Contribution

The Impact of Artificial Intelligence on the Work of Human Rights Defenders

Manuel Brunner[1]

[1]Prof. Dr., LL.M., HSPV NRW

Contents

Abstract

Artificial Intelligence (AI) is one of the most important technologies of the present day. It is already influencing the day-to-day life of many people and is shaping the decision-making of enterprises, governments, humanitarian organisations and other actors. This article discusses the potentials of the use of AI by and against human rights defenders. On the one hand, AI can offer immense benefits for such individuals or groups with its timesaving, simplifying and analytical capabilities. On the other hand, AI can be used to track and attack human rights defenders and their activities. It offers an additional way, especially for governments, to curtail the rights of human rights defenders and to hinder them in their work.

Keywords

I. Introduction[*]

Current discussions about artificial intelligence (AI) often revolve around questions of the impact of this capability of computational systems on the future of work or on the dangers AI poses for the enjoyment of human rights.[1] However, a rarely discussed point in those debates is the impact that AI has on the work of those individuals or organisations who endeavour to improve the human rights situation for others, namely the impact on human rights defenders. On the one hand, AI can be a powerful tool for human rights defenders supporting them in fields like data analysis or campaigning or can be used to provide services for people who face situations which are problematic in terms of the enjoyment of human rights. Moreover, AI might be used to visualize human rights violations and AI-driven application can serve as warning tools when human rights defenders are under threat. On the other hand, states might use AI to persecute and curtail the rights of those actors. While the field of AI and its risks to human rights have already been discussed by some Special Rapporteurs of the United Nations Human Rights Council regarding their respective mandates, an articulation by the Special Rapporteur on the situation of human rights defenders is yet missing.[2] For the purpose of the following, the term "artificial intelligence" is understood along the lines of the definition of AI systems in Article 2 of the Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law (Framework Convention).[3] Therefore, AI "means a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations or decisions that may influence physical or virtual environments. Different artificial intelligence systems vary in their levels of autonomy and adaptiveness after deployment."

This article discusses the various aspects of the impact of AI on the work of human rights defenders as outlined above. To add context to the following descriptions and analyses the explanations start with a definition of the term "human rights defender" and an introduction to the legal framework under which those actors operate. In the following section AI is discussed as a useful tool in the hands of human rights

---

[*] This paper was presented at the 30th Anniversary Conference of the Human Rights Centre of the University of Potsdam "Human Rights and Artificial Intelligence Addressing challenges, enabling rights", 7–8th November 2024 in Potsdam, Germany.

[1] See for instance *Thomas H. Davenport*, The AI Advantage: How to Put the Artificial Intelligence Revolution to Work, 2018; International Monetary Fund (ed.), Gen-AI: Artificial Intelligence and the Future of Work: Staff Discussion Note, 2024; *Jean-Philippe Deranty/Thomas Corbin*, Artificial intelligence and work: a critical review of recent research from the social sciences, in: AI & Society 39 (2024), pp. 675–691; *Steven Livington/Mathias Risse*, The Future Impact of Artificial Intelligence on Humans and Human Rights, in: Ethics & International Affairs 33 (2019), pp. 141–158; *Rowena Rodrigues*, Legal and human rights issues of AI: Gaps, challenges and vulnerabilities, in: Journal of Responsible Technology 4 (2020), 100005; *Onur Bakiner*, The promises and challenges of addressing artificial intelligence with human rights, Big Data & Society 10 (2023).

[2] Reports about AI in the fields of their respective mandates have for instance been: UNHRC, Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression of 29 August 2018, UN Doc. A/73/348; UNHRC, Report of the Special Rapporteur on the right to education of 16 October 2024, UN Doc. A/79/520; UNHRC, Report of the Special Rapporteur in the field of cultural rights of 30 July 2025, UN Doc. A/80/278.

[3] Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law of 5 September 2024, CETS No. 225.

defenders. Several potential and actual applications are presented in that regard. Topics like AI-driven data analysis, exemplified by the work of non-governmental human rights organisations, and the creation of a warning tool for human rights defenders are discussed. The next section of the article explores the use of AI against human rights defenders. Examples from known practices of authoritarian governments are used to illustrate the dangers of AI-driven technologies for the work of human rights defenders. Thereafter, the challenges of current international legislation on AI regarding human rights defenders are discussed. The article ends with a conclusion.

II. Definition of and Legal Framework for Human Rights Defenders

The term "human rights defender" has gained prominence in human rights work, politics, international relations and legal studies for over a quarter century now.[4] Before that, individuals or groups who would today be labelled as human rights defenders were more regularly called human rights activists, professionals, workers or monitors.[5] While those terms have

not vanished, the new term is arguably seen and used as the more common one today. The origin of the described change in the usage of terms lies in the year 1998. After 14 years of negotiations, the General Assembly of the United Nations adopted the Declaration on the Right and Responsibility of Individuals, Groups and Organs of Society to Promote and Protect Universally Recognized Human Rights and Fundamental Freedoms on 9 December of that year.[6] This official title of the instrument is often shortened to "Declaration on human rights defenders".[7] The shortening serves practical purposes as it clearly underlines who benefits from its regulations. The term "human rights defender" is not defined in the declaration itself. However, in the practice of the United Nations human rights defenders are considered "people who, individually or with others, act to promote or protect human rights in a peaceful manner."[8]

The Declaration on human rights defenders is an international instrument of the right to defend human rights. It reaffirms rights that are of utmost importance for the defence of the rights of others.[9] Those rights include the freedom of association, the freedom to peaceful assembly, the freedom of opinion and expression, the right to gain access to information, the right to provide legal aid and the right to develop and discuss new ideas in the area of human rights.[10] The implementation of the Declaration on human rights de-

---

4   See for instance the contributions in *Alice M. Nah (ed.)*, Protecting Human Rights Defenders at Risk, 2020; *Benjamin Beuerle*, Zur Umsetzung der „Erklärung zu den Menschenrechtsverteidigern" fünf Jahre nach ihrer Verabschiedung – eine Bestandsaufnahme, in: MRM 9 (2004), pp. 47–52; *Norman Weiß*, Schutz von Menschenrechtsverteidigern – neuere Entwicklungen, in: MRM 21 (2016), pp. 29–44.

5   OHCHR, About human rights defenders, available at: https://www.ohchr.org/en/special-procedures/sr-human-rights-defenders/about-human-rights-defenders (last visited 15 November 2025).

---

6   UN Doc. A/RES/53/144.

7   For example UN, Commentary to the Declaration on the Right and Responsibility of Individuals, Groups and Organs of Society to Promote and Protect Universally Recognized Human Rights and Fundamental Freedoms of July 2011, p. 5.

8   OHCHR (fn. 5).

9   UN (fn. 9).

10 Ibid.

fenders is a precondition for the creation of an environment that enables human rights defenders to carry out their work.[11] As the Declaration on human rights defenders is a document that was adopted by the General Assembly of the United Nations it is not legally binding and is therefore a soft law document.[12] However, as individuals or groups, human rights defenders also benefit from the rights enshrined in treaties and other international instruments in the field of human rights such as the International Covenant on Civil and Political Rights[13] or the Universal Declaration of Human Rights[14] and from fundamental rights in national constitutional law. Furthermore, some states have introduced legislation or policies for the benefit of human rights defenders.[15]

---

[11] Ibid.; UN Doc. E/CN.4/2006/95, para. 50.

[12] See for a detailed analysis *Stephen M. Schwebel*, The Effect of Resolutions of the U.N. General Assembly on Customary International Law, in: Proceedings of the Annual Meeting of the American Society of International Law 73 (1979), pp. 301–309; *Ludovic Hennebel/Hélène Tigroudja*, International Human Rights Law: A Treatise, 2025, pp. 92 ff.

[13] International Covenant on Civil and Political Rights of 16 December 1966, UNTS vol. 999, p. 171

[14] Universal Declaration of Human Rights of 10 December 1948, UN Doc. A/RES/217 A (III).

[15] See for instance the following legislation: United Mexican States, Ley para la protección de personas defensoras de derechos humanos y periodistas of 25 June 2012, Diario Oficial de la Federación ID 270; Republic of Honduras, Ley de la protección para las y los defensores de derechos humanos, periodistas, communicadores sociales y operadores de justicia of 15 May 2015, Diario Oficial de la República de Honduras, Número 33, 730, A. 1–21 or Republic of Niger, Loi fixant les droits et devoirs des défenseurs des droit de Í hommes au Niger of 20 June 2022, Journal officiel de la République du Niger, Loi Nº 2022-27; see also International Service of Human Rights, Model Law for the Recognition and Protection of Human Rights Defenders of 21 June 2016, available at: https://ishr.ch/defenders-toolbox/resources/model-law/ (last visited 15

Guidelines and policies for the protection of human rights defenders have also been issued by several international organisations and institutions.[16]

III. AI as a Tool for Human Rights Defenders

As stated above, AI can serve various needs of human rights defenders in carrying out their work. AI tools and applications may be used for analysing data, preparing and carrying out human rights campaigns and providing services to individuals who find themselves in situations which are critical

---

November 2025); and the following guidelines: the Guidelines for Irish Embassies and Missions on Human Rights Defenders of 2010, available at: https://www.humanrights.ch/cms/upload/pdf/150415_irish_hrd_guidelines_en.pdf (last visited 15 November 2025); the Swiss Guidelines on the Protection of Human Rights Defenders of 2013, available at: https://www.eda.admin.ch/eda/en/fdfa/fdfa/publikationen/alle-publikationen.html/content/publikationen/en/eda/menschenrechte-humanitaeres-migration/Leitlinien-zum-Schutz-von-HRD / (last visited 15 November 2025) or the Guidelines on Supporting Human Rights Defenders of Canada of 2017, available at: https://www.international.gc.ca/world-monde/issues_development-enjeux_developpement/human_rights-droits_homme/rights_defenders_guide_defenseurs_droits.aspx?lang=eng (last visited 15 November 2025).

[16] See for example Ensuring Protection - European Union Guidelines on Human Rights Defenders of 2008, available at: https://www.eeas.europa.eu/sites/default/files/eu_guidelines_hrd_en.pdf (last visited 15 November 2025); Guidelines on the Protection of Human Rights Defenders of the Organization for Security and Co-operation in Europe of 10 June 2014, available at: https://odihr.osce.org/odihr/guidelines-on-the-protection-of-human-rights-defenders (last visited 15 November 2025); Committee of Ministers of the Council of Europe, Recommendation to member States on the need to strengthen the protection and promotion of civil society space in Europe, CoE Doc. CM/Rec(2018) 11 of 28 November 2018.

concerning the enjoyment of human rights of those persons. Visualizing human rights violations and usage as a warning tool complements the usage of AI by and for human rights defenders. In the following section of this article, some applications of AI technology in the work of human rights defenders are discussed. The list of applications presented here is, of course, illustrative, not exhaustive.

1. Data Analysis

Human rights defenders often undertake to collect and analyse data of events that might be human rights violations or violations of international humanitarian law. Data about potential human rights violations might come from a myriad of diverse sources. Such sources may include interviews, photographs, sound files, video footage, written accounts, maps or satellite imagery. Once the data is analysed, it might be used to produce reports, as evidence in court procedures, in educational materials or in other ways. Especially in situations in which a great number of human rights violations take place, such as in armed conflicts or within the context of authoritarian governments, many potential pieces of evidence might be available. Analysing such huge quantity of data can be done by humans exclusively. However, the process can be time-consuming and requires many human recourses. Therefore, non-governmental organisations in the field of human rights make use of AI-driven technologies to support their experts and researchers to analyse data in these kinds of situations.

An example for such data analysis can be found in the practice of the non-governmental human rights organisation Amnesty In-

ternational.[17] The organisation used AI within the context of the human rights situation in Sudan, more precisely in the country's western region of Darfur.[18] The armed conflict in Darfur between the Sudanese government and opposition groups in the region has started in 2003. A core element of the situation is high level violence against civilians. The atrocities committed in Darfur include killings, torture, enforced disappearances, rape and other forms of sexualized violence, pillage, forced displacement and the destruction of villages.[19] In 2016, Amnesty International documented a new wave of violence which included attacks on and the destruction of many villages in Darfur.[20] To understand the scale of the destruction, Amnesty International launched the projects "Decode Darfur" and "Decode the Difference". The organisation called upon volunteers to scan satellite images of Darfur and to identify destroyed villages in the process. The task set for those volunteers was to study the available images and to determine if villages had been

---

[17] *Anne Dulka*, The Use of Artificial Intelligence in International Human Rights Law, in: Stanford Technology Law Review 26 (2023), pp. 316-366 (330).

[18] Ibid., p. 331.

[19] See for instance *Roberto Belloni*, The Tragedy of Darfur and the Limits of the 'Responsibility to Protect', in: Ethnopolitics 5 (2006), pp. 327–346; *Alex de Waal*, Darfur and the failure of the responsibility to protect, in: International Affairs 83 (2007), pp. 1039–1054; *Joyce Apsel*, The Complexity of Destruction in Darfur: Historical Processes and Regional Dynamics, in: Human Rights Review 10 (2009), pp. 239-259.

[20] *Milena Marin/Freddie Kalaitzis/Buffy Price*, Using artificial intelligence to scale up human rights research: a case study on Darfur, Amnesty International Evidence Lab of 6 July 2020, available at: https://citizenevidence.org/2020/07/06/using-artificial-intelligence-to-scale-up-human-rights-research-a-case-study-on-darfur/ (last visited 15 November 2025). The following paragraphs are based on that report.

destroyed between 2014 and 2016. The project was joined by 28,600 individuals from 147 countries. They generated 13 million annotations covering 2.6 million satellite images of approximately 100x100 meters each. With this method an area of 300,000 square kilometres was covered. However, this was not enough to cover the entire 493,180 square kilometres of the region. To scale up the work, Amnesty International made use of AI.

Before sharing the collected data with partners and engineers, Amnesty International developed a risk assessment model. In this model four criteria for sharing the date were identified, namely sensitivity and graphic nature of the data, transparency, data validation and risks of malicious attacks and misuse of data. Amnesty International then worked with experts on machine learning from the University College London in the United Kingdom during the next steps of the project. The large dataset of annotated chips of satellite imagery from the two prior projects on Darfur was used to train a machine learning model to automatically map the vast desert areas of the region. The model was able to identify "human presence", as well as "destroyed" or "mixed" (i.e., partially destroyed) villages. This was done for the whole of Darfur. The results were embedded in a web-based mapping application. This tool allowed the researchers of Amnesty International to refine the annotations and visualize at scale the patterns of habitation and the destruction. However, the work still faced considerable limitations due to a small team with limited resources, a reliance on a mosaic of archival images that were stitched together, and the fact that some tiles dated back more than ten years. To overcome those challenges, Amnesty International entered into a relationship with the then-existing Canadian AI company Element AI.

Through the collaboration of Element AI with the Satellite Applications Catapults' ML Use Case Programme, Amnesty International was able to access high resolution commercial data. These data allowed the Element AI team to assess the viability of training deep learning models for detecting destroyed villages in Darfur. With a dataset of images and crowdsourced annotations, indicating whether a human settlement was visible in an image, and if so, whether the settlement was destroyed, it was the aim of Amnesty International to learn the task of mapping an image to the correct annotation. This meant that the team of Amnesty International wanted to accurately predict whether a human annotator would mark that image as a destroyed village. After more fine-tuning, the organisation concluded that the best used system deployed in the general Darfur area would successfully identify 82% of all destruction cases (and miss 18% of them), while only 15% of the destruction "alerts" would be false positives. These metrics indicate that the deployed model could thereby save significant time and operational costs for intelligence research.

After the exercise, Amnesty International felt that large scale human rights research and continuous monitoring of conflicts aided by AI were within reach. However, numerous challenges occur when such technology is used for those tasks. According to Amnesty International, four main lessons were learned during the exercise. First, human rights organisations need to carefully evaluate the risks of such research against potential benefits. If effective migration strategies cannot be put in place, human rights organisations should be comfortable dropping such projects altogether. Secondly, training data for human rights AI algorithms is often not easily available, and data collection processes are not straightforward. Amnesty Inter-

national benefited from its already-large base of volunteers to generate such data. In the future, the organisation wants to use a combination of volunteer-driven analysis and analysis by AI. The idea for such work is that the AI can supply functionality in scale, whereas the volunteers provide depth to the analysis. Thirdly, Amnesty International pointed out that projects like the one on Darfur would require a large constellation of partnerships and pro-bono or heavily subsidized collaborations. Fourthly, the accuracy of algorithms should be carefully considered. Human rights researchers would need technical literacy to understand and interpret the data and to make crucial design choices such as calibration of accuracy metrics such as precision and recall.

Another example is the work of Human Rights Watch regarding Myanmar.[21] The human rights non-governmental organisation also partnered with Element AI to monitor ethnic violence in the country.[22] This was done against the background of violence against the Rohingya population in the state of Rakhine.[23] Human Rights Watch worked with the company to design a machine learning tool that could use satellite imagery and remote sensing thermal data to spot, track, and catalogue violations of human rights against the Rohingya.[24]

Human Rights Watch used thermal data to monitor the violence with a focus on attacks on settlements.[25] Indicators such as smoke plumes captured by environmental satellites were used in the procedure. The researchers were then able to make use of AI to combine this data with aerial images to identify spots where violence had taken place. Furthermore, AI was used to combine the data with other publicly available information, like videos and photos which were posted on social media. Thereby, Human Rights Watch was able to pinpoint where burnings of villages were conducted. This then allowed it to corroborate the testimony of individuals who were victims of human rights violations and to identify perpetrators. Human Rights Watch was able to identify at least 214 villages that were nearly totally destroyed as part on the ethnic cleansing campaign against the Rohingya by the Burmese military.

2. Human Rights Campaigning

Another important aspect of the work of human rights defenders is campaigning. Campaigning can for instance concern a particular situation in which human rights are in danger at a large scale, the fate of an individual who is in grave danger of having his or her human rights violated or general information about human rights. A campaign can bind many resources which cannot be used for other tasks while the campaign needs to be prepared and executed. AI can be used during the preparation and the running phase of a human rights campaign to support the work of human rights defenders in several ways.

---

[21] *Dulka* (fn. 17), p. 337.

[22] Ibid.

[23] See for instance *Anthony Ware/Costas Laoutides*, Myanmaŕ s 'Rohingya' Conflict, 2018; *Ken MacLean*, The Rohingya Crisis and the Practices of Erasure, in: Journal of Genocide Research 21 (2019), pp. 83-95.

[24] *Dulka* (fn. 17), pp. 337, 338.

[25] Human Rights Watch, Burma: Satellite Imagery Shows Mass Destruction - 214 Villages Almost Totally Destroyed in Rakhine State of 19 September 2017, available at: https://www.hrw.org/news/2017/09/19/burma-satellite-imagery-shows-mass-destruction (last visited 15 November 2025). The following paragraph is based on that report.

AI might be used as a tool for strategic communication.[26] In that regard AI tools can help human rights defenders with their ability to produce ideas within an instant and thereby reducing the time and resources needed to mobilize support and drive social change. Furthermore, it is a typical feature of human rights campaigns that they call to action for individuals. The individuals who are targeted with such a campaign might hold different values, different views on political, economic or social questions which are relevant for the specific campaign and have made different experiences in life. AI can assist human rights defenders in explaining the issues at stake in the campaign, which might involve complex points, and in convincing individuals of the goals of the effort. Such a complex point might be the relationship between climate change and the enjoyment of human rights. If human rights defenders, for instance, endeavour to convince individuals who are sceptical of the existence of climate change, they can harness the power of AI prompt engineering to better tailor their respective messages in a way that resonates more with the targeted audience.

AI technology can also play a role for human rights defenders in administrative and logistical tasks during campaigning. Through the use of such technologies, human rights defenders can allocate more time and energy to strategic and tactical planning. Repetitive tasks that may be fulfilled by AI tools or in which such tools might be helpful to include generating responses to e-mails, summarizing long texts and reports, project management or creating presentations. Moreover, AI-driven audio transcription technologies can help to make produced content more easily accessible in a variety of formats. In addition, AI-driven chatbots can be used as a first point of contact. The chatbots can engage with supporters and stakeholders and address frequently asked questions or answer initial inquiries.

Furthermore, AI can be used by human rights defenders to create political satire. The combination of humour and political protest has long been a catalyst for social and political change.[27] Furthermore, the use of humour, satire or parody can lead critical discourse about the absurdity and consequences of policy decisions.[28] Human rights defenders might use the capabilities of AI to create satirical songs, pictures, videos, small games or internet memes.[29] For instance, AI pictures might be generated that show political leaders in mock scenarios where they are presented as low-wage workers or refugees to visualize the impact of their political decisions or to unmask their rhetoric.[30] One example which was used by a human rights group concerned British politician Suella Braverman of the Conservative and Unity

---

26 *Melissa McNeilly*, The Human Rights Opportunities of Artificial Intelligence (AI), New Tactics in Human Rights of 28 August 2023, available at: https://www.newtactics.org/perspectives/human-rights-opportunities-artificial-intelligence-ai/ (last visited 15 November 2025). The following paragraphs are based on that article.

27 Ibid.; see also on this topic *Atyaf Fakhri Hassan Hussein*, The Use of Satire in Addressing Corruption Issues and the Benefit of Artificial Intelligence: An Analytical Study in the Program "Al-Rashid and the People", in: Fausto Pedro García Márquez/Alaa Ali Hameed/Akhtar Jamil (ed.), Pattern Recognition and Artificial Intelligence: Selected papers from the 6th Mediterranean Conference on Pattern Recognition and Artificial Intelligence (MedPRAI24), 2026, pp. 359–376; *Hang Lu/Shupei Yuan*, "I know it's a deepfake": the role of AI disclaimers and comprehension in the processing of deepfake parodies, in: Journal of Communication 74 (2024), pp. 359-373.

28 *McNeilly* (fn. 26).

29 Ibid.

30 Ibid.

Party. She has been a Member of Parliament and served as Secretary of State for the Home Department in the cabinets of Prime Ministers Mary Elizabeth Truss and Rishi Sunak.[31] Braverman is known for her very critical stance on immigration, refugee issues and multiculturalism.[32] The human rights group has let an AI tool create a picture of a woman who looks like Braverman.[33] This woman has a toddler in her arms and wears a life jacket. The picture shows her in front of a body of water with a coastline in the background. Thereby, it is intended that the woman in the picture looks like a refugee. The AI-generated image was meant to criticize Braverman's harsh immigration policies. Using such pictures in human rights campaigns might be seen as controversial; however, it can succeed in capturing public attention.[34]

3. Visualizing Human Rights Violations

AI-generated images or films cannot only be used by human rights defenders as a tool for tasks in campaigns as outlined in the previous paragraph. Such media files might prove also highly useful to visualize human rights violations. The ever-growing capabilities of AI applications in the sphere of generating imagery, videos and sound files can play a pivotal role in this regard.

Human rights violations often take place away from the public eye. Authoritarian governments, non-state armed groups or other actors might try to hide their atrocities to avoid prosecution by national or international institutions of criminal justice or condemnation by the international community or public outrage and thereby, to steer clear of losing support for their respective causes. An example for such behaviour includes the usage of the infamous Colonia Dignidad, a settlement of the likewise named Christian cult by German immigrants near the city of Parral in Chile. The Chilean government under dictator General Augusto Pinochet used the areal as a centre for internment, torture and murder of dissidents during the 1970s.[35] Another example was the secret prison complex El Vesubio in the metropolitan area of Buenos Aires. The military dictatorship of Argentina used the site to torture and murder political rivals from 1976 to 1978.[36] Human rights violations may also occur in the context of a general policy of public secrecy in and isolation of a country. This is, for instance, the case for North Korea where many human rights violations were made public by individuals who have managed to flee the country.[37] A

---

[31] Government of the United Kingdom, The Rt Hon Suella Braverman KC MP, available at: https://www.gov.uk/government/people/braverman (last visited 15 November 2025).

[32] See *Philip Hubbard*, Suella Braverman's talk of a refugee 'invasion' is a dangerous political gambit gone wrong, King's College London of 3 November 2022, available at: https://www.kcl.ac.uk/suella-bravermans-talk-of-a-refugee-invasion-is-a-dangerous-political-gambit-gone-wrong (last visited 15 November 2025).

[33] *McNeilly* (fn. 26).

[34] Ibid.

---

[35] See for a detailed analysis *Evelyn Hevia Jordán*, Colonia Dignidad: Lights and Shadows in the Recognition of the Victims, in: Elizabeth Lira/Marcela Cornejo/Germán Morales (ed.), Human Rights Violations in Latin America: Reparation and Rehabilitation, 2022, pp. 223–236; *Caroline Moine*, Denouncing or Supporting the Chilean Dictatorship in West Germany? Local Associations of Solidarity and Their Transnational Networks Since the 1970s, in: Global Society 33 (2019), pp. 332-347.

[36] See in greater detail *Gonzalo Conte*, A topography of memory: Reconstructing the architectures of terror in the Argentine dictatorship, in: Memory Studies 8 (2015), pp. 86-101.

[37] See for more detailed accounts the UNHRC, Report of the Commission of Inquiry on Human Rights in the Democratic People's Republic of Ko-

context in which human rights violations also take place away from the public eye is the situation of refugees who try to reach the territory of member states of the European Union via the Mediterranean Sea. Human rights violations against refugees on board of ships, boats and skiffs by the Libyan Coast Guard have been reported on numerous occasions.[38] The list of such human rights violations could be extended with ease.

Reporting on human rights violations that take place away from the public eye often relies on the testimony of victims or witnesses alone or to a great degree. While such accounts can be powerful in themselves, a visualization of the told stories has the potential to add additional weight to the reports. Human rights defenders can use AI-based tools to reconstruct the interior of a prison camp, the site of a massacre or the conditions on board of coast guard vessels in images or films. The so created media files can then for instance be used to raise public awareness for a specific situation, in educational events about human rights issues or in presentations at hearings of parliamentary committees. Usually, the task of visualizing human rights violations was carried out by draftsman who manually, or by using graphic software, produced images based on the accounts of the victims or witnesses. However, using AI for such a role offers the advantage that images can be created more quickly and in

greater numbers, while reducing costs for human rights defenders.

4. Warning Tool for Human Rights Defenders

While human rights defenders work for the rights of others, they themselves can come under attack from those who want to curtail their activities. Such attacks can be of a violent or a non-violent character. While such attacks – for example the tracking of online activities of human rights defenders or the automated blocking of online content – are addressed regarding AI in a paragraph below, the usage of AI as a warning tool for human rights defenders is also worth discussing.

The matter of using AI as a warning tool for human rights defenders became prominent in the work of the Office of the High Commissioner for Human Rights of the United Nations (OHCHR).[39] The duties of the OHCHR encompass the monitoring and the protection of human rights defenders against attacks. The Human Rights Indicators and Data Unit of the OHCHR increasingly focuses on using data to strengthen both monitoring and reporting of threats to human rights and, thereby, also to human rights defenders. It is a difficult undertaking to gather the necessary data for accurate reports on threats to and attacks on human rights defenders.

To find better and more efficient ways to protect human rights defenders the OHCHR entered into a collaboration with

---

rea of 7 February 2014, UN Doc. A/HRC/25/CRP.1; *Jin Woong Kang*, Human Rights and Refugee Status of the North Korean Diaspora, in: North Korean Review 9/2 (2013), pp. 4-17.

[38] See *Andreina De Leo*, Fostering Accountability for Human Rights Violations in EU Border Externalization Through the European Ombudsman: The Case of Contesting Financial Support to the Libyan Coast Guard, in: Journal of Immigrant & Refugee Studies 23 (2025), pp. 104-120.

[39] *Amy Lynn Smith*, Building a Collaboration to Protect Human Rights Defenders, Medium of 16 January 2023, available at: https://unhumanrights.medium.com/building-a-collaboration-to-protect-human-rights-defenders-26457ae8abd0 (last visited 15 November 2025). The following paragraphs are based on that article.

the US-American AI company Dataminr in April 2022. Dataminr had already worked with the United Nations in the "United Nations Global Pulse" initiative which attempts to bring real-time monitoring and prediction to development and aid programs. The company's products include "Dataminr Pulse", a tool to monitor real-time events and to support crisis responses by providing playbooks, messaging tools and post-event documentation.[40] The "First Alert" platform of Dataminr enables first responders to quickly act in cases of emergencies. It is designed as a real-time critical event discovery solution which is intended to maintain situational awareness and make decisions with confidence.[41]

5. Improving the Human Rights Situation on the Ground Directly

AI can also be used by human rights defenders to improve the human rights situation on the ground directly. Individuals or groups who face human rights challenges can be equipped with tools to improve their respective situation based on the capabilities of AI. AI-based speech-to-text applications and translation services are of utmost importance in that regard.

The potential of such AI applications in human rights fieldwork was demonstrated by the Norwegian Refugee Council. The humanitarian organisation used chatbots to help migrants from Venezuela in Colombia. Those chatbots were designed to help the migrants to learn about their rights under Columbian immigration laws and policies.[42] AI applications may also be used in humanitarian activities regarding resource management. This is demonstrated by the AI application "Dr. Tania". This tool was developed by Neurafarm, a farming startup founded by IT experts and engineers from Indonesia, starting in 2018.[43] "Dr. Tania" is designed to help Indonesian farmers to tackle crop diseases that threaten their harvest. Thereby, the application also has a link to the right to food. The name of the tool comes from the word "tani" which means "farmer" in Bahasa Indonesia. Farmers can upload a picture of diseased plant to the application. "Dr. Tania" then compares the image of the plant to other pictures in its database and offers the user a diagnosis. In the next step, the farmer is offered information on how to manage the disease and on treatment procedures. The application is designed as a chatbot to ensure easy handling for the user. The uploaded pictures are added to the database, increasing the accuracy of diagnoses in the future. While this is not a usage of an AI application by human rights defenders in a narrow sense it nonetheless shows the potential of such applications

---

[40] Dataminr Pulse for Corporate Security, available at: https://www.dataminr.com/products/pulse/corporate-security/ (last visited 15 November 2025).

[41] Dataminr First Alert, available at: https://www.dataminr.com/products/first-alert/ (last visited 15 November 2025).

[42] *Leila Toplic*, AI in the Humanitarian Sector, NetHope of 6 October 2020, available at: https://nethope.org/articles/ai-in-the-humanitarian-sector/ (last visited 15 November 2025); *Meheret Takele Mandefro*, AI-powered knowledge Chatbot (Norwegian Refugee Council), United Kingdom Humanitarian Innovation Hub of 19 June 2025, available at: https://www.ukhih.org/news/ai-powered-knowledge-chatbot-norwegian-refugee-council/ (last visited 15 November 2025).

[43] *Leander Jones*, Dr Tania: An Indonesian AI Chatbot Helps Farmers Identify and Treat Crop Disease, RESET – Digital for Good of 6 August 2020, available at: https://en.reset.org/indonesian-ai-powered-app-helps-farmers-identify-crop-disease-05252020/ (last visited 15 November 2025). The following paragraph is based on that article.

in the field of resource management for people affected on the ground.

IV. AI as a Threat to Human Rights Defenders

AI does not only offer great potential for human rights defenders. It can also be a threat to them, as already hinted at above. Human rights defenders can face infringements of their specific rights as human rights defenders for instance by governments using AI. It is known that governments have used AI applications for surveillance of human rights defenders or to block their content in social media platforms.[44]

An example for such usage of AI is the "Oculus" system.[45] It became known in 2023 that the government of the Russian Federation introduced this system, which helps its authorities to scan the internet for "illegal content".[46] The government has stated that the main task of "Oculus" is to recognize violations of Russian laws in pictures and videos on the internet. The system is capable of analysing texts which are shown within pictures and videos. It was also claimed that the system could evaluate up to 200,000 pictures per day. Before, this task was done manually by employees of the General Radio Frequency Centre, a subordinate regulatory authority of the Federal Service for Supervision of Communications, Information Technology and Mass Media (Roskomnadzor). According to government officials, an employee is able to process 106 pictures and 101 videos per day on average. "Oculus" would also be able to detect "extremist content", calls for "illegal demonstrations", "drug-promoting content" as well as "LGBTQ propaganda". The introduction of "Oculus" and its described usage must be seen as a threat for human rights defenders in Russia. Since the beginning of the full-scale invasion of Ukraine by the Russian Federation in February 2022, the authorities are trying to suppress undesirable opinions even more.[47] Those activities do not only concern questions of the war, violations of international humanitarian law and human rights within it and its consequences for Russian politics, economy and society, but also the suppression of expressions of LGBTQ rights. Therefore, "Oculus" has the potential to make human rights defenders in Russia turn away from using the internet and especially social media to shed light on human rights issues within the country or in countries, like Ukraine, which are impacted by human rights violations by the Russian state.

Another example concerns Egypt.[48] The authoritarian government of the North African state uses AI as a tool for political repression and surveillance since the events of the Arab Spring in 2011. AI is used increasingly to act against the free-

---

[44] See *Rasma Kaskina/Angelina Cvetkovska*, Artificial intelligence (AI) and human rights: Using AI as a weapon of repression and its impact on human rights, 2024, available at: https://www.europarl.europa.eu/RegData/etudes/IDAN/2024/754450/EXPO_IDA%282024%29754450%28SUM01%29_EN.pdf (last visited 15 November 2025).

[45] Ibid., p. 3.

[46] *Matthias Kremp*, Russland automatisiert Suche nach »verbotenen Inhalten« im Internet, Spiegel Netzwelt of 13 February 2023, available at: https://www.spiegel.de/netzwelt/web/russland-automatisiert-suche-nach-verbotenen-inhalten-im-internet-projekt-oculus-a-ac0d1273-5a71-4035-9a39-757c5d0de9ef (last visited 15 November 2025). The following explanations are based on that article.

[47] *Kaskina/Cvetkovska* (fn. 44), p. 3.

[48] Ibid., p. 4. The following paragraph is based on that report.

dom of expression and the freedom of assembly. Under the Anti-Cyber and Information Technology Crimes Law, which was enacted in 2018, authorities have the legal power to monitor online content and to block websites to protect national security or the economy. While the intended use of the law is to combat political extremists and terrorist groups, in practice it is used against dissidents and other citizens alike. The AI-driven cyber espionage and surveillance techniques of the Egyptian government were and are employed against Egyptian journalists, academics, lawyers, opposition politicians and human rights defenders within the country and those living abroad as well.

V.　International AI Legislation and Human Rights Defenders

The specific position of human rights defenders has not been explicitly addressed in recent international legislation concerning the regulation of AI. In the field of human rights, the main instrument at the moment is the aforementioned Framework Convention which was adopted under the auspices of the Council of Europa. The treaty was signed on 5 September 2024. According to its Article 1 para. 1, the object und purpose of the convention is "to ensure that activities within the lifecycle of artificial intelligence systems are fully consistent with human rights, democracy and the rule of law". As far as can be seen, rights and vulnerabilities of human rights defenders were not discussed during the drafting process. This is not necessarily astonishing as human rights defenders enjoy human rights themselves and therefore can be seen as already being included

in human rights codification at the international level. However, during that process critical remarks concerning the rights of individuals and groups who are engaged in protective activities were made by Michel Frost, the United Nations Special Rapporteur on Environmental Defenders under the Aarhus Convention.[49] In a statement on the proposed Council of Europe Framework Convention,[50] addressed to the Committee of Ministers of the Council of Europe and the permanent representatives of the member states of the Council of Europe in Strasbourg, he explicitly criticized the provisions on national security and national defence in the draft convention which are now also present in the adopted instrument.[51] According to Article 3 para. 2 Framework Convention "A Party shall not be required to apply this Convention to activities within the lifecycle of artificial intelligence systems related to the protection of its national security interests, with the understanding that such activities are conducted in a manner consistent with applicable international law, including international human rights law obligations, and with respect for its democratic institutions and processes." Furthermore, according to Article 3 para. 4 Framework Convention "Matters relating to national defence do not fall within the scope of this Convention." Frost argues in his statement

---

[49] Convention on Access to Information, Public Participation in Decision-making and Access to Justice in Environmental Matters of 25 June 1998, UNTS vol. 2161, p. 447.

[50] Committee of Minister of the Council of Europe, Draft Framework Convention on artificial intelligence, human rights, democracy and the rule of law, CoE Doc. CM(2024)52-prov1 of 15 March 2024 (Draft Framework Convention).

[51] The statement by *Michel Frost* can be accessed here: https://unece.org/sites/default/files/2024-05/SR_EnvDefenders_Statement_CoE_FrameworkConvention_AI%20and%20Human%20Rights_08.05.2024.pdf (last visited 15 November 2025).

that "(the) complete exemption for matters relating to national defence and the vaguely-worded exemption of matters relating to national security (...) creates a significant risk for abuse and legal loopholes. Indeed, many AI systems used for surveillance and monitoring of the activities of environmental defenders could thereby be excluded from the scope of the Framework Convention on the basis that the activities are allegedly necessary for 'the protection of national security interests'."[52] In line with the recommendations by the Parliamentary Assembly of the Council of Europe,[53] he, therefore, called on member states as a matter of absolute urgency "when finalizing the draft Framework Convention, to revise current articles 3 (2) and (4) by limiting the national security interest and national defence exceptions. While such exceptions may be warranted under certain circumstances, a blanket exception for matters of national defence is not. Instead, the text of the Framework Convention must provide in unequivocal terms that AI activities necessary to protect national security interests or national defence must be conducted strictly in line with international human rights law and other international obligations, including article 3 (8) of the Aarhus Convention. This means also that any exception to the rules and principles under the Framework Convention, including in relation to matters of national interest or national defence, must pass the tripartite test of legality, proportionality and necessity under international human rights law."[54] While his mandate only concerns environmental de-

fenders under the Aarhus Convention,[55] the same apprehensions are true for human rights defenders as both are engaged in activities of defending vulnerable conditions and people and often use the same methods and techniques for their work.

VI. Conclusion

The explanations and examples presented in this article have shown on the one hand, that human rights defenders can profit in their work from the usage of AI-driven technologies much in the same ways as individuals, groups and organisations in other work environments. AI technologies can be used to simplify task, finish tasks quicker and add practical and creative potential to human rights work. However, using AI by human rights defenders also has the same limitations and challenges as in other work environments.[56] There is a necessity to keep humans in the loop for checking flaws in the output of AI technologies. Moreover, for the sake of transparency, AI-created content must be labelled as such to avoid allegations of misconduct or deception.[57] Furthermore, the necessary data to feed the AI mechanisms

---

[52] Ibid., p. 2.

[53] Parliamentary Assembly of the Council of Europe, Opinion on Draft Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law, Opinion 303 (2024) of 18 April 2024.

[54] *Frost* (fn. 51), pp. 2, 3.

[55] The mandate can be found in Decision VII/9, adopted by the meeting of the parties to the Aarhus Convention, UN Doc. ECE/MP.PP/2021/2/Add.1.

[56] See also *Sam Bowman*, The role of artificial intelligence in predicting human rights violations, Open Global Rights of 14 November 2024, available at: https://www.openglobalrights.org/the-role-of-ai-in-predicting-human-rights-violations/ (last visited 15 November 2025).

[57] See for an example *Luke Taylor*, Amnesty International criticised for using AI-generated images, The Guardian of 2 May 2023, available at: https://www.theguardian.com/world/2023/may/02/amnesty-international-ai-generated-images-criticism (last visited 15 November 2025).

must be obtained, which can be a challenging process. Depending on legislation at the national and the international level, further legal requirements must be fulfilled. On the other hand, AI technologies are already used by authoritarian governments to hinder human rights work and to persecute human rights defenders. As the capabilities of AI are growing at an ever-faster rate and AI-driven tools and applications become more available, there is no doubt that AI will have a great impact on the work of human rights defenders in the future, to their benefit as well as to their detriment. More scholarly work will be needed to accompany these developments.

Vita

The author is a professor of constitutional law at the University of Applied Sciences for Police and Public Administration in North Rhine-Westphalia, Germany.