

EVALUATION REPORT

Final Formative Decentralised Evaluation

RFoM's Extra-budgetary project "The Impact
of Artificial Intelligence on Freedom of Expression" (SAIFE)

January 2024

Authors: Diana Mirza-Grisco & Ben Greenacre

Acknowledgements

This report was authored by Diana Mirza-Grisco and Ben Greenacre, under the guidance and with the support of Deniz Wagner, SAIFE project manager and Julia Haas, SAIFE project officer.

The authors would like to thank all key informants for their time and insights.

Executive Summary

Responding to significant interest from a wide range of stakeholders on the topic of the impact of AI on the information environment, the OSCE RFoM implemented the 'Spotlight on Artificial Intelligence and Freedom of Expression' (SAIFE) project between July 2019 - December 2023. The project focused on the impact of AI on freedom of media, opinion and expression split across four distinctive thematic areas: security, hate speech, media pluralism, and surveillance. The goal of the project was at first, to frame a complex emerging technology and its impacts in human rights terms. Secondly, the project sought to catalyse human rights-based law and policy regulation of AI systems to ensure a fair, inclusive, and transparent information space.

The objective of the following decentralised evaluation was to provide a final assessment of the project's results and impact, as well as to identify lessons learned and concrete recommendations to increase the impact of future RFoM activities and projects pertaining to technology, digital rights and online free speech. The exercise was guided by the OSCE Evaluation Policy of 14 November 2022 and the OECD Development Assistance Committee (DAC) evaluation revised criteria of relevance, coherence, effectiveness, efficiency, impact, sustainability, inclusion and equity. This evaluation had a mixed – methods approach to data collection, and was informed by a desk review of over 80 documents, interviews with 26 stakeholders, and an online survey that generated 42 responses across different categories of stakeholders.

Based on the different lines of evidence, the evaluation made several major findings. They are summarised below, structured according to the evaluation criteria.

Relevance

Project stakeholders consider the project highly relevant for raising awareness about AI's impact on FoE and FoM, as well as for the provision of recommendations and guidance for pS on how to develop and enforce policy and regulation on AI's impact on freedom of expression.

Civil society & participating States (pS) took differing views to the non-inclusion of tech companies in SAIFE, but both groups agreed they would like to see industry participation in next steps.

Moving forward, most participants felt that the 4 key thematic areas identified by SAIFE remain relevant. The only novel use of AI technologies that was consistently raised for further investigation was 'generative AI' (*on which as of the time of drafting, the RFoM team were issuing a publication*). Many interlocutors emphasised that follow-up and implementation must be a focus of any subsequent project.

Coherence

The growing interest of OSCE entities in AI as a cross-cutting comprehensive security issue relevant for all units/dimensions contributed to the creation of internal OSCE synergies. The close collaboration with the field operations supported the achievement of project results. In addition to being first on the scene, the project was unique in its modalities and outputs, due, in part, to the geographic focus of the OSCE, the unique 'comprehensive security' mandate of the OSCE, and the specificities of the RFoM mandate.

SAIFE was one of the earliest regional and international initiatives to engage with the human rights impact of AI. This enabled the project to set the terms for future discussions in international fora. In doing so, both States and civil society stakeholders appreciated that SAIFE established a practical, and policy-oriented framework for discussion. Representatives of International Organisations noted that early and on-going collaboration with SAIFE directly influenced the development of their own programmes on AI and human rights.

Effectiveness

The majority of the survey participants feel that their understanding on AI impact on freedom of expression, and freedom of media was enhanced to a very large extent as a result of their participation in the SAIFE project. Additionally, the findings show that the project was very effective in raising awareness on AI impact on media freedom and pluralism, as well as on the discrimination and gender divides.

The policy manual is considered very useful. Recommendations are generally actionable and detailed, but should be tailored to specific actors in future projects. Respondents also consider that all the other project outputs were also helpful, especially where they were tailored to local contexts, for example through translation into local languages.

Additionally, SAIFE successfully brought together more than 120 experts across various disciplines and geographical contexts, and managed to create coalitions of experts worldwide interested in this topic.

Efficiency

Overall, the project followed its implementation plan, and accommodated the tech evolutions and needs of the beneficiaries. Formal processes did not influence majorly the timeliness of the project implementation. However, on occasion, they have taken time away from substantive work. Uncertainty linked to staff fluctuation in the office and in the project team had minor impact on efficiency.

Less formalized M&E processes were employed throughout the project. There were no formal M&E tools developed for the project, but this did not represent a major hindrance. Key Informant Interviews (KIIs) consider that the informal M&E approach fit well this project. However, in a different setting a more structured approach would have been beneficial.

Impact

State actors report significant use of SAIFE outputs in a number of ways including in briefing parliamentarians and senior civil servants, framing internal and external advocacy and consultations or educational purposes. The project also contributed to the enhancement of the collaboration between non-state and state stakeholders on AI policy development. At least one International Organisation reported that SAIFE directly influenced the development of their own work programme on AI. Stakeholders have also reported on building up on SAIFE work by conducting studies, analyses, and using SAIFE outputs for teaching purposes.

Sustainability

SAIFE sustainability plan was based on the project's outputs sustainability potential, including mainly the manual, the resource hub and a capacity building module (masterclasses). At the same time, the project

established connections and networks between experts at regional and global level which will ensure the project outcomes are furthered and built on.

Through the advocacy work of RFoM (e.g. regional and national visits), the project outputs will potentially reach a wide range of national stakeholders. KIIIs believe that the RFoM is best placed to combine different perspectives. Many stakeholders consider that a continuation of the project is necessary to strengthen the sustainability of the policy oriented SAIFE outputs.

Inclusion and Equity

Gender mainstreaming was maintained throughout the project's activities, and analyses did not isolate specific aspects of identity, but rather looked holistically at how different aspects of one's identity may intersect and modify experiences of discrimination (i.e the team applied an 'intersectional' analysis). However, some participants feel that future iterations of the project could further deepen the intersectional analysis, and further increase the outputs relevance for specific communities. SAIFE captured the perspectives from underrepresented regions, however, participants felt that more work could have been done to broaden geographic participation. In addition to tech companies, interviewees made several recommendations for inclusion in future SAIFE consultations.