again that we cannot view AI applications as tools that provide a neutral perception of the conflict and the parties involved on which trust can be built almost automatically. Algorithmic biases, biased data, and the dangers of misuse of data need particular attention. Bias, be it actual or perceived, and the misuse of data, in particular that related to discrimination and violations of privacy, can have a huge impact on trust in relation to mediation.

Mediators operate in an environment where the **dangers of misinformation and disinformation** have already become more prominent due to new technology and various (mis-)uses of social media.⁸³ Problems around trust are already heightened and will become further exacerbated by deep fakes. In this sense, AI is already shifting the environment in which mediation is practised. Indeed, debates around trust in new technology, and AI in particular, have gained considerable prominence. For example, in his speech at the 2018 Internet Governance Forum, French president Macron called for building the 'Internet of trust'.⁸⁴

To make sense of trust in the context of mediation and AI, it is useful to distinguish between (a) trust in the tools of mediation, (b) trust enjoyed by mediators and their teams, and (c) trust between the parties in conflict.

Trust in the tools of mediation starts with consent. At its core, mediation depends on the consent of those involved. Elements of consent include 'the integrity of the mediation process, security and confidentiality [as well as] the acceptability of the mediator and the mediating entity'.⁸⁵ In the context of cybermediation and in particular the use of AI, this need for consent should be extended to include the technological tools used as part of the mediation. At the most basic level, this means that the parties to the conflict are informed about the tools that the mediation team uses. Further, the parties

to the conflict should be given a say in decisions related to more complex AI tools, especially those dealing with sensitive information. Data privacy and security need to be ensured and sensitive cases explicitly discussed by all the parties involved. Mediators and their teams need to be able to give credible assurances regarding their practices when it comes to data. Transparency regarding algorithms, in line with, for example, the EU's ethics guidelines on trustworthy AI, is also an important factor to consider. Explainability of the algorithms behind certain AI tools can become a crucial factor in certain cases. Parties to the conflict should be encouraged to ask for an explanation of how an algorithmic recommendation or decision was reached and mediators and their teams should be able to present this in an easy-to-understand way.

Trust in the tools of mediation and **trust in mediators and their teams** are closely intertwined. Some argue that the impartiality of the mediator and trust in the mediator are often linked and that both are key for successful mediation.⁸⁶ Others maintain that bias in mediators, particularly when it comes to representatives of states acting as a third party, is common and has no bearing on the success of mediation.⁸⁷ More again stress that the leverage the mediator has in relation to the parties in conflict is more important than trust in order to achieve success in mediation.⁸⁸ For this study, we assume that trust in the mediators and their teams is an element contributing to mediation success. A negative perception of the tools that mediators use can influence the perception of mediators and their teams and erode the trust they enjoy. As we have discussed, AI tools always carry concerns regarding bias. Hence, it will be particularly important to assure the parties to the conflict that such biases are well understood by mediators and their teams and that measures are being taken to minimise their occurrence and impact. Some might argue that algorithmic biases and biased data sets as technical issues, which, if detected, can be corrected. While this is certainly true to an extent, it is equally important to acknowledge that in the sensitive context of mediation, technical issues have the potential to become politicised and damage the mediator's reputation.

Finally, a profound lack of trust between the parties to the conflict is a key reason why mediation becomes necessary. **Trust between the parties in conflict** can be built as part of the process of mediation and the implementation of the agreement.⁸⁹ The potential role played by AI applications in this context is worth considering.

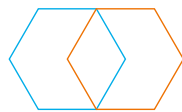


If the parties to the conflict can be brought together to decide on some key aspects of the AI tools intended for use to analyse the conflict or post-conflict monitoring, trust in these tools can be built and, more importantly, the first steps towards trust between the parties can be taken. By making decisions on some of the technical tools involved in the process, a 'degree of working trust',⁹⁰ understood as 'trust that the other side is

genuinely committed, largely out of its own interest, to finding an accommodation'⁹¹ can be built. Further, mistrust between the parties in conflict tends to flare up again during the discussion of the details of implementation.⁹² Jointly agreeing on tools for implementation and monitoring, including the use of AI tools, might be a practical way of inserting confidence-building measures into this critical phase of the process of mediation.



Conclusions and further recommendations



The overview provided in this report makes it clear that mediators and their teams cannot afford to be ignorant regarding AI. To illustrate this, we introduced a **three-part typology to map the relationship between mediation and AI**. The typology includes (a) AI as a topic for negotiation, (b) As a tool for mediation, and (c) AI as something that impacts the environment in which mediation is practised. This means that regardless of whether mediators intend to actually use AI tools to support the process of mediation, they need to have a basic AI literacy.

Looking more specifically at AI as a tool for mediation, the report identified **three areas in which AI applications can usefully support the work of mediators and their teams**: (a) supporting knowledge management and background research, (b) generating a good understanding of the conflict and the parties to the conflict, and (c) creating greater inclusivity of mediation processes. As new AI tools are developed and their application in related fields, such as development co-operation and humanitarian assistance, is further explored, additional applications for the context of mediation are likely to emerge and the scope of potential applicability to mediation is likely to widen.

The report also highlighted **three areas of consideration and precaution** when using AI as a tool for mediation. First, the potential for bias – both related to data and algorithms – needs to be acknowledged and actively addressed by mediators and their teams. Transparency about these questions, especially in communication with the parties to the conflict, is crucial. Second, since AI tools depend on vast amounts of data, data privacy and security, especially in relation to sensitive data related to the conflict or the parties in conflict, need to be addressed and the highest standards need to be upheld. Lastly, since the use of AI tools almost inevitably involves some form of collaboration with the private sector, mediators need to give this relationship careful thought.

The report identified **three ways in which trust in the context of mediation might be impacted by the use of AI tools**. First, if AI tools are used by mediators and their teams, it is important to undertake steps to build trust in these tools, in particular through explicit communication about them and transparency about data sets and algorithms. Second, the link between trust in the AI tools used and trust in the person of the mediator needs to be acknowledged. Negative perceptions of the AI tools used might impact the credibility of the mediator. Third, AI tools might also be used to build a ‘working trust’ between the parties in conflict, in particular when it comes to monitoring and implementation.

Building on these points, **four additional recommendations with practical implications** for the work of mediators can be made. These might also serve as inspiration for the next steps towards the use of AI tools to support conflict resolution.

Capacity building for AI in mediation: The report argued that AI literacy for mediators and their teams is crucial. In essence, this means that a basic understanding of AI applications and their applications needs to be built. In the future, this might be taken further to include concrete measures towards capacity building for the use of AI in mediation. On an individual level, this can be envisioned as a three-tier structure including (a) a foundation level which focuses on developing basic AI literacy, (b) a practitioner level which focuses on training individuals in the use of AI tools and related data security and privacy questions, and (c) an expert level which includes individuals with skills to develop and implement AI tools.⁹³ The UN Department of Political and Peacebuilding Affairs seems a natural starting point for thinking about capacity building for AI in mediation.

Considering new members for mediation teams: The UN Guidance for Effective Mediation advises parties to ‘reinforce the mediator with a team of specialists, particularly experts in the design of mediation processes,



country/regional specialists, and legal advisers, as well as with logistics, administrative, and security support. Thematic experts should be deployed as required'.⁹⁴ In light of the narrative of this report, careful thought should be given to adding AI experts to mediation teams where needed, either regarding AI as a topic for mediation or AI as a tool for mediation. While not every mediation team will be in need of such topical or technical experts, the need for people trained and specialised in this area might arise more prominently in the future. For example, a conflict might involve the deployment of deep fakes, cyber-attacks involving the applications of AI, or AI-powered weapons systems (including LAWS). The mediation will benefit from topical experts who can help develop an effective mediation strategy and a way towards conflict resolution in these cases. Similarly, technical experts can usefully be added when it comes to tailoring existing AI applications or developing new ones.

Working with the private sector: If the aim is to make specific AI tools useful for mediation and to incorporate them into the work of mediators and their teams, some form of collaboration with the private sector will most likely be needed. It is crucial to define the relationship clearly, based on a needs assessment, and to clarify the

terms of this collaboration, especially when it comes to data security and data privacy. Those active in the field of mediation should consider working towards developing best practices and ethical guidelines for this type of collaboration.

Engaging in foresight scenario planning and proof of concept:

As indicated in this report, there are few examples of AI tools being currently employed in the context of mediation. Therefore, it will be important to develop further ideas around the potential application of existing and future AI tools in the context of mediation. Foresight scenario planning can be a useful method in this regard.⁹⁵ Further, test cases for the application of specific tools should be run independent of actual ongoing negotiations to illustrate the potential of AI tools to support mediators and to deliver proof of concept for specific applications before employing them in the field.

While mediation will, at its core, remain a profoundly human endeavour, this report has shown that it is useful to carefully think through the potential of new technology, AI in particular, for mediation. Any tool that usefully supports the work of international conflict resolution should be explored.



Notes

- 1 For an overview of recent activities and events: DiploFoundation (2019) *#CyberMediation*. Available at <https://www.diplomacy.edu/cybermediation> [accessed 10 November 2019].
- 2 Rosen Jacobson B et al. (2018) *Data Diplomacy: Updating Diplomacy to the Big Data Era*. Available at https://www.diplomacy.edu/sites/default/files/Data_Diplomacy_Report_2018.pdf [accessed 10 November 2019].
- 3 DiploFoundation (2019) *Mapping the Challenges and Opportunities of AI for the Conduct of Diplomacy*. Available at <https://www.diplomacy.edu/sites/default/files/AI-diplo-report.pdf> [accessed 10 November 2019].
- 4 DiploFoundation (2019) *Diplo's AI Lab*. Available at <https://www.diplomacy.edu/AI> [accessed 10 November 2019].
- 5 Gavrilović A (2019) [WebDebate summary] AI for peacemaking: New tools. *Diplo* blog, 17 June. Available at <https://www.diplomacy.edu/blog/webdebate-summary-ai-peacemaking-new-tools> [accessed 10 November 2019].
- 6 Berridge GR and Lloyd L (2012) *The Palgrave Macmillan Dictionary of Diplomacy*. [3rd ed]. Houndmills, Basingstoke: Palgrave Macmillan, p. 239.
- 7 Rosen Jacobson B et al. (2018) *Data Diplomacy: Updating Diplomacy to the Big Data Era*. Available at https://www.diplomacy.edu/sites/default/files/Data_Diplomacy_Report_2018.pdf [accessed 10 November 2019].
- 8 DiploFoundation (2019) *Mapping the Challenges and Opportunities of AI for the Conduct of Diplomacy*. Available at <https://www.diplomacy.edu/sites/default/files/AI-diplo-report.pdf> [accessed 10 November 2019].
- 9 Marr B (2018) The key definitions of artificial intelligence (AI) that explain its importance. *Forbes*, 14 February. Available at <https://www.forbes.com/sites/bernardmarr/2018/02/14/the-key-definitions-of-artificial-intelligence-ai-that-explain-its-importance/#3b6263494f5d> [accessed 10 November 2019].
- 10 Walsh T (2018) *Machines that Think. The Future of Artificial Intelligence*. Amherst, NY: Prometheus.
- 11 Brooks R (2017) Machine learning explained. *MIT* blog, 28 August. Available at <https://rodnebrooks.com/forai-machine-learning-explained/> [accessed 10 November 2019].
- 12 Enhance Data Science (2017) Machine Learning Explained: Supervised Learning, Unsupervised Learning, and Reinforcement Learning. Available at <http://enhancedatascience.com/2017/07/19/machine-learning-explained-supervised-learning-unsupervised-learning-and-reinforcement-learning/> [accessed 10 November 2019].
- 13 Machado G (2016) ML basics: Supervised, unsupervised and reinforcement learning. *Medium* blog, 6 October. Available at <https://medium.com/@machadogj/ml-basics-supervised-unsupervised-and-reinforcement-learning-b18108487c5a> [accessed 10 November 2019].
- 14 Enhance Data Science (2017) *Machine Learning Explained: Supervised Learning, Unsupervised Learning, and Reinforcement Learning*. Available at <http://enhancedatascience.com/2017/07/19/machine-learning-explained-supervised-learning-unsupervised-learning-and-reinforcement-learning/> [accessed 10 November 2019].
- 15 Machado G (2016) ML basics: Supervised, unsupervised and reinforcement learning. *Medium* blog, 6 October. <https://medium.com/@machadogj/ml-basics-supervised-unsupervised-and-reinforcement-learning-b18108487c5a> [accessed 10 November 2019].
- 16 Enhance Data Science (2017) *Machine Learning Explained: Supervised Learning, Unsupervised Learning, and Reinforcement Learning*. Available at <http://enhancedatascience.com/2017/07/19/machine-learning-explained-supervised-learning-unsupervised-learning-and-reinforcement-learning/> [accessed 10 November 2019].
- 17 Machado G (2016) ML basics: Supervised, unsupervised and reinforcement learning. *Medium* blog, 6 October. <https://medium.com/@machadogj/ml-basics-supervised-unsupervised-and-reinforcement-learning-b18108487c5a> [accessed 10 November 2019].
- 18 Sing S et al. (2017) Learning to play Go from scratch. *Nature* 550, pp. 336–337. Available at <https://www.nature.com/articles/550336a> [accessed 10 November 2019].
- 19 Silver D et al. (2017) Mastering the game of Go without human knowledge. *Nature* 550, pp. 354–359. Available at <https://www.nature.com/articles/nature24270> [accessed 10 November 2019].
- 20 Rosen Jacobson B et al. (2018) *Data Diplomacy: Updating Diplomacy to the Big Data Era*, p. 18. Available at https://www.diplomacy.edu/sites/default/files/Data_Diplomacy_Report_2018.pdf [accessed 10 November 2019].
- 21 Elkus A (2016) The futility of (narrow) speculation about machines and jobs. *Medium* blog, 19 October. Available at <https://medium.com/@Aelkus/the-futility-of-narrow-speculation-about-machines-and-jobs-a116a0672659> [accessed 10 November 2019].
- 22 Bailey K (2016) Reframing the 'AI effect'. *Medium* blog, 27 October. Available at <https://medium.com/@katherinebailey/reframing-the-ai-effect-c445f87ea98b> [accessed 10 November 2019].
- 23 Future Today Institute (2018) *2018 Tech Trends Report*, p. 47. Available at <https://futuretodayinstitute.com/2018-tech-trends-annual-report/> [accessed 10 November 2019].
- 24 Pandya J (2019) The dual-use dilemma of Artificial Intelligence. *Forbes*, 7 January. Available at <https://www.forbes.com/sites/cognitiveworld/2019/01/07/the-dual-use-dilemma-of-artificial-intelligence/#2376bbb26cf0> [accessed 10 November 2019].
- 25 Murphy H (2019) Cyber security companies race to combat 'deep fake' technology. *Financial Times*, 16 August. Available at <https://www.ft.com/content/63cd4010-bfce-11e9-b350-db00d509634e> [accessed 10 November 2019].
- 26 At Diplo, we use this typology to explore the relationship between diplomacy and big data. Rosen Jacobson B et al. (2018) *Data Diplomacy: Updating Diplomacy to the Big Data Era*, p. 18. Available at https://www.diplomacy.edu/sites/default/files/Data_Diplomacy_Report_2018.pdf [accessed 10 November 2019].
- 27 At Diplo, we applied this typology to think through the relationship between AI and diplomatic practice. DiploFoundation (2019) *Mapping the Challenges and Opportunities of AI for the Conduct of Diplomacy*. Available at <https://www.diplomacy.edu/sites/default/files/AI-diplo-report.pdf> [accessed 10 November 2019].
- 28 DiploFoundation (2019) *Mapping the Challenges and Opportunities of AI for the Conduct of Diplomacy*, pp. 14–15. Available at <https://www.diplomacy.edu/sites/default/files/AI-diplo-report.pdf> [accessed 10 November 2019].
- 29 Jenny J et al. (2018) *Peacemaking and New Technologies. Dilemmas and Options for Mediators*. Mediation Practice Series, Centre for Humanitarian Dialogue, p. 6. Available at <https://www.hcdcentre.org/wp-content/uploads/2018/12/MPS-8-Peacemaking-and-New-Technologies.pdf> [accessed 10 November 2019].
- 30 UN (2018) UN Secretary-General's Strategy on New Technologies, pp. 4–5. Available at <https://www.un.org/en/newtechnologies/images/pdf/SGs-Strategy-on-New-Technologies.pdf> [accessed 10 November 2019].
- 31 Hayden MA (1989) What is technological literacy? *Bulletin of Science, Technology & Society* 9, pp. 228–233.
- 32 Waetjen WB (1993) Technological Literacy reconsidered. *Journal of Technology Education* 4(2). Available at <https://scholar.lib.vt.edu/ejournals/JTE/v4n2/jte-v4n2/waetjen.jte-v4n2.html> [accessed 10 November 2019].
- 33 Perućica N et al. (no date) *Digital Literacy for Digital Natives*. Available at https://www.diplomacy.edu/sites/default/files/Digital_literacy_for_digital_natives.pdf [accessed 10 November 2019].
- 34 Rosen Jacobson B (2017) *Lethal Autonomous Weapons Systems: Mapping the GGE Debate*. DiploFoundation Policy Papers and Briefs, No. 8. Available at https://www.diplomacy.edu/sites/default/files/Data_Diplomacy_Report_2018.pdf [accessed 10 November 2019].
- 35 PwC (2018) Künstliche Intelligenz als Innovationsbeschleuniger in Unternehmen [Artificial Intelligence as an Accelerator of Innovation in Business]. Available at <https://www.pwc.de/de/>



- digitale-transformation/ki-als-innovationsbeschleuniger-in-unternehmen-whitepaper.pdf [accessed 10 November 2019].
- ³⁶ Daugherty PR and Wilson HJ (2018) *Human + Machine: Reimagining Work in the Age of AI*. Boston, MA: Accenture Global Solutions [e-pub].
- ³⁷ PwC (2018) Künstliche Intelligenz als Innovationsbeschleuniger in Unternehmen [Artificial Intelligence as an Accelerator of Innovation in Business], p. 7. Available at <https://www.pwc.de/de/digitale-transformation/ki-als-innovationsbeschleuniger-in-unternehmen-whitepaper.pdf> [accessed 10 November 2019].
- ³⁸ Höne KE and Chivot E (2019) [Event summary] The impact of AI on diplomacy and international relations, *Diplo* blog, 1 February. Available at <https://www.diplomacy.edu/blog/event-summary-impact-ai-diplomacy-and-international-relations> [accessed 10 November 2019].
- ³⁹ Barnett J and Treleaven P (2018) Algorithmic Dispute Resolution: The Automation of Professional Dispute Resolution Using AI and Blockchain Technologies, *The Computer Journal* 61(3), pp. 399-408.
- ⁴⁰ UN Peacemaker (2019) Mediation Resources Overview. Available at <https://peacemaker.un.org/resources> [accessed 10 November 2019].
- ⁴¹ Editor (2018) CyberMediation: The impact of digital technologies on the prevention and resolution of violent conflicts. *Diplo* blog, 20 December. Available at <https://www.diplomacy.edu/blog/cybermediation-impact-digital-technologies-prevention-and-resolution-violent-conflicts> [accessed 10 November 2019].
- ⁴² Editor (2018) Cybermediation: What role for blockchain and artificial intelligence? *Diplo* blog, 12 October. Available at <https://www.diplomacy.edu/blog/cybermediation-what-role-blockchain-and-artificial-intelligence> [accessed 10 November 2019].
- ⁴³ Aquino Santos MR (2018) Cognitive trading using Watson. *IBM* blog, 12 December. Available at <https://www.ibm.com/cloud/blog/cognitive-trading-using-watson> [accessed 10 November 2019].
- ⁴⁴ DiploFoundation (2019) Mapping the Challenges and Opportunities of AI for the Conduct of Diplomacy, p. 26. Available at <https://www.diplomacy.edu/sites/default/files/AI-diplo-report.pdf> [accessed 10 November 2019].
- ⁴⁵ Höne KE (2018) Shaping an intelligent tech & trade initiative – ITTI driving AI-trade interactions to boost global prosperity. *Digital Watch*. Available at <https://dig.watch/resources/shaping-intelligent-tech-trade-initiative-itti-driving-ai-trade-interactions-boost> [accessed 10 November 2019].
- ⁴⁶ DiploFoundation (2019) *Mapping the Challenges and Opportunities of AI for the Conduct of Diplomacy*, p. 26. Available at <https://www.diplomacy.edu/sites/default/files/AI-diplo-report.pdf> [accessed 10 November 2019].
- ⁴⁷ Ibid. pp. 27-28.
- ⁴⁸ Amon C et al. (2018) *The Role of Private Technology Companies in Support of Mediation for the Prevention and Resolution of Violent Conflicts*. Geneva: Capstone Global Security, Final Report, Geneva Graduate Institute, p. 7.
- ⁴⁹ Jenny J et al. (2018) *Peacemaking and New Technologies. Dilemmas and Options for Mediators*. Mediation Practice Series, Centre for Humanitarian Dialogue, p. 17. Available at <https://www.hdcentre.org/wp-content/uploads/2018/12/MPS-8-Peacemaking-and-New-Technologies.pdf> [accessed 10 November 2019].
- ⁵⁰ Batrinca B and Treleaven PC (2015) Social media analytics: a survey of techniques, tools and platforms. *AI & Society* 30(1), pp 89-116.
- ⁵¹ Rosen Jacobson B et al. (2018) Data Diplomacy: Updating Diplomacy to the Big Data Era, p. 23. Available at https://www.diplomacy.edu/sites/default/files/Data_Diplomacy_Report_2018.pdf [accessed 10 November 2019].
- ⁵² Dickow M and Jacob D (2018) *The Global Debate on the Future of Artificial Intelligence*. Stiftung Wissenschaft und Politik, p. 7. Available at https://www.swp-berlin.org/fileadmin/contents/products/comments/2018C23_dkw_job.pdf [accessed 10 November 2019].
- ⁵³ UN Global Pulse (2016) *Haze Crisis Analysis and Visualization Tool*. Available at https://www.slideshare.net/unglobalpulse/haze-gazer-tool-overview?from_action=save [accessed 10 November 2019].
- ⁵⁴ Osorio D (interview 10 September 2019).
- ⁵⁵ Ibid.
- ⁵⁶ Strugar M (interview 10 September 2019).
- ⁵⁷ UN Peacemaker (2012) *UN Guidance for Effective Mediation*. New York: United Nations, p. 4. Available at https://peacemaker.un.org/sites/peacemaker.un.org/files/GuidanceEffectiveMediation_UNDPA2012%28english%29_0.pdf [accessed 10 November 2019].
- ⁵⁸ Ibid.
- ⁵⁹ Jenny J et al. (2018) *Peacemaking and New Technologies. Dilemmas and Options for Mediators*. Mediation Practice Series, Centre for Humanitarian Dialogue, p. 31. Available at <https://www.hdcentre.org/wp-content/uploads/2018/12/MPS-8-Peacemaking-and-New-Technologies.pdf> [accessed 10 November 2019].
- ⁶⁰ Ibid..
- ⁶¹ UN Department of Political and Peacebuilding Affairs and Centre for Humanitarian Dialogue (2019) *Digital Technologies and Mediation in Armed Conflict*. Available at <https://peacemaker.un.org/sites/peacemaker.un.org/files/DigitalToolkitReport.pdf> [accessed 10 November 2019].
- ⁶² Ibid., p. 14.
- ⁶³ Ibid.
- ⁶⁴ Rosenthal A (2019) When old technology meets new: How UN Global Pulse is using radio and AI to leave no voice behind. *UN Global Pulse* blog, 23 April. Available at <https://www.unglobalpulse.org/news/when-old-technology-meets-new-how-unglobal-pulse-using-radio-and-ai-leave-no-voice-behind> [accessed 10 November 2019].
- ⁶⁵ Slonim N (2019) IBM Project Debater visits the UN. *IBM Research* blog, 6 May. Available at <https://www.ibm.com/blogs/research/2019/05/project-debater-un/> [accessed 10 November 2019].
- ⁶⁶ Seabrook J (2019) The next word. Where will predictive text take us? *The New Yorker*, 14 October. Available at <https://www.newyorker.com/magazine/2019/10/14/can-a-machine-learn-to-write-for-the-new-yorker#> [accessed 10 November 2019].
- ⁶⁷ Amon C et al. (2018) *The Role of Private Technology Companies in Support of Mediation for the Prevention and Resolution of Violent Conflicts*. Geneva: Capstone Global Security, Final Report, Geneva Graduate Institute, p. 8.
- ⁶⁸ UN Peacemaker (2012) *UN Guidance for Effective Mediation*. New York: United Nations, p. 10. Available at https://peacemaker.un.org/sites/peacemaker.un.org/files/GuidanceEffectiveMediation_UNDPA2012%28english%29_0.pdf [accessed 10 November 2019].
- ⁶⁹ Tutt A (2016) An FDA for algorithms. *69 Administrative Law Review* 83, pp. 84-123. Available at <https://ssrn.com/abstract=2747994> [accessed 10 November 2019].
- ⁷⁰ Vincent J (2016) Twitter taught Microsoft's AI chatbot to be a racist asshole in less than a day. *The Verge*, 24 March. Available at <https://www.theverge.com/2016/3/24/11297050/tay-microsoft-chatbot-racist> [accessed 10 November 2019].
- ⁷¹ Mijatović D (2018) In the era of artificial intelligence: safeguarding human rights. *Open Democracy* blog, 3 July. Available at <https://www.opendemocracy.net/digitaliberties/dunja-mijatovic/in-era-of-artificial-intelligence-safeguarding-human-rights> [accessed 10 November 2019].
- ⁷² Independent High-Level Expert Group on Artificial Intelligence (2019) *Ethics Guidelines for Trustworthy AI*, p. 18. Available at <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai> [accessed 10 November 2019].
- ⁷³ DiploFoundation (2019) *Mapping the Challenges and Opportunities of AI for the Conduct of Diplomacy*, p. 31. Available at <https://www.diplomacy.edu/sites/default/files/AI-diplo-report.pdf> [accessed 10 November 2019].
- ⁷⁴ Ibid.
- ⁷⁵ Rosen Jacobson B et al. (2018) *Data Diplomacy: Updating Diplomacy to the Big Data Era*, pp. 61-62. Available at https://www.diplomacy.edu/sites/default/files/Data_Diplomacy_Report_2018.pdf [accessed 10 November 2019].
- ⁷⁶ Ibid., p. 62.
- ⁷⁷ Amon C et al. (2018) *The Role of Private Technology Companies in Support of Mediation for the Prevention and Resolution of Violent Conflicts*. Geneva: Capstone Global Security, Final Report, Geneva Graduate Institute, pp. 3 and 10.
- ⁷⁸ DiploFoundation (2019) *Mapping the Challenges and Opportunities of AI for the Conduct of Diplomacy*, p. 28. Available at <https://www.diplomacy.edu/sites/default/files/AI-diplo-report.pdf> [accessed 10 November 2019].



- www.diplomacy.edu/sites/default/files/AI-diplo-report.pdf [accessed 10 November 2019].
- ⁷⁹ Ibid..
- ⁸⁰ Jenny J et al. (2018) *Peacemaking and New Technologies. Dilemmas and Options for Mediators*. Mediation Practice Series, Centre for Humanitarian Dialogue, p. 46. Available at <https://www.hdcentre.org/wp-content/uploads/2018/12/MPS-8-Peacemaking-and-New-Technologies.pdf> [accessed 10 November 2019].
- ⁸¹ Zartman IW (2018) Mediation roles for large small countries. In Carment D and Hoffman E [eds] *International Mediation in Fragile World*. Abingdon: Routledge [e-pub].
- ⁸² Salem R (2003) Trust in Mediation. Available at https://www.beyondintractability.org/essay/trust_mediation [accessed 10 November 2019].
- ⁸³ Jenny J et al. (2018) *Peacemaking and New Technologies. Dilemmas and Options for Mediators*. Mediation Practice Series, Centre for Humanitarian Dialogue, p. 11. Available at <https://www.hdcentre.org/wp-content/uploads/2018/12/MPS-8-Peacemaking-and-New-Technologies.pdf> [accessed 10 November 2019].
- ⁸⁴ The 'Internet of trust' is described as 'trust in the protection of privacy, trust in the legality and quality of content, and trust in the network itself'. IGF (2018) IGF 2018 Speech by French president Emmanuel Macron. Available at <https://www.intgovforum.org/multilingual/content/igf-2018-speech-by-french-president-emmanuel-macron> [accessed 10 November 2019].
- ⁸⁵ UN Peacemaker (2012) *UN Guidance for Effective Mediation*. New York: United Nations, p. 8. Available at https://peacemaker.un.org/sites/peacemaker.un.org/files/GuidanceEffectiveMediation_UNDPA2012%28english%29_0.pdf [accessed 10 November 2019].
- ⁸⁶ Berridge GR (2015) *Diplomacy. Theory and Practice* [5th ed]. Houndmills, Basingstoke: Palgrave Macmillan, pp. 260-261.
- ⁸⁷ Tuval S (1975) Biased intermediaries: Theoretical and historical considerations. *Jerusalem Journal of International Relations* 1(Fall), pp. 51-69.
- ⁸⁸ Strugar M ((interview 10 September 2019).
- ⁸⁹ Zartman IW (2018) Mediation roles for large small countries. In Carment D and Hoffman E [eds] *International Mediation in Fragile World*. Abingdon: Routledge [e-pub].
- ⁹⁰ Kydd AH (2006) When can mediators build trust. *The American Political Science Review* 100:3, p. 449.
- ⁹¹ Ibid.
- ⁹² Ibid.
- ⁹³ Compare Rosen Jacobson B et al. (2018) *Data Diplomacy: Updating Diplomacy to the Big Data Era*, pp. 51. Available at https://www.diplomacy.edu/sites/default/files/Data_Diplomacy_Report_2018.pdf [accessed 10 November 2019].
- ⁹⁴ UN Peacemaker (2012) *UN Guidance for Effective Mediation*. New York: United Nations, p. 7. Available at https://peacemaker.un.org/sites/peacemaker.un.org/files/GuidanceEffectiveMediation_UNDPA2012%28english%29_0.pdf [accessed 10 November 2019].
- ⁹⁵ Pauwels E (2019) *The New Geopolitics of Converging Risks. The UN and Prevention in the Era of AI*. New York: United Nations University. Centre for Policy Research. Available at <https://collections.unu.edu/eserv/UNU:7308/PauwelsAIGeopolitics.pdf> [accessed 10 November 2019].





