Cornelius Friesendorf | Argyro Kartsonaki (Eds.)

OSCE Insights

Securing States and People

Nomos

IFSH Institute for Peace Research and Security Policy at the University of Hamburg

Cornelius Friesendorf | Argyro Kartsonaki (Eds.)

OSCE Insights

Securing States and People

Nomos

IFSH Institute for Peace Research and Security Policy at the University of Hamburg

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at http://dnb.d-nb.de

ISBN 978-3-7560-1852-9 (Print)
978-3-7489-4585-7 (ePDF)

1st Edition 2025

Published by
Nomos Verlagsgesellschaft mbH & Co. KG
Waldseestraße 3–5 | 76530 Baden-Baden
www.nomos.de

Production of the printed version:
Nomos Verlagsgesellschaft mbH & Co. KG
Waldseestraße 3–5 | 76530 Baden-Baden

ISBN (Print): 978-3-7560-1852-9
ISBN (ePDF): 978-3-7489-4585-7

DOI: https://doi.org/10.5771/9783748945857

Onlineversion
Nomos eLibrary

The Vienna Document 2011 and Military Applications of Artificial Intelligence

Nicolò Miotto[*]

Abstract

The development and deployment of military applications of artificial intelligence (AI) is raising concerns about their negative implications for international security. Misperception, unintended escalation, and proliferation are some of the key potential risks stemming from military uses of AI. This article argues that states within and outside the OSCE region should draw on the OSCE Vienna Document 2011 to develop confidence- and security-building measures (CSBMs) applicable to the military uses of AI. Such CSBMs could help foster dialogue and co-operation by increasing transparency and predictability concerning military applications of AI.

Keywords

OSCE, artificial intelligence, Vienna Document 2011, CSBMs, military transparency

To cite this publication: Nicolò Miotto, "The Vienna Document 2011 and Military Applications of Artificial Intelligence," in *OSCE Insights*, eds. Cornelius Friesendorf and Argyro Kartsonaki (Baden-Baden: Nomos, 2025), https://doi.org/10.5771/9783748945857-01

Introduction

Artificial intelligence (AI) is expected to bring about unprecedented innovation in numerous sectors of society, including defense.[1] Its use in the military promises various technical benefits, including improvements in data collection, strengthened analytical capabilities, and faster decision-making processes. As several countries have manifested their interest in developing military applications of AI, a fierce public debate surrounding their potential technical, (geo)political, and ethical risks has been taking place. While some observers have highlighted that, despite the risks, AI can improve key military capabilities such as early warning and target identification, others have warned against potential risks such as misperception, unintended escalation, and proliferation.[2] In noting these challenges, many have engaged in reflection on potential means of mitigating such threats.

Among other tools, diverse stakeholders have suggested developing confidence- and security-building measures (CSBMs) for military applications of AI to increase transparency, enhance predictability, and avert escalation. Hence, research on CSBMs is expanding, receiving contributions from academia, governments, and the private sector.[3] With that said, these

studies mainly focus on developing new measures that can address both the technical limitations of AI and their potential implications for international security. Little attention has been paid to exploring the applicability of well-established CSBMs to the military uses of AI. In particular, what is lacking—with the single exception of a rather general study[4]—is an analysis of the contribution that the OSCE Vienna Document 2011 (VD11) could make in this regard.[5]

Reflecting on the contributions of the VD11 to the multilateral governance of military uses of AI is of the utmost importance at a time when international discussions on the matter have stalled.[6] Due to the erosion of trust and confidence caused by Russia's war of aggression against Ukraine, it is unlikely that the VD11 will be updated any time soon to cover military applications of AI. Nonetheless, this study argues that states within and outside the OSCE region should draw upon the VD11 to implement CSBMs to increase the transparency and predictability of military uses of AI.

This paper starts by outlining the definitions of AI and CSBMs adopted in this research. It then addresses prominent issues pertaining to military uses of AI and key CSBMs that have been recommended to mitigate related threats. It then explores the main problems underlying the application of CSBMs to military uses of AI, noting that despite these challenges, certain arrangements could likely be implemented successfully. Finally, it shows how key VD11 provisions could be drawn on to establish CSBMs for military uses of AI and provides recommendations in this direction.

Definitions and terminology

Artificial intelligence and its military applications

AI is a much-used umbrella concept that incorporates numerous related technologies and areas of research, including machine learning (ML) and deep learning (DL). Definitions of AI vary depending on the capabilities of the systems in question and their functionalities.[7] Despite their diversity, however, these definitions point to certain general features related to the overall rationale and objectives of AI technologies. Such characteristics include the capacity to simulate human reasoning and perform cognitive tasks that are generally associated with human intelligence.[8]

A closer look at the quantity and quality of the cognitive tasks simulated by these technologies helps to further clarify what AI is by marking the difference between so-called "artificial general intelligence" (AGI)/"artificial super intelligence" (ASI) and "narrow AI." AGI/ASI represents a strictly hypothetical form of AI which would be capable of equaling or surpassing human intelligence and behavior, becoming self-conscious and acquiring the ability to perform tasks, learn, and plan autonomously as humans do.[9] The category of narrow AI, to which current uses of AI belong, comprises "complex software programs that can execute discrete 'intelligent' tasks such

as recognizing objects or people from images, translating language, or playing games."[10] Narrow AI programs include ML and its sub-field, DL.

This paper looks at military applications of AI as an ensemble of narrow AI programs used to carry out specific military tasks such as image recognition, autonomous navigation, and training. This research only considers uses of narrow AI to enhance the capabilities of the weapon and equipment systems covered by the VD11 (e.g., battle tanks, armored combat vehicles, and combat aircrafts).[11] Therefore, certain conventional and non-conventional weapon and equipment systems not covered by the VD11, such as warships and nuclear command, control, and communications, are not considered by this study.

Confidence- and security-building measures (CSBMs)

This paper adopts a general definition of CSBMs, as outlined in early research, as arrangements designed to enhance

> an assurance of mind and belief in the trustworthiness of the announced intentions of other states in respect of their security policies, and the facts with regard to military activities and capacities which are designed to further the objectives of a nation's security policy.[12]

The main objectives of CSBMs are to increase transparency by publicly displaying a state's non-aggressive posture and to enhance predictability by allowing for the detection of inconsistencies in other states' behavior vis-à-vis established CSBMs.[13] The ultimate intended impact of CSBMs is to reduce the risk of unintended escalation and conflict between countries, which could be triggered by misperceptions about other states' military postures and activities. Examples of CSBMs include the notification of military exercises, the observation of military activities, the establishment of communication channels between countries, inspections of military facilities, and the exchange of information on military forces and budgets.[14] These cases mirror the principles and practices outlined in pivotal OSCE documents such as the 1975 Helsinki Final Act[15] and the VD11.

Military applications of AI, associated risks, and CSBMs

Several countries, including the United States, Russia, and China, are heavily investing in AI to modernize their military capabilities.[16] This interest in developing military applications of AI stems from the technical opportunities they offer (such as improvements in target identification and the acceleration of decision-making processes)[17] and from the ambition to equal or surpass competitors' actual and/or perceived capabilities.[18] Projects aimed at integrating AI into military systems encompass a wide range of tools, including unmanned aerial and maritime vehicles, missile technology, nuclear capabilities, and space systems. AI is being developed and tested to support other military tasks,

including command and control, information management, logistics, and training.[19] Existing AI capabilities in these sectors include collateral damage estimation, the geolocation of images, the provision of recommendations on best paths and transport modes, and the tracking of individuals' learning progress.[20] The strong interest in further improving these tools and developing new ones is driven by the advantages AI offers, such as enhanced assessment accuracy, faster analysis and communication, and lower logistics costs.[21]

Despite these promising opportunities, researchers, public institutions, and civil society organizations have expressed several concerns about the military uses of AI. Indeed, the technology is vulnerable to several limitations. For instance, technical issues such as changes in the data distribution can negatively impact the performance of AI models.[22] Furthermore, malicious actors can affect the integrity of data by manipulating the training datasets, thus leading AI models to fail or to act differently than expected.[23] Additional issues such as psychological constraints can affect human-machine interactions; for example, end-users can act upon erroneous analytical outputs due to unconditional trust in AI data analysis capabilities.[24]

In a military context, these and further issues can have serious security implications, potentially undermining international security. Possible technical failures range from errors in autonomous navigation to target misidentification, paving the way for concerning scenarios such as diplomatic tensions, escalation, and even overt military conflict.[25] In response to these challenges, academics, policymakers, and private companies have recommended different types of CSBMs. These can be grouped into two main categories based on the issues they aim to address: (1) CSBMs that address potential technical issues with AI software; and (2) CSBMs that address inter-state security dynamics underlying the development and deployment of military applications of AI. The first category includes measures such as the publication of system cards[26] to provide information about the capabilities and limitations of AI models and the use of content provenance and watermarking methods to verify the authenticity and integrity of AI-generated data.[27]

CSBMs from the second category include broader arrangements such as the establishment of Track II initiatives[28] to promote dialogue on the risks posed by military uses of AI and the releasing of joint political declarations on the maintenance of human control over decisions concerning target engagement.[29] Additional measures include tabletop exercises to simulate crisis scenarios and develop tailored responses, the establishment of hotlines between countries, and the development of incident sharing agreements to consolidate knowledge of AI technical failures and their impact on security.[30]

These CSBMs represent valuable measures to mitigate key potential threats. However, their effective implementation faces several challenges stemming from the current geopolitical environment and the intrinsic characteristics of AI technol-

ogy. Analyzing these limitations can help us to understand which CSBMs are more likely to contribute to the goals of enhancing transparency and predictability.

Challenges and opportunities for the application of CSBMs to the military uses of AI

Geopolitical and technical challenges

While the need to engage in talks about military applications of AI and their regulation has been recognized by the academic and policymaking community, several dilemmas continue to pose obstacles to the implementation of concrete measures. Geopolitical tensions following Russia's war of aggression against Ukraine represent a prominent example of the challenges affecting the negotiation of CSBMs. Indeed, CSBMs can be seen as the ultimate representation of a shared understanding of what constitutes common security concerns.[31] Their effective negotiation depends on the establishment of confidence and trust between states. Hence, their development is conditional on rebuilding trust and confidence and achieving a common notion of which issues pertaining to military applications of AI represent security matters of reciprocal interest.

Moreover, in such a contested environment, it is unlikely that states will adopt intrusive AI software–focused CSBMs such as system cards. This has already been highlighted in the research on cyber CSBMs, which notes that nonlikeminded countries are unlikely to implement intrusive measures such as the observation of cyber exercises in order to maintain a degree of secrecy over cyber capabilities.[32] Indeed, states that have deployed cutting-edge military applications of AI are unlikely to publicly acknowledge the limitations or potential biases that affect their functioning, especially vis-à-vis adversaries' deployment of such technologies. This would be detrimental to their security interests and could reveal gaps in military effectiveness. When AI software transparency is weighed up against the projection of military power, the balance often tips in favor of the latter.

Dilemmas inherent to the technology only add to these geopolitical challenges. As noted by recent research, there is much uncertainty about whether AI and its military applications can be effectively tested to verify that systems are functioning and behaving as originally intended, designed, and expected and about which techniques and methods can be employed to best conduct technical assessments.[33] This overall uncertainty has serious implications for CSBMs as it calls into doubt what can be verified with certainty about the military uses of AI. In the face of this uncertainty, not only are countries likely to refrain from implementing AI software–related CSBMs, but, even if circumstances were different, they would face technical challenges to effectively ensuring the safety of military uses of AI.

Despite these notable challenges, shedding light on existing co-operative dynamics between states in the international environment and shifting

the focus from AI software to military hardware can help us to assess whether less intrusive measures are more feasible and can be effectively implemented.

Opportunities for politically and technically feasible CSBMs

While the security environment is competitive and characterized by strong tensions, multilateral discussions on the military applications of AI have already taken place at intergovernmental fora before and following Russia's war of aggression against Ukraine, including at the OSCE. At the OSCE, formal and informal discussions have been particularly focused on the impact of AI on law enforcement and crime,[34] freedom of expression and media pluralism,[35] human rights,[36] and international law.[37] Attention has also been paid to the military uses of AI. For example, informal discussions on these issues took place between 2014 and 2021, bringing to the table governmental and non-governmental representatives from OSCE participating States.[38]

Most importantly, from 2019 to 2021 the OSCE Parliamentary Assembly (PA) and the Forum for Security Co-operation (FSC) hosted formal political discussions between OSCE participating States on the military uses of AI.[39] Such engagement also included discussions on whether existing arms control frameworks, including the VD11, should be updated to account for the military uses of AI. While such discussions have not taken place at either the PA or the FSC recently, they have continued in other formats, expanding formal political engagement beyond Europe by including the OSCE Asian Partners for Co-operation.[40]

Therefore, while geopolitical tensions are hindering in-depth discussions on the overall arms control architecture and eroding trust and confidence, evidence also points to the fact that more limited but important informal and formal discussions are already taking place at the multilateral level within and outside the OSCE region. Although such engagement primarily involves like-minded countries, it nevertheless represents an important step, paving the way for future discussions when the security environment allows.

Technical issues concerning the verification and validation of AI software should not overshadow the potential benefits of applying less intrusive and more technically feasible CSBMs to AI-integrated military hardware.[41] Research on cyber CSBMs has shown that arrangements such as the exchange of information on cyber doctrines and the organization of cyber forces are likely to be implemented, even among non-likeminded countries.[42] Moreover, likeminded states are more open to discussing and implementing even intrusive CSBMs such as those concerning the prior notification and observation of military cyber exercises.[43] This is not mere theory, as the OSCE already represents an existing successful model. Between 2013 and 2016, the Organization served as a platform for adopting a total of sixteen voluntary cyber CBMs which encompass a wide set of arrangements, ranging from information exchanges on cyber doctrines, strategies,

and policies to the voluntary reporting of cyber vulnerabilities.[44]

Furthermore, key CSBMs can be applied to AI-integrated military hardware. For example, if a state were to deploy an unmanned aerial vehicle (UAV) equipped with AI autonomous navigation software to better conduct military intelligence gathering at its borders, its neighbors may be more interested in why it deployed such technology and whether this indicates a change in its military posture than in whether the UAV's AI software works effectively. This observation opens the door for the implementation of certain CSBMs to increase transparency between states by signaling a non-aggressive military posture and to enhance predictability by helping to detect anomalies in states' behavior. If the AI software cannot be inspected due to security concerns, secrecy requirements, and lack of effective methodologies, then measures should focus on the deployment of military hardware and its implications. In this sense, the VD11 could serve as a basis for implementing concrete measures to mitigate certain detrimental inter-state security dynamics underlying the development and deployment of military applications of AI.

CSBMs for military uses of AI: The VD11 as a source

The VD11 does not cover military uses of AI, and therefore its applicability to this domain is strictly dependent on future updates to the document. Due to existing politico-military tensions, it is unlikely that the VD11 will be amended in the near future. Nonetheless, OSCE participating States should draw upon VD11 provisions to create voluntary CSBMs to increase transparency and predictability concerning the military uses of AI. Similarly, states outside the OSCE region should use the VD11 as an inspiration for similar measures. The feasibility of applying the various CSBMs outlined in VD11 to military uses of AI can be assessed following the same logic as that used in the previous section's discussion of which measures are more likely to be implemented in the near future. The CSBMs set out in the VD11 offer a crucial means of improving transparency, allowing states to assess each other's intentions and military postures. They could also enhance predictability by providing diplomatic channels for discussing states' behavior with regard to the development and employment of military applications of AI.

Because it is unlikely that states will adopt intrusive CSBMs allowing for the inspection of AI software, other more feasible VD11 arrangements could be considered. Moreover, because it is highly difficult to validate and verify AI models,[45] such arrangements would need to tackle other issues first. For example, states could address the destabilizing implications of reciprocal uncertainty concerning military budget allocations and weapons development.[46] Additionally, countries could dispel concerns related to newly developed military doctrines that contemplate the use of new and emerging technologies.[47] If they are not addressed, these matters risk destabilizing

inter-state relations, leading to misperceptions and erroneous assessments of other countries' intentions and military postures. These uncertainties are particularly impactful in the case of AI since states are competing to develop its military applications and, consequently, are heavily investing in this endeavor.[48] The VD11 contains numerous CSBMs to shed light on military expenditure, military research and development, and military doctrines and strategies, thus providing an effective means of assessing countries' intentions.

While it is unlikely that states will implement CSBMs concerning the demonstration of military cyber capabilities,[49] this does not necessarily apply to the military uses of AI. Indeed, if the capabilities are looked at from a hardware (rather than a software) perspective, states may be interested in showcasing how AI is being employed to enhance the performance of a given weapon and equipment system. For instance, a state might be interested in demonstrating (including to its adversaries) its use of AI to improve the navigation capabilities of an armored vehicle, as a means of showcasing advances in its defense capabilities. In doing so, it would not need to share the technical characteristics of the AI software, the algorithm underlying the ML model, or the training dataset used. Certainly, such a demonstration would be limited in scope, but it would provide insight into how that state intends to use military applications of AI. The VD11 therefore offers an important basis for providing general information about AI-integrated weapon and equipment systems.

Although intrusive CSBMs are less likely to be implemented, this does not mean that arrangements should not consider the security implications of potential technical failures of AI software. Indeed, a mere technical failure could be read as a discrepancy in a state's behavior and military posture and could thus generate tensions. If the autonomous navigation system of an AI-powered UAV were to fail, for example, causing it to accidentally cruise into the airspace of a rival neighboring country, this could be mistakenly interpreted as a hostile act. In such cases, there is a need to quickly reassure adversaries in order to dispel concerns and avert unintended escalation. In this sense, crisis hotlines are a valuable means of responding to such emergencies. The VD11 provides for well-structured measures that could support states under these circumstances.

Recommendations

The following recommendations focus on often overlooked but prominent VD11 CSBMs, in particular key provisions outlined in Chapter II ("Defence Planning"), Chapter III ("Risk Reduction"), and Chapter IV ("Contacts"). These measures, in contrast to provisions such as the annual exchange of military information, have yet to receive sufficient attention. In addition, they provide a feasible field for action in contrast to other VD11 provisions such as Chapter VI ("Observation of Certain Military Activities"), which would likely be perceived as particularly sensitive and

intrusive. Drawing on the CSBMs set out in the VD11, states within and outside the OSCE region should consider:

Implementing information exchange on defense planning concerning military applications of AI. VD11 Chapter II, "Defence Planning," foresees information exchange between OSCE participating States regarding their

> intentions in the medium to long term as regards size, structure, training and equipment of [their] armed forces, as well as defence policy, doctrines and budgets related thereto.[50]

The exchange of such information aims to increase transparency and promote dialogue between participating States. These provisions require participating States to exchange information on the "training programmes for their armed forces and planned changes thereto in the forthcoming years," as well as the "procurement of major equipment and major military construction programmes […], either ongoing or starting in the forthcoming years."[51] In addition, if information is available, participating States are expected to provide "the best estimates specifying the total and figures for […] research and development" with regard to the last two years of the forthcoming five fiscal years.[52] As part of their information exchange, OSCE participating States should consider the voluntary provision of details and estimates on budget allocations, military research and development, AI-integrated weapon and equipment systems, and new military doctrines that include the employment of military applications of AI. States outside the OSCE region should establish similar mechanisms to provide insights into their intensions and military postures in the medium and long term.

Using existing platforms and/or developing new ones to discuss the information exchanged. According to VD11 Chapter II, any participating State can ask for clarification on the defense planning–related information provided by another participating State. High-level discussions on the information are envisaged in the format of the Annual Implementation Assessment Meeting (AIAM), the High-Level Military Doctrine Seminar (HLMDS), and study visits.[53] The HLMDS is a particularly relevant format for discussing such matters. It brings together high-level military and civilian representatives such as chiefs of defense and/or chiefs of general staff, diplomats, and academics, who discuss doctrinal changes, their impact on military structures, and the military information exchanged. OSCE participating States should consider voluntarily discussing the information exchanged at the HLMDS. States outside the OSCE region should use similar structures or develop new ones to engage in dialogue on the impact of AI on military structures and doctrines, exchanging views on white papers, defense policies, and military doctrines.

Establishing co-operation as regards hazardous incidents of a military nature involving military applications of AI. VD11 Chapter III.17, "Co-operation as Regards Hazardous Incidents of a Military Nature," outlines measures to prevent possible misunderstandings in the event

of a military incident.[54] If a hazardous incident of a military nature occurs, the participating State whose military forces are involved in the incident should provide information to other participating States, and any participating State affected by the incident can also request clarification. This general mechanism could be employed in the event of incidents involving military applications of AI such as the hypothetical cases concerning AI-powered UAVs outlined in the previous sections. In line with the provisions of this chapter, participating States have an established point of contact (PoC) to better co-ordinate communications in the event of a hazardous incident of a military nature. In the context of military uses of AI, participating States should employ this mechanism to dispel concerns. States outside the OSCE region should develop similar measures, such as crisis hotlines, thus reducing the risk of accidental military escalation. PoCs can quickly provide both technical and political information to the relevant counterpart(s), warning against potential weapon system failures and dispelling concerns about the nature of the military activity.

Holding discussions on hazardous incidents of a military nature involving military applications of AI. As outlined in Chapter III.17, hazardous incidents of a military nature can be discussed at the FSC and at the AIAM.[55] In the context of the military applications of AI, these discussions could help to clarify the nature of the incidents and to pave the way for broader dialogue on the security risks posed by AI and means of averting escalation. In particular, discussions could address the possible repercussions of diverse technical malfunctions for international security. OSCE participating States should hold these talks at the AIAM to foster dialogue. States outside the OSCE region should bring discussions to existing venues or create new platforms for discussing such matters.

Using existing data-sharing tools and/or developing new ones as incident sharing repositories. Details on incidents involving military uses of AI such as location, type of weapon or equipment system involved, and the nature of the incident (for example airspace infringement, target misidentification) should be shared between states within and outside the OSCE region. An example of a data-sharing tool that participating States could employ is the OSCE Communications Network, which is used for information exchange under the VD11. Following the example of the Communications Network, states outside the OSCE region should develop data-sharing tools to share information on the incidents and engage in political discussions informed by accurate, evidence-based analyses.

Organizing demonstrations of new types of AI-integrated major weapon and equipment systems. VD11 Chapter IV.31, "Demonstration of New Types of Major Weapon and Equipment Systems," requires any participating State that deploys "a new type of major weapon and equipment system" to "arrange […] a demonstration for representatives of all other participating States."[56] As countries are deploying military applications of AI, these demonstrations could be particularly helpful in creating occasions for dialogue and

co-operation. Participating States should consider applying this CSBM to the military uses of AI. Accordingly, participating States that deploy new types of AI-integrated major weapon and equipment systems should arrange demonstrations for the representatives of all other participating States. For instance, a participating State could demonstrate how new types of armored vehicles employ autonomous navigation for path planning and real-time path adjustment and explain how these new types of weapon and equipment systems fill the gaps of previous versions of military hardware. States outside the OSCE region should consider implementing similar measures at the bilateral and multilateral levels. Notably, such demonstrations would still allow countries to maintain their technological advantage, as general information about the relevant military hardware capabilities could be shared without requiring the sharing of AI software.

Discussing the results of the demonstrations. According to VD11 provisions, following up on the demonstrations, participating States can discuss observations and results at key OSCE fora such as the FSC and the AIAM. States outside the OSCE region should bring these discussions to existing regional fora or develop new venues for such engagement. Such discussions could be particularly valuable as opportunities not only for addressing present concerns but also for raising technical and political matters related to future deployments of military applications of AI.[57]

Notes

1 Darrell M. West and John R. Allen, *Turning Point: Policymaking in the Era of Artificial Intelligence* (Washington: Brookings Institution Press, 2020).

2 See Jessica Cox and Heather Williams, "The Unavoidable Technology: How Artificial Intelligence Can Strengthen Nuclear Stability," *Washington Quarterly* 44, no. 1 (2021): 69–85; István Szabadföldi, "Artificial Intelligence in Military Application: Opportunities and Challenges," *Land Forces Academy Review* 26, no. 2 (2021): 157–65.

3 Michael C. Horowitz and Paul Scharre, AI and International Stability: Risks and Confidence-Building Measures (Washington, DC: Center for a New American Security, 2021), https://s3.us-east-1.amazonaws.com/files.cnas.org/documents/AI-and-International-Stability-Risks-and-Confidence-Building-Measures.pdf; Marina Favaro, Strengthening the OSCE's Role in Strategic Stability (Atlantic Council, 2022), https://www.atlanticcouncil.org/wp-content/uploads/2022/01/Strategic-Insights-Memo_OSCE-and-Strategic-Stability_1.12.22-1.pdf; Anna Nadibaidze, Commitment to Control Weaponised Artificial Intelligence: A Step Forward for the OSCE and European Security (Geneva: Geneva Centre for Security Policy, 2022), https://www.gcsp.ch/publications/commitment-control-weaponised-artificial-intelligence-step-forward-osce-and-european; Sarah Shoker et al., "Confidence-Building Measures for Artificial Intelligence: Workshop Proceedings," arXiv:2308.00862 [cs.CY], arXiv, August 3, 2023, https://arxiv.org/abs/2308.00862

4 Favaro, cited above (Note 3).

5 OSCE, Vienna Document 2011 on Confidence- and Security-Building Measures, FSC.DOC/1/11 (Vienna: November 30, 2011), https://www.osce.org/fsc/86597

6 Ingvild Bode et al., "Prospects for the Global Governance of Autonomous Weapons: Comparing Chinese, Russian, and US Practices," *Ethics and Information Technology* 25, no. 1 (2023): Article 5.

7 IBM Data and AI Team, "Understanding the Different Types of Artificial Intelligence," IBM, October 12, 2023, https://www.ibm.com/blog/understanding-the-different-types-of-artificial-intelligence/

8 Ralf T. Kreutzer and Marie Sirrenberg, "What Is Artificial Intelligence and How to Exploit It?," in *Understanding Artificial Intelligence: Fundamentals, Use Cases and Methods for a Corporate AI Journey* (Cham: Springer, 2020), 1–57.

9 Scott McLean et al., "The Risks Associated with Artificial General Intelligence: A Systematic Review," *Journal of Experimental & Theoretical Artificial Intelligence* 35, no. 5 (2021): 649–63.

10 Vincent Boulanin, ed., *The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk*, vol. 1, *Euro-Atlantic Perspectives* (Stockholm: SIPRI, 2019), 14, https://www.sipri.org/sites/default/files/2019-05/sipri1905-ai-strategic-stability-nuclear-risk.pdf

11 The full list of weapon and equipment systems is reported in Annex III of the VD11.

12 Johan Jørgen Holst, "Confidence-Building Measures: A Conceptual Framework," *Survival* 25, no. 1 (1983): 2.

13 Abbott A. Brayton, "Confidence-Building Measures in European Security," *The World Today* 36, no. 10 (1980): 382–91; Erica D. Borghard and Shawn W. Lonergan, "Confidence Building Measures for the Cyber Domain," *Strategic Studies Quarterly* 12, no. 3 (2018): 10–49.

14 Holst, cited above (Note 12); Brayton, cited above (Note 13).

15 CSCE, Helsinki Final Act (Helsinki: 1975), https://www.osce.org/helsinki-final-act

16 Margarita Konaev et al., U.S. Military Investments in Autonomy and AI: A Strategic Assessment (Center for Security and Emerging Technology, 2020), https://cset.georgetown.edu/wp-content/uploads/U.S.-Military-Investments-in-Autonomy-and-AI_Strategic-Assessment-1.pdf; Samuel Bendett et al., Advanced Military Technology in Russia: Capabilities and Implications (London: Chatham House, 2021), https://www.chathamhouse.org/sites/default/files/2021-09/2021-09-23-advanced-military-technology-in-russia-bendett-et-al.pdf

17 Eric Robinson, Daniel Egel, and George Bailey, Machine Learning for Operational Decisionmaking in Competition and Conflict: A Demonstration Using the Conflict in Eastern Ukraine (Santa Monica, CA: RAND Corporation, 2023), https://www.rand.org/pubs/research_reports/RRA815-1.html

18 Anna Nadibaidze and Nicolò Miotto, "The Impact of AI on Strategic Stability Is What States Make of It: Comparing US and Russian Discourses," *Journal for Peace and Nuclear Disarmament* 6, no. 1 (2023): 47–67.

19 Elsa B. Kania, "Chinese Military Innovation in the AI Revolution," *RUSI Journal* 164, no. 5–6 (2019): 26–34; Thomas Reinhold and Niklas Schörnig, eds., *Armament, Arms Control and Artificial Intelligence: The Janus-Faced Nature of Machine Learning in the Military Realm* (Cham: Springer Nature, 2022); Sarah Grand-Clément, Artificial Intelligence beyond Weapons: Application and Impact of AI in the Military Domain (Geneva: UNIDIR, 2023), https://unidir.org/wp-content/uploads/2023/10/UNIDIR_AI_Beyond_Weapons_Application_Impact_AI_in_the_Military_Domain.pdf

20 Kania, cited above (Note 19); Reinhold and Schörnig, cited above (Note 19); Grand-Clément, cited above (Note 19).

21 Grand-Clément, cited above (Note 19).

22 For a detailed overview of the issue, see Joaquin Quiñonero-Candela et al., *Dataset Shift in Machine Learning* (Cambridge, MA: The MIT Press, 2022).

23 Maaike Verbruggen, "No, Not That Verification: Challenges Posed by Testing, Evaluation, Validation and Verification of Artificial Intelligence in Weapon Systems," in *Armament, Arms Control and Artificial Intelligence: The Janus-Faced Nature of Machine Learning in the Military Realm*, eds. Thomas Reinhold and Niklas Schörnig (Cham: Springer Nature, 2022), 175–91; Ioana Puscas, AI and International Security: Understanding the Risks and Paving the Path for Confidence-Building Measures (Geneva: UNIDIR, 2023), 22–26, https://unidir.org/wp-content/uploads/2023/10/UNIDIR_AI-international-security_understanding_risks_paving_the_path_for_confidence_building_measures.pdf

24 James Johnson, "The AI Commander Problem: Ethical, Political, and Psychological Dilemmas of Human-Machine Interactions in AI-Enabled Warfare," *Journal of Military Ethics* 21, no. 3–4 (2022): 246–71.

25 Puscas, cited above (Note 23).

26 System cards are documents that report the intended uses and limitations of AI models. They can also provide the results of red teaming exercises. An example is the system card of the Generative Pre-trained Transformer 4 (GPT-4) released by OpenAI. See OpenAI, "GPT-4 System Card," March 23, 2023, https://cdn.openai.com/papers/gpt-4-system-card.pdf

27 Shoker et al., cited above (Note 3); Furkan Gursoy and Ioannis A. Kakadiaris, "System Cards for AI-Based Decision-Making for Public Policy," arXiv:2303.04754 [cs.CY], arXiv, March 1, 2022, https://arxiv.org/abs/2203.04754

28 Track II diplomacy typically involves experts or influential individuals who engage in dialogue on crucial matters but do not represent official capacities.

29 Nadibaidze, cited above (Note 3).

30 Horowitz and Sharre, cited above (Note 3); Favaro, cited above (Note 3); Shoker et al., cited above (Note 3).

31 Holst, cited above (Note 12), 3.

32 Jürgen Altmann, "Confidence and Security Building Measures for Cyber Forces," in *Information Technology for Peace and Security: IT Applications and Infrastructures in Conflicts, Crises, War, and Peace*, ed. Christian Reuter (Wiesbaden: Springer Vieweg, 2019), 197.

33 Verbruggen, cited above (Note 23).

34 OSCE, "2019 OSCE Annual Police Experts Meeting: Artificial Intelligence and Law Enforcement—an Ally or Adversary?," https://www.osce.org/event/2019-annual-police-experts-meeting

35 Eliska Pirkova et al., Spotlight on Artificial Intelligence and Freedom of Expression: A Policy Manual, eds. Deniz Wagner and Julia Haas (Vienna: OSCE Office of the Representative on Freedom of the Media, 2021), https://www.osce.org/files/f/documents/8/f/510332_1.pdf

36 OSCE, "Artificial Intelligence Poses Risks but Can Also Contribute to More Open and Inclusive Societies, Say Participants at ODIHR Event," October 6, 2023, https://www.osce.org/odihr/554413

37 OSCE, "OSCE Court of Conciliation and Arbitration Moot Court Explores Space Activities and Artificial Intelligence," November 16, 2022, https://www.osce.org/court-of-conciliation-and-arbitration/531383; OSCE, "Moot Court in the Framework of MUNLAWS Conference," https://www.osce.org/court-of-conciliation-and-arbitration/553960

38 See: OSCE, "Panel of Eminent Persons," https://www.osce.org/networks/pep; OSCE, "OSCE Security Days," https://www.osce.org/sg/secdays

39 OSCE Parliamentary Assembly, Luxembourg Declaration (Luxembourg: July 4–8, 2019), 4, https://www.oscepa.org/en

/documents/annual-sessions/2019-luxembourg/3882-luxembourg-declaration-eng/file; European Union, OSCE Forum for Security Co-operation N°955, Vienna, 23 September 2020, EU Statement on New Technologies, FSC.DEL/207/20 (Vienna: September 25, 2020), https://www.osce.org/files/f/documents/5/5/466311.pdf; European Union, OSCE Forum for Security Co-operation N°975, Vienna, 12 May 2021, EU Statement on Challenges of New Generation Warfare, FSC.DEL/173/21 (Vienna: May 14, 2021), https://www.osce.org/files/f/documents/7/a/487063.pdf

40 OSCE, "Inter-Regional Conference on the Impact of Emerging Technologies on International Security and Democracy, in the OSCE Asian Partnership for Co-Operation Group Framework," https://www.osce.org/partners-for-cooperation/asian/544219

41 The terms "software" and "hardware" refer to the instructions run by a computer and the physical components constituting the computer system, respectively. In a UAV equipped with AI for autonomous navigation, for example, the software would be the algorithms, while the hardware would be the UAV itself.

42 Altmann, cited above (Note 32).

43 Borghard and Lonergan, cited above (Note 13), 23–24.

44 For a comprehensive list of OSCE cyber confidence-building measures, see OSCE, Decision No. 1202 OSCE Confidence-Building Measures to Reduce the Risks of Conflict Stemming from the Use of Information and Communication Technologies, PC.DEC/1202 (March 10, 2016), https://www.osce.org/pc/227281

45 Verbruggen, cited above (Note 23).

46 Holst, cited above (Note 12).

47 Panel of Eminent Persons on European Security as a Common Project, Back to Diplomacy: Final Report and Recommendations of the Panel (Panel of Eminent Persons on European Security as a Common Project, 2015), 15, https://www.osce.org/files/f/documents/2/5/205846.pdf

48 Nadibaidze and Miotto, cited above (Note 18).

49 Altmann, cited above (Note 32).

50 OSCE, cited above (Note 5), 7.

51 OSCE, cited above (Note 5), 8.

52 OSCE, cited above (Note 5), 9.

53 According to VD11 Chapter XI, "Annual Implementation Assessment Meeting," the AIAM is to be held each year. The VD11 also encourages the holding of "periodic" military seminars, without specifying how often they should take place. See OSCE, cited above (Note 5), 9.

54 OSCE, cited above (Note 5), 13.

55 OSCE, cited above (Note 5), 13.

56 OSCE, cited above (Note 5), 18–19.

57 The author would like to express his gratitude to Lara Maria Guedes and Andrea Miotto for discussions on the implications of artificial intelligence for international security. He is also grateful to Anna Nadibaidze, Argyro Kartsonaki, two anonymous reviewers and the language editor for their valuable feedback on previous drafts. All opinions expressed in this paper are those of the author alone and do not reflect the positions of the OSCE.

The Moscow Mechanism of the OSCE: Rules, Practice, and Possible Improvements

Wolfgang Benedek[*]

Abstract

While the Moscow Mechanism has remained relatively dormant for some time, the emergence of major challenges for the human dimension of the OSCE and a lack of alternative means of quickly investigating issues concerning alleged violations of human rights have resulted in a resurgence of its use. After outlining the rules governing its construction and the different ways in which it has been invoked, this contribution analyzes how the Moscow Mechanism is used in practice. It provides an overview of past missions and considers the advantages and challenges associated with its application. It then assesses the missions' outcomes and follow-up activities, explores the Mechanism's strengths and weaknesses, and closes with recommendations for its future implementation.

Keywords

Moscow Mechanism, Vienna Mechanism, OSCE, ODIHR, human dimension, human rights fact-finding

To cite this publication: Wolfgang Benedek, "The Moscow Mechanism of the OSCE: Rules, Practice, and Possible Improvements," in *OSCE Insights*, eds. Cornelius Friesendorf and Argyro Kartsonaki (Baden-Baden: Nomos, 2025), https://doi.org/10.5771/9783748945857-02

Introduction

The Moscow Mechanism allows OSCE participating States to obtain, without the need for consensus, a fact-finding report written entirely by independent experts on an issue or situation related to the human dimension commitments of the OSCE.[1] It was adopted at the Moscow Meeting of the CSCE on the human dimension in 1991, when the CSCE was seeking new tools to address the challenge of protecting its human dimension commitments. This meeting reconfirmed previous agreements stemming from the Vienna Follow-up Conference (ending in 1989), known as the Vienna Mechanism, and from the Charter of Paris (1990).[2] It adopted additional rules to strengthen them, providing for the possibility of investigating their alleged violations.

The Moscow Mechanism was applied a number of times in the 1990s, mostly in the context of the war in the Balkans, and then very rarely until 2018, when it was used to investigate reports of a clampdown on LGBTQ+ people in Chechnya. According to the OSCE, the Moscow Mechanism has been invoked fifteen times.[3] Its increased use since 2018 has revealed its advantages and shortcomings, as well as the challenges associated

with its application. One such challenge, for example, pertains to the refusal of some participating States, such as Russia and Belarus, to co-operate when the issue to be investigated concerns activities under their purview. These states have sought to justify their non-cooperation by arguing that the Mechanism has become outdated and that the OSCE can discuss these matters in its bodies.[4] The vast majority of OSCE participating States have taken a different position, however, and have increasingly made use of the Mechanism. Its reports serve a wider function than merely informing debates in the OSCE. Nonetheless, experience also shows that there is room for methodological improvement. This contribution therefore ends with several recommendations for how to make the Mechanism more effective.[5]

The rules

The Vienna Follow-up Conference, which ended in 1989, decided to hold three meetings on the human dimension of the CSCE, to take place in Paris (1989), Copenhagen (1990), and Moscow (1991). The Vienna Mechanism, agreed in the Vienna Concluding Document of 1989,[6] was a first step toward improving the implementation of commitments in the human dimension. It set out an obligation to provide a written response to requests for information by other participating States. In Moscow, in order to enhance the effectiveness of the Document on the Copenhagen Meeting on the Human Dimension of 1990[7] and to strengthen and expand the Vienna Mechanism, the deadlines first introduced by the Copenhagen Meeting were shortened.[8] Upon the issuing of a formal request, participating States now had to respond within ten days, while requests for bilateral meetings had to be replied to as soon as possible, as a rule within one week. In addition, the Moscow Document laid out the elements of a new mechanism that would allow for the establishment of ad hoc missions by independent experts to investigate alleged violations of human dimension commitments, i.e. the "Moscow Mechanism."[9] The final version of the Moscow Mechanism contains minor amendments by the CSCE made in Helsinki (1992) and in Rome (1993).[10]

There are different ways of invoking the Moscow Mechanism and very strict rules for its application.[11] It may in certain cases be preceded by an invocation of the Vienna Mechanism. In general, the Moscow Mechanism can be applied via self-invocation or the invocation of another participating State (or States), and the process can take either a co-operative or a contentious approach. In the case of self-invocation, the aim is "to address or contribute to the resolution of questions in [a state's own] territory relating to the human dimension" (Moscow Mechanism, para. 4).[12] Such an approach would be co-operative. Table 1 at the end of this section summarizes the main terms associated with the invocation of the Moscow Mechanism.

One or more participating States may also request that another participating State invite a mission of experts "to address a particular, clearly defined

question on its territory relating to the human dimension" (para. 8). If the other state agrees, the mission of experts is established according to the same procedure as self-invocation, which again falls under the co-operative category. In such cases, the inviting state selects the experts who will take part in the mission, which in practice is done in consultation with the initiating state(s). The report must be provided within three weeks. When a situation requiring investigation arises in the territory of another state and no invitation is issued, however, this is deemed a refusal to co-operate, and the approach thus falls under the "contentious" category. In most such cases in the past, the states to be investigated recognized that they had a duty to provide information according to the Vienna Mechanism but chose not to co-operate with the Moscow Mechanism procedure.

For contentious cases, the rules provide that the requesting state, with the support of at least five other participating States, may initiate a mission of up to three rapporteurs to investigate the facts and give advice on possible solutions (paras 9–11). Again, the expectation is that the Vienna Mechanism will have already been applied. The report must be submitted within two weeks following the appointment of the last rapporteur. In principle, the requesting states and the requested state may each appoint one rapporteur from the resource list, and the two should agree on a third, forming a joint mission. Should the requested state fail to co-operate and to appoint its rapporteur within the six-day deadline following notification by the first rapporteur, however, the expert appointed by the requesting states must submit the report as a single rapporteur. The experts selected must not be nationals of the requesting or the requested state.

As a fast-track procedure, if a participating State requests an investigation of "a particularly serious threat to the fulfilment of the provisions of the CSCE human dimension" in another participating State, it can, with the support of at least nine other participating States, request an expert mission as described above without first resorting to the Vienna Mechanism (para. 12). A mission of experts may also be by established the OSCE Permanent Council upon the request of any participating State (para. 13). This option has never been applied, mainly because in such cases consensus would be required, which is unlikely. The main advantage of the Moscow Mechanism is that, except in such a case, no consensus is required, and the Mechanism cannot be blocked at any point.[13]

In order to avoid disputes on the selection of experts, the Moscow Mechanism provides for the establishment of a resource list or roster of experts, which is managed by the OSCE Office for Democratic Institutions and Human Rights (ODIHR). For this purpose, each participating State may appoint up to six experts who are eligible to serve for one or two mandates of three years each. No particular qualifications are required. Other participating States may voice reservations about up to two experts, in response to which the appointing state may either make other appointments or insist on its appointments, in which case

the appointed experts cannot take part in missions related to the state that voiced the reservation. In order to be operational, at least forty-five experts must be appointed to the roster (para. 3).[14] In the case of self-invocation or the invitation of a mission of experts upon request, the three experts are selected by the inviting state; in contentious cases, the first expert or rapporteur is selected by the invoking state(s). In the event of a lack of co-operation, he or she may remain a single expert/rapporteur.

The terms of reference are determined by the requesting and/or inviting state(s). In the case of self-invocation, paragraph 5 of the Moscow Mechanism provides that the state concerned will agree with the mission on the precise terms of reference, which may include fact-finding and advisory services to facilitate the observance of OSCE commitments. In practice, the experts play no role in defining the mandate, although they do have some discretion in interpreting it in light of feasibility considerations (for example, they may limit themselves to what they consider possible in view of time and resources). The purpose is indicated as facilitating the resolution of a particular question or problem related to the human dimension. If invited, the mission may even use its good offices and mediation services to promote dialogue and co-operation among the interested parties. In contentious cases, the establishment of facts, proposals, and advice on possible solutions is expected (para. 11). Accordingly, the report should also include a number of recommendations.

The cost of the mission is covered by the requesting states (para. 14), which usually distribute the costs among themselves. This includes operative costs for services provided to the experts by ODIHR, such as travel, translation, and light editing, while staff costs for administrative and logistical support must be covered by ODIHR. ODIHR also provides the experts with a list of useful contacts and establishes a mailbox through which they can receive relevant information. Neither ODIHR nor the OSCE in general provides substantive support to the experts, however, as this is not their role. For their work, the experts receive a lump sum from which they are to cover the costs of personal assistants, whom they are free to hire.

While the required co-operation of an inviting state is usually not a problem, when the process is contentious the requested state cannot be forced to co-operate. The Moscow Mechanism only provides that the participating States must refrain from taking reprisals against persons, organizations, or institutions who make contact with or submit information to the experts. Only the inviting state must provide the mission with state officials to accompany it, facilitate its work, and guarantee its safety (para. 6).

Regarding the drafting of the report, it is written by the experts themselves, and ODIHR only assists with light editing. In the case of self-invocation, the report is first shared with the invoking state, which has two weeks to provide its own comments on it, which it can add to the report. In contentious cases, the report is first shared with the requested

state, which has two weeks to provide its own observations, should there be any. The report must then be placed on the agenda of the next Permanent Council to be discussed. There is no need for a formal adoption, which would be difficult given the consensus requirement on all Permanent Council decisions. In practice, the report is generally published on the OSCE website immediately following the discussion and is thus made publicly available.[15]

Vienna Mechanism	Obligation of participating States to provide written information on a human dimension issue upon the request of other participating States within ten days and to engage in a bilateral dialogue within one week.
Moscow Mechanism	Right of a participating State to invite an expert mission to facilitate the resolution of questions related to the human dimension on its own territory or of a certain number of invoking states to send an expert mission to address a particular question regarding, or a serious threat to, the fulfillment of the human dimension provisions on the territory of another participating State.
Requesting (or invoking) state(s)	Participating State(s) that invoke(s) the Mechanism; possibility of self-invocation for the resolution of questions in a participating State's own territory.
Requested state	State subject to the invocation of the Mechanism.
Co-operative approach	Mission of experts is established and undertaken with the co-operation of the requested state.
Contentious approach	Mission of experts is established and undertaken without the co-operation of the requested state.
Rapporteur(s)	Expert(s) who serve(s) on the mission to facilitate the resolution of a human dimension issue through a fact-finding report and advisory services or to investigate a particular question or a particularly serious threat related to the human dimension and who produce(s) a report with recommendations.
Resource list	Roster of experts, nominated by participating States, from which experts can be chosen to serve on a Moscow Mechanism mission.
Terms of reference	Mandate of the expert missions, to be defined by the requesting state(s).
Deadlines	Strict timelines regulating the composition of the expert missions, the delivery of the reports, and the opportunity to comment on them.

Table 1. Definitions of major terms.

The practice

Application of the Vienna Mechanism

The Vienna Mechanism may be employed on its own or as a first step toward the use of the Moscow Mechanism (para. 8). In 1989, for example, the Vienna Mechanism was used by sixteen countries to inquire into the arrest of the playwright Vaclav Havel.[16] In the case of the Chechen Republic, the Vienna Mechanism was used by the invoking states first. Unsatisfied with the results,

they subsequently invoked the Moscow Mechanism. The Vienna Mechanism was invoked in November 2021 by thirty-five participating States to gain information on the implementation of the recommendations made by the Moscow Mechanism rapporteur to Belarus.[17] While the Russian Federation and Belarus have sought to justify their non-cooperation by arguing that the Moscow Mechanism is outdated and obsolete, they claim to recognize the Vienna Mechanism and, at least in principle, co-operated with it in the above examples. However, in the case of the invocation of the Vienna Mechanism by forty-one participating States in March 2024 as a follow-up to the Moscow Mechanism report on alleged human rights violations in the Russian Federation in 2022, Russia refused to respond to the questions asked.[18]

Application of the Moscow Mechanism: Cases

According to a list maintained by the OSCE, the Moscow Mechanism has been invoked fifteen times thus far. This list also contains an invocation by the Russian Federation in the case of NATO strikes on Yugoslavia in 1999, for which no report is available (this despite the fact that other sources only consider it an invocation of the Vienna Mechanism).[19] Among the four other cases from the 1990s, two were related to the war in the former Yugoslavia. They were requested by twelve members of the European Community and the United States and concerned reports on atrocities and attacks on unarmed civilians in Croatia and Bosnia-Herzegovina. The request resulted in a report on Croatia alone, as the mission could not be sent to Bosnia-Herzegovina for security reasons. A follow-up mission in 1993 at the request of the Ministers of Foreign Affairs of CSCE participating States led to a proposal for the establishment of an International War Crimes Tribunal for the Former Yugoslavia and thus contributed to its later creation by the UN Security Council. Also in 1993, the CSCE Committee of Senior Officials established a mission to investigate human rights violations in Serbia and Montenegro, which, however, was unable to deliver due to a lack of co-operation on the part of the Federal Republic of Yugoslavia.[20]

The list includes two self-invocations: in 1992, Estonia requested a review of the conformity of Estonian legislation on citizenship with universal human rights norms, and in 1993 Moldova requested an examination of its legislation and policies regarding the implementation of minority rights.

In the case of Turkmenistan in 2003, ten OSCE participating States requested a report on the November 2002 attack on Turkmen President Saparmurat Niyazov and related investigations. In 2011, fourteen OSCE participating States invoked the Mechanism with regard to human rights violations following the presidential elections in Belarus of December 19, 2010. In both cases, the country under investigation did not co-operate, but the rapporteur was able to produce a substantive report based on multiple sources,

which was discussed in the Permanent Council.

After another period of non-use, the Moscow Mechanism was invoked again in 2018 by sixteen OSCE participating States to investigate alleged human rights violations, mainly against LGBTQ+ people in the Chechen Republic of the Russian Federation. In 2020, seventeen participating States invoked the Mechanism against Belarus under paragraph 12 to examine alleged human rights violations related to the presidential elections of August 9, 2020. In these cases as well, the requested states refused to co-operate. As a result, there was no opportunity to form a commission of experts, and the single rapporteur had to provide a report to the Permanent Council within the two-week deadline.[21]

Since 2020, the popularity of the Moscow Mechanism has increased, leading to a growing number of cases. In 2022, following consultation with Ukraine, forty-five OSCE participating States invoked the Mechanism under paragraph 8 to investigate "the human rights and humanitarian impacts of the Russian Federation's invasion and acts of war, supported by Belarus, on the people of Ukraine, within Ukraine's internationally recognized borders and territorial waters."[22] A commission of three experts was established by Ukraine, which presented its report on alleged violations of international humanitarian and human rights law, war crimes, and crimes against humanity committed in Ukraine since February 24, 2022, within the three-week deadline. There was no co-operation from the Russian Federation, although it was invited to share information in accordance with paragraph 6 of the Moscow Mechanism. As the inviting state, however, Ukraine pledged full co-operation. It also made use of its right to attach its comments to the report (para. 7). Due to the urgency of the matter, the commission of experts presented its report at a special meeting of the Permanent Council convened by the Polish Chairpersonship on April 13, 2022.[23] In the debate, only Russia and Belarus criticized the report. Because of the report's narrow deadline, it could not investigate the atrocities and other human rights violations committed by Russian soldiers in Bucha and other locations (such as Hostomel) in any depth. Therefore, the same states triggered a follow-up report under the Moscow Mechanism, which was delivered by a different commission in July 2022.[24]

From July 2022 to February 2024, the Moscow Mechanism was invoked four more times: in July 2022 by thirty-eight participating States (under para. 12) on alleged human rights violations in the Russian Federation; in March 2023 by thirty-eight participating States to examine human rights violations and abuses in Belarus; in March 2023 by forty-five participating States following consultation with Ukraine on the forcible transfer of children from occupied Ukrainian territories and their deportation to the Russian Federation;[25] and in February 2024 by forty-five participating States following consultation with Ukraine on the arbitrary deprivation of liberty of Ukrainian civilians by the Russian Federation.[26] Consequently, the Moscow Mechanism has already been used four times

to investigate allegations of violations of human dimension commitments related to the Russian invasion of Ukraine. In both the implementation of the Moscow Mechanism and the relevant follow-up processes, but also with regard to reflections on how to strengthen it, civil society organizations have played an important role.[27]

Application of the Moscow Mechanism: Main issues

The application of the Moscow Mechanism raises several practical issues. The mandate, as indicated in the terms of reference, is usually too broad to be fully covered. Agreement is more likely when the mandate is broad, covering the concerns of all invoking states; nevertheless, the purpose of the Moscow Mechanism is to facilitate the resolution of a particular question or problem (para. 5) or of a particular, clearly defined question (para. 8). Only in the case of the fast-track or emergency mode, when a particularly serious threat to the provisions of the human dimension is at issue (para. 12), is a wider approach foreseen. In practice, not least because of the narrow deadlines, the experts are free to write their report in a way that allows for the mandate to be met in its main substance. For example, the report on the mission carried out in 2022 to investigate alleged human rights violations in the Russian Federation, which was given a very broad mandate by the invoking states, limited its scope to assessing Russia's legal and administrative practice in light of its OSCE human dimension commitments.[28]

The methodology used by the experts is crucial to reaching results within tight deadlines. This requires co-operation with trusted local and international human rights nongovernmental organizations (NGOs) and gaining access to victims and witnesses. Although human rights fact-finding methodologies have evolved significantly in recent years,[29] thanks in part to the availability of online open-source information and the use of geolocation and satellite imagery,[30] there are obvious limits to what can be done by the experts of the Moscow Mechanism within the given time and resource constraints. However, besides their own investigations, they may be able to draw on interviews conducted and analytical reports produced by local and international NGOs. All this material, as well as reports from investigative media, need to be cross-checked with other sources. These sources can be diverse, including interviews conducted by the mission as well as reports and material from international organizations (such as the United Nations and the Council of Europe) and university research teams. With an eye to transparency and credibility, it is important to indicate the sources in the methodology section of the reports, albeit in a way that does not put anyone at risk.

The tight deadlines are an obvious challenge for any serious report. They may be explained by the original purpose of addressing "a particular, clearly defined question" (para. 8). In co-operative cases, the deadline can be prolonged, if necessary, as paragraph 7 indicates a

deadline of "preferably" three weeks. In practice, however, this has been avoided as far as possible, as an important advantage of the Moscow Mechanism is that it provides quick results. Ideally, the experts envisaged will have been alerted to their task before the formal decision on the invocation is taken, providing them with extra time, yet in concrete cases how well the experts perform their task within the limited time will depend on their expertise and network.

The situation regarding the experts is aggravated by the fact that, unlike most other international missions, the organization hosting the experts does not, as a matter of principle, provide substantive input, as ODIHR (and the OSCE in general) does not see this as their role and has no budget for such input. While the strong commitment of ODIHR/OSCE staff to assisting the experts logistically must be recognized, the rules of the Moscow Mechanism do not prohibit the provision of more substantive support for experts, and there is no reason to think that such support would jeopardize their full independence. This could take the form of a focal point which assists the rapporteurs in pinpointing relevant information. Experts also benefit from information received from other international organizations, such as the United Nations and the Council of Europe. Certainly, the short timelines make any co-operation on issues of substance difficult; nevertheless, providing access to existing knowledge within OSCE executive structures should be possible.[31]

Outcomes of the reports under the Moscow Mechanism

Following their presentation and discussion in the OSCE Permanent Council, the reports are published on the OSCE website in English, and where appropriate also in Russian and the local language of the relevant state (for example Ukrainian). Because of the consensus requirement, which gives de facto veto power to each participating State, it is nearly impossible to agree on common OSCE follow-up activities. However, this does not mean that the implementation of further activities is impossible. In practice, based on the reports under the Moscow Mechanism, side events have taken place at the subsequent annual OSCE Ministerial Conferences, and the reports were also discussed at the Warsaw Human Dimension Conferences in 2022 and 2023, which were held despite Russia's blocking of the annual Human Dimension Implementation Meeting. As noted above in the case of Belarus and the Russian Federation, the Vienna Mechanism has been invoked as a follow-up mechanism for inquiring into whether the recommendations of the report were taken up. As another type of follow-up, the Moscow Mechanism was invoked a second time to investigate repression and political detentions in Belarus since the first report of October 2020.[32] In the case of the reports on Ukraine, it is worth noting that since June 2022, ODIHR has published semiannual reports on violations of international humanitarian law and human rights in Ukraine. This has been made possible through an extrabudgetary fund

for rapid monitoring missions which has supported other missions in the past. The monitoring, which began right after the military attack on Ukraine, has been stepped up with the deployment of monitors on the ground since May 2022. These regular reports could also serve as follow-up for the ad hoc missions carried out under the Moscow Mechanism.[33]

The reports under the Moscow Mechanism have an even broader set of uses, however. As the (co-)author of three reports, for example, I have been invited to present reports at hearings in the US Senate, informal meetings of the Political and Security Committee of the European Union, Arria formula meetings of the UN Security Council and side events of the UN General Assembly, and various pertinent academic and other conferences, in addition to responding to numerous media requests. The purpose of this engagement is to share the results contained in the reports, which may be taken into account in the political and legal decisions of these organizations and institutions. In all these activities, the rapporteurs are free to accept or decline invitations and in how to present their report. However, they may only speak about their findings following the publication of the report. When accepting their mandate, rapporteurs may not be fully aware of this part of their role, which is not regulated in any way.

Finally, the reports are widely read and used by a variety of actors, including local and international NGOs, whose work the reports both confirm and encourage and who can also draw on the reports in their consultations with policymakers. The Council of Europe and the Human Rights Council have acknowledged the reports in their own work. As an example of best practice, the establishment of the International Accountability Platform for Belarus (IAPB) has served as a follow-up to the report on human rights violations related to the presidential elections of 2020. It is based on a joint declaration by nineteen states, seventeen of which had already invoked the Moscow Mechanism in the case of Belarus, and was also supported by the European Union.[34] The IAPB was founded in response to a recommendation made in the report on Belarus to ensure accountability for human rights violations and to prevent a culture of impunity. It was formed as a coalition of independent international and Belarusian NGOs with the purpose of "collect[ing], consolidat[ing], verify[ing] and preserv[ing] evidence of gross human rights violations constituting crimes under international law."[35] It is led by the Danish Institute against Torture (DIGNITY), the Viasna Human Rights Centre, the International Committee for the Investigation of Torture in Belarus, and REDRESS, and it co-operates with additional international and local human rights NGOs on its advisory council. Its professional legal and medical staff has experience with criminal investigations and prosecutions and with a victim- or survivor-centered approach. It may also share its findings with the Office of the United Nations High Commissioner for Human Rights in its examination of the human rights situation in Belarus and with national prosecution authorities.

The particularities of the Moscow Mechanism

Applying the Moscow Mechanism comes with both advantages and challenges. Among the advantages is the fact that the Moscow Mechanism is relatively easy to invoke, ensures a fast procedure with quick results, cannot be obstructed, and is very flexible in its implementation. In addition, the operational costs of the missions are mainly covered by the invoking states, and the report is swiftly discussed in the Permanent Council and published on the OSCE website. Importantly, the speed with which the procedure is carried out also signals to victims and human rights defenders that their situation will be given the necessary attention.

Among the challenges are the often overly broad mandates, the very narrow deadlines, the limited resources, the lack of experienced staff, the frequent lack of co-operation, and the lack of regulations regarding the protection of witnesses and evidence. Regarding the selection of rapporteurs, more information should be provided on their expertise, although the invoking states certainly examine the pool closely before choosing an expert for a mission. Beyond the report itself, there is no other record of the collection of evidence relied on by the rapporteurs. There are no specific security arrangements for the rapporteurs and no rules (and only limited guidance) governing the activities of the rapporteurs following the completion of their missions. The ad hoc nature of the investigation only allows for an assessment of the situation at a given time. Finally, there is no established monitoring procedure regarding the implementation of the reports' recommendations.

Recommendations

Narrowing the mandate of the missions. The mandates under the Moscow Mechanism are generally too broad. It would be preferable to be more specific, so as not to raise unrealistic expectations. The possibility foreseen in the rules of the Moscow Mechanism to the effect that the state concerned "will agree with the mission on the precise terms of reference" (para. 5) has yet to be put into practice but could be in the future.

Implementing a thorough expert selection process. In view of the highly demanding task carried out by the experts/rapporteurs, their selection should take their experience and networks, as well as their ability to present the results following the missions, into account.

Supporting the experts. The experts should be well briefed on their role and on the support available from the invoking/requesting states and ODIHR, regarding both their mission and possible follow-up activities. Meetings with former experts could be organized by ODIHR to share pertinent experience. Relevant knowledge gleaned by OSCE structures should also be shared.

Improving co-ordination among experts. In the case of missions comprised of three experts, there is a need for co-ordination regarding both the sharing of tasks and follow-up activities such as media engagements. ODIHR could assist in this, but

in the end, it is the responsibility of the three experts to ensure a consistent approach to responding to requests. For this purpose, the experts should co-ordinate their activities using safe channels of communication.

Engaging in more structured follow-up activities. Follow-up activities ought to be made more structured, for example by holding regular debriefings and debates on the implementation of the recommendations, by using either the Vienna Mechanism or the regular meetings of the Permanent Council. The practice of holding side events on the reports at ministerial meetings and the Human Dimension Conferences should be continued and could be extended to involve the OSCE Human Dimension Committee. The International Accountability Platform for Belarus offers an example of how to institutionalize a professional follow-up mechanism, although it was organized outside the auspices of ODIHR and the OSCE for reasons of ensuring its independence, but also in view of ODIHR's limited engagement.

Finally, in view of the recent increase in Moscow Mechanism missions, ODIHR and interested participating States could arrange meetings of former experts to discuss best practices and consult on ways to strengthen the Moscow Mechanism.

Notes

1 See OSCE ODIHR, *OSCE Human Dimension Commitments*, vol. 1, *Thematic Compilation*, 4th ed. (Warsaw: OSCE Office for Democratic Institutions and Human Rights, 2022), https://www.osce.org/odihr/human-dimension-commitments-thematic; OSCE ODIHR, *OSCE Human Dimension Commitments*, vol. 2, *Chronological Compilation*, 4th ed. (Warsaw: OSCE Office for Democratic Institutions and Human Rights, 2022), https://www.osce.org/odihr/human-dimension-commitments-chronological

2 See CSCE, Charter of Paris for a New Europe (Paris: November 21, 1990), https://www.osce.org/mc/39516

3 For an overview, see OSCE, "Human Dimension Mechanisms," https://www.osce.org/odihr/human-dimension-mechanisms

4 See the responses received by the rapporteur to his requests for co-operation by the representatives of the Russian Federation and of Belarus, annexed to the reports related to Chechnya (2018) and Belarus (2020). The reports can be accessed at: OSCE, cited above (Note 3).

5 The author has had the opportunity to serve as a single expert for the reports on Chechnya (2018) and on Belarus (2020) and as a member of the first Commission of Experts on Ukraine (April 2022).

6 CSCE, Concluding Document of the Vienna Meeting 1986 of Representatives of the Participating States of the Conference on Security and Co-operation in Europe, Held on the Basis of the Provisions of the Final Act Relating to the Follow-up to the Conference (Vienna: January 19, 1989), https://www.osce.org/mc/40881

7 See CSCE, Document of the Copenhagen Meeting of the Conference on the Human Dimension of the CSCE (Copenhagen: June 29, 1990), paragraph 42, https://www.osce.org/odihr/elections/14304

8 The description on the OSCE website does not contain the Moscow amendments. See CSCE, Vienna Mechanism (Vienna: December 1, 1989), https://www.osce.org/odihr/20064

9 See CSCE, Document of the Moscow Meeting of the Conference on the Human Dimension of the CSCE (Moscow: October 3, 1991), https://www.osce.org/odihr/elections/14310

10 See CSCE, Moscow Mechanism (Moscow: December 1, 1991), https://www.osce.org/odihr/20066

11 Among the few descriptions available, the US Helsinki Commission Report Brief on the Moscow Mechanism of July 18, 2017, stands out. See US Helsinki Commission, Helsinki Commission Report: The OSCE Moscow Mechanism (July 18, 2017), https://www.csce.gov/wp-content/uploads/2017/07/Report-Moscow-Mechanism-FINAL.pdf

12 Here and below, all paragraph references are to the Moscow Mechanism, cited above (Note 10).

13 On the issue of consensus in the OSCE, see Rick Fawn, "The Limits and Potential of Consensus in Times of Crisis," in *OSCE Insights*, eds. Cornelius Friesendorf and Argyro Kartsonaki (Baden-Baden: Nomos, 2024), https://www.nomos-elibrary.de/10.5771/9783748917366-04/the-limits-and-potential-of-consensus-in-times-of-crisis?page=1

14 For a list of the experts, see OSCE, List of Experts for the Human Dimension Mechanism Appointed by OSCE Participating States (April 8, 2024), https://www.osce.org/odihr/20062

15 See OSCE, cited above (Note 3).

16 See US Helsinki Commission, cited above (Note 11).

17 See the joint letter of the thirty-five participating States to Belarus at: https://assets.publishing.service.gov.uk/media/6186bdbbd3bf7f56080b1d32/20211104_Joint_letter_to_Belarus__Vienna_Mechanism.pdf. See also the statement on the response by Belarus at: https://www.gov.uk/government/speeches/response-by-belarus-to-the-vienna-mechanism-joint-statement

18 See the invocation at: https://delegazioneosce.esteri.it/wp-content/uploads/2024/03/Vienna-Mechanism-statement_22032024.pdf

19 US Helsinki Commission, cited above (Note 11).

20 US Helsinki Commission, cited above (Note 11).

21 See Wolfgang Benedek, "OSCE Moscow Mechanism: Situation of Human Rights in Chechnya," in *European Yearbook on Human Rights 2019*, eds. Philip Czech et al. (Cambridge: Intersentia, 2020), 419–38; Wolfgang Benedek, "Accountability for Grave Human Rights Violations in Belarus: Case Study on the OSCE Moscow Mechanism," *Defence Horizon Journal* Special Edition (May 2021): 32–35.

22 See OSCE, cited above (Note 3).

23 See Wolfgang Benedek, Veronika Bilkova, and Marco Sassòli, "Violations of International Humanitarian and Human Rights Law, War Crimes and Crimes Against Humanity Committed in Ukraine since 24 February 2022: Summary of the Report by a Mission of Experts under the OSCE Moscow Mechanism," ESIL Talk!, April 19, 2022, https://www.ejiltalk.org/violations-of-international-humanitarian-and-human-rights-law-war-crimes-and-crimes-against-humanity-committed-in-ukraine-since-24-february-2022-summary-of-the-report-by-a-mission-of-experts-under-t/; Marco Sassòli, "Un premier rapport sur les violations commises pendant la guerre en Ukraine," *Jusletter*, June 20, 2022.

24 See Vasilka Sancin, "The Role and Impact of the OSCE Moscow Mechanism Reports Following the Russian Invasion of Ukraine," *International and Comparative Law Review* 23, no. 1 (2023): 210–27.

25 See OSCE, cited above (Note 3).

26 See OSCE, "Ukraine Appoints Three Experts Following Invocation of the OSCE's Moscow Mechanism," March 15,

2024, https://www.osce.org/odihr/564851; OSCE, Report on Violations and Abuses of International Humanitarian and Human Rights Law, War Crimes and Crimes against Humanity, related to the Arbitrary Deprivation of Liberty of Ukrainian Civilians by the Russian Federation, prepared by Veronika Bilkova, Cecilie Hellestveit, and Elina Steinerte, ODIHR.GAL/19/24/Corr.2* (April 25, 2024), https://www.osce.org/files/f/documents/f/4/567367.pdf

27 See, for example, The Netherlands Helsinki Committee and the Civic Solidarity Platform, Strengthening OSCE Instruments in the Human Dimension (The Netherlands Helsinki Committee, 2019), https://www.civicsolidarity.org/sites/default/files/strengthening_the_use_of_osce_human_dimension_instruments_csp_report_2019.pdf

28 See OSCE, Report on Russia's Legal and Administrative Practice in Light of Its OSCE Human Dimension Commitments, prepared by Angelika Nußberger, ODIHR.GAL/58/22/Rev.1 (September 22, 2022), https://www.osce.org/files/f/documents/7/5/526720_0.pdf

29 Wolfgang Benedek, "International Fact-Finding on Human Rights Violations: The Moscow Mechanism of OSCE," in *Der Schutz des Individuums durch das Recht, Festschrift für Rainer Hofmann zum 70. Geburtstag,* eds. Philipp B. Donath et al. (Berlin: Springer, 2023), 129–41.

30 Eliot Higgins, *We Are Bellingcat: An Intelligence Agency for the People* (London: Bloomsbury, 2021).

31 See Wolfgang Benedek, The Use of the OSCE Moscow Mechanism and Its Potential, Graz Law Working Paper Series No. 22 (Graz: University of Graz Faculty of Law, 2021), 1–7, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3972128

32 See OSCE, Report on the Serious Threat to the OSCE Human Dimension in Belarus since 5 November 2020, prepared by Hervé Ascensio, ODIHR.GAL/39/23/Corr.1* (May 11, 2023), https://www.osce.org/files/f/documents/d/5/543240_0.pdf

33 See all interim reports on reported violations of international humanitarian law and human rights in Ukraine available thus far at: OSCE, "Interim Reports on Reported Violations of International Humanitarian Law and International Human Rights Law in Ukraine," https://www.osce.org/odihr/537287

34 See US Department of State, "Joint Statement by 19 States in Support of the Establishment of the International Accountability Platform for Belarus," March 24, 2021, https://www.state.gov/joint-statement-by-19-states-in-support-of-the-establishment-of-the-international-accountability-platform-for-belarus/

35 For more details on the work of the IAPB, see https://iapbelarus.org