Commentary

HUMAN RIGHTS WEEKEND: ARTIFICIAL INTELLIGENCE, BIG DATA & HUMAN RIGHTS: PROGRESS OR SETBACK?

Isabella Banks [*] *and Leonore ten Hulsen* [**]

I. Introduction

The "Artificial Intelligence, Big Data & Human Rights" lecture took place on Sunday, February 10, 2019, the final day of Human Rights Weekend. Human Rights Weekend is an annual event hosted by Human Rights Watch, an international non-governmental organisation that conducts research and advocacy on human rights. This year's Human Rights Weekend was located at De Balie, a cultural centre in Amsterdam where artists, politicians, thinkers, opinion leader, scientists, and the public can meet and comment on developments in society.

The lecture brought together experts from a variety of disciplines to share their views on the interplay of technology, artificial intelligence (AI), big data, and human rights. The speakers discussed the advantages and disadvantages of our increasingly interconnected society and the overflow of open-source information: data that is collected from publicly available sources for use in an intelligence context. They considered the ways in which the internet can both threaten human rights and help to protect them. The purpose of the lecture was to raise awareness about the implications of technological developments for the human rights field. This commentary is a brief review of the views expressed by the speakers in their individual talks and in the panel discussion that followed.

II. Open-Source Investigation: Leveraging Big Data to Pursue Accountability for Human Rights Violations

Benjamin Strick is an open-source investigator for the BBC and a BellingCat contributor. Mr. Strick opened the event with a step-by-step explanation and demonstration of his investigative work at BellingCat. BellingCat is a platform that uses open-source information and social media to investigate war crimes and human rights abuses in various conflicts around the world. Moreover, it creates guides and case studies so that citizen investigators, like Mr. Strick, may learn to do the same.

Mr. Strick began by stating that in conducting open-source investigations, it is important not only to collect data, but also to track the way it was collected. This transparency makes open-source investigations more credible and helps others learn the investigative process. The purpose of Mr. Strick's work with BellingCat is to rebuke fake news and uncover evidence that supports human rights investigations and may contribute to criminal prosecutions.

---

[*] Isabella Banks is currently pursuing a Master of Science (MSc) in International Crimes, Conflict and Criminology.

[**] Leonore ten Hulsen is pursuing a Master of Laws (LLM) in Internet, Intellectual Property and ICT-law and a Master of Laws (LLM) in International Technology Law. Both authors are editors of the Amsterdam Law Forum and students at the Vrije Universiteit Amsterdam.

According to Mr. Strick, open-source information is everything that is freely available online. This includes data from the BBC, Instagram, Facebook, Twitter, Strava, OKCupid, Google Maps, as well as lesser known applications like Stalkscan.com, Flightradar 24, and Yandex.

Mr. Strick then showed the audience an example of an open-source human rights investigation that he and other BellingCat investigators around the world carried out. The investigation began with a video clip filmed on July 10, 2018. This video – which ultimately went viral – showed two civilian women and their children being shot and killed by men who appeared to be soldiers. The incident appeared to take place in Cameroon, but the Cameroonian government quickly issued a statement declaring it "an instance of gross misinformation". The government denied involvement in the crime captured in the video on the following grounds: the physical landscape in the video was different to that of Cameroon; the weapons and uniforms of the soldiers in the video were not used by Cameroon's military; and the men involved were not members of the Cameroon military.

Mr. Strick showed the audience how open-source information and social media could be used to investigate the "where", "when", "how", and "who" of the human rights violation featured in the video and test the veracity of each of the Cameroonian government's claims.

The "where" of the crime was verified using geolocation and satellite imagery. Mr. Strick and his colleagues used screenshots to analyse the mountain range in the background of the video and compared its topography with other mountain ranges documented by Google Earth. Over the course of three weeks, they were able to match it with a mountain range in north Cameroon. The location of the crime was further corroborated by an analysis of the features of the landscape that appeared in the video, such as buildings, paths, trees, and terrace lines. In doing so, Mr. Strick and his colleagues refuted the government's first claim.

The "when" of the crime was ascertained using chronolocation. To determine the year, Mr. Strick and his colleagues used Google Earth to analyse how the location they had identified changed over time and compared it to the features of the landscape that appear in the video. To narrow down the time further, they used the height of one of the soldiers and the length of his shadow to calculate the angle of the sun. With the help of a free online application called SunCalc, they determined that the crime took place between March 20 and April 5, 2015.

To find out how the women and children were killed, Mr. Strick and his colleagues compared screenshots of the weapons in the video with weapons on Google Images. Once they had identified the type of weapon the soldiers in the video used, they used videos on YouTube and images on Google, Instagram and Facebook to determine that this was indeed the same type of weapon used by the Cameroon military. In so doing, they refuted the Cameroonian government's second claim that the weapons in the video were not used by the Cameroon military.

The government's claim that the uniform worn by the men in the video were not those of the Cameroon military was similarly disproven. Mr. Strick and his colleagues again used social media to find images of Cameroonian soldiers geotagged to the same area that the crime was committed wearing the same uniforms as the men in the video were wearing. This finding was verified with reference to online databases of military uniforms of military units all over the world.

Finally, Mr. Strick and his colleagues used Facebook to track members of the Cameroon military and identify the soldiers that appeared in the video. Mr. Strick explained that military ranks and

photos are often proudly and publicly shared on social media. Mr. Strick and his colleagues were ultimately able to identify the man who shot the four civilians.

When the Cameroon government learned of BellingCat's investigation, it issued an arrest warrant for the men involved and stated that it would investigate the human rights violation. A year later, no investigation has been initiated. Despite the inaction on the part of the Cameroon government, Mr. Strick's presentation clearly demonstrated the power of open-source investigation and the massive amount of information that can be uncovered online without using any paid services.

III. Artificial Intelligence and the Law: The Importance of Regulating the Use of Technology

Marlies van Eck is an assistant professor at the eLaw Institute of Leiden University and a principal consultant at Hooghiemstra & Partners. Professor van Eck began by stating that her talk would take a legal perspective on AI. She pointed out that this perspective is often left out in the process of building technological systems, which can be problematic given technologists' optimistic tendency to focus on averages rather than anomalies and "what might go wrong". This has resulted in a failure to critically examine the use of technology in addition to the technology itself.

To illustrate this point, Professor van Eck introduced the theories of historian Melvin Kranzberg and political theorist Langdon Winner. The first of Kranzberg's "six laws of technology" states that "technology is neither good nor bad; nor is it neutral". In the same vein, Winner argues that certain technologies are inherently political. Professor van Eck cited Winner's famous example of Long Island bridges that city planner Robert Moses specifically designed to prevent public buses from passing under them. This, in turn, limited the access of black and low-income communities to an acclaimed public park. Like these bridges, technologies may be designed (consciously or subconsciously) to have a particular social effect.

Professor van Eck challenged the audience to consider who should bear responsibility for bad algorithmic outcomes. She argued that the use of technology is a human decision and not one that should be explained away by an unpredictable algorithmic "glitch". She further criticized the tendency to refer to "ethics" as a silver bullet to address this issue. Rather, she emphasized the need to develop legal rules and regulations.

Professor van Eck stressed that in order for lawyers to keep up with and regulate the use of technology, they must learn more about AI and initiate more jurisprudence-developing lawsuits. This is well within the realm of possibility, given that the Netherlands has been inspired by Finland's example to provide a free online course in AI for anyone who is interested in learning to build it responsibly. She concluded by emphasising the importance of interdisciplinary co-operation and collaboration – particularly between supervisory authorities, such as the EU Agency for Fundamental Rights and national data protection – in order to "rule AI" rather than "letting AI rule us."

IV. Government Monitoring of Online Platforms: How Automatic Upload Filters Limit Freedom of Expression, Data Protection, and Privacy

Evelyn Austin is a movement builder at Bits of Freedom, a digital rights organisation based in the Netherlands that focuses on privacy and communication at the national and EU level. The subject of her talk was a worrying trend: EU institutions requiring online platforms to monitor

everything we upload to the internet. This is made possible by a technology known as "upload filters" and comes with various consequences.

First, it means that companies like Facebook and Google decide the limits of our freedom of expression. Ms. Austin gave an example of how Facebook prevented an American woman from sharing an article published by the *New York Times* – which contained photographs of children who were starving as a result of Saudi Arabia's war in Yemen – on the basis that it was in breach of their community guidelines against "nudity or sexual behaviour". Despite the journalist's explanation of the humanitarian importance of sharing the article and accompanying photographs, the upload filter bot would not revise its decision. Faced with significant public pressure, Facebook eventually reinstated the post without any explanation of the filtering process or transparency around what had gone wrong. Not long after, the article was taken down again and the woman's account was blocked – again without explanation.

Ms. Austin's second example of this phenomenon concerned the Dutch advocacy organisation Women on Waves, which provides reliable medical information on the termination of pregnancies. Its YouTube videos were repeatedly taken down and reinstated without explanation. The organisation went through YouTube's redress procedure, but it seemed to have no effect. Eventually, its YouTube account was blocked altogether, at which point Bits of Freedom intervened and got it reinstated. YouTube considered the problem solved, but never explained its reasons for blocking Women on Waves' video content and account.

Ms. Austin expressed serious concern about the lack of transparency around the flawed content monitoring mechanism of these platforms. She argued that upload filters severely impact citizens' freedom of expression, data protection, and privacy and called for takedown mechanisms that have a basis in the law. Ms. Austin criticised the EU institutions that require this kind of monitoring for treating complex social problems as technological issues and in so doing, strip European citizens of their rights. She concluded with the following warning: in the best-case scenario, it is Facebook or Google that allows us to speak.

V. Panel Discussion Hosted by Human Rights Watch

At the conclusion of Ms. Austin's remarks, the three speakers were invited to return to the stage for a panel discussion. The discussion was hosted by **Sarah St. Vincent,** a researcher and advocate on national security, surveillance, and domestic law enforcement for the US program of Human Rights Watch, and moderated by **Bahram Sadeghi**, an Amsterdam-based freelance television director and host.

Mr. Sadeghi invited Ms. St. Vincent to comment on the speakers' remarks based on her background in US government surveillance. She began by stating that Mr. Strick's presentation – which demonstrated the "magic" of open-source investigation – made her think about the extent to which governments use those same tools to achieve repressive goals. She pointed out that this is worrisome because, unlike Mr. Strick, governments have a monopoly on the use of force and imprisonment. This highlights the need for strong safeguards and limits on what governments can do.

In response to Ms. Austin's presentation on upload filters, Ms. St. Vincent noted that governments often use issues that are important to the public – such as terrorism – to justify the use of such mechanisms and other breaches of privacy. She pointed out governments' failure to use these mechanisms to address issues such as white supremacy or misogyny. She questioned

why more policies were not in place to protect against government misuse of the information and to preserve the freedom of expression as well as other fundamental human rights.

Sarah St. Vincent then introduced a "human rights framework" for preventing government interference with the right to privacy, which ensures that measures are lawful, necessary, and proportionate. In discussing the "lawful" requirement, she pointed out that even in the context of a lawsuit, governments sometimes choose not to disclose what information they have or how they have used it. This deliberate withholding of evidence can make the legal action Professor van Eck advocated for difficult, and further highlights the need for strong legal protections and limitations on what governments can do. Ms. St. Vincent explained that the "necessary" criterion means that the measure is necessary to achieve a legitimate goal, and that there is no less intrusive way of achieving that goal. She noted that the European Court of Human Rights' standard is 'strictly necessary in a democratic society'. Ms. St. Vincent defined a "proportionate" measure as one that does not disproportionately impact citizens' rights or have a chilling effect on people's choices and behaviour. She concluded by stating that surveillance and AI are extremely powerful tools that must be regulated in a free society.

The ensuing panel discussion centred on digital privacy and the extent to which individuals are personally responsible for preventing the misuse of their data. The majority of the speakers felt that expecting citizens to educate themselves about these issues and avoid sharing sensitive information online was unfairly burdensome. The panel again discussed the importance of strong laws and regulations to protect citizens that cannot or are not interested in protecting themselves. Mr. Strick added that the existence of privacy rules may not be enough, and that the terms of service of platforms like Facebook and Twitter should be written in a way that non-lawyers can understand. Ms. St. Vincent responded that if these companies are *de facto* monopolies, individuals who want to be connected in the world may have no meaningful choice about whether or not to use their platforms, regardless of their terms of service.

To conclude the event, Ms. St. Vincent returned to the concept of self-education. She emphasised the importance of grassroots advocacy and recommended reading two books to learn more about the consequences of AI for poor and minority communities: *Automating Inequality* by Virginia Eubanks and *Weapons of Math Destruction* by Cathy O'Neill.

VI. Discussion and Conclusion

Taken together, the "Artificial Intelligence, Big Data & Human Rights" lecture provided a balanced overview of the implications of technological developments in an increasingly digitalised society. However, one element that may have been lacking from the panel discussion that followed was a more critical examination of open-source investigations, their effectiveness, and their impact on the presumption of innocence and the right to privacy. Despite the moderator's efforts to press the speakers on this topic, they appeared reluctant to fully elaborate their views – perhaps out of respect for Mr. Strick.

Open-source investigations like those conducted by BellingCat are currently on the rise due to the increasing availability of open-source tools and data.[1] In light of this trend, it is worth considering three questions that were either not raised or not explicitly answered during the panel discussion.

[1] E. Moerman, 'Burgers in het digitale opsporingstijdperk', *NJB* 2019-94, p. 1.

First, are open-source investigations an effective way to hold perpetrators accountable? Mr. Strick's compelling human rights investigation of four murders in Cameroon ultimately did not result in any meaningful action by the government to hold the identified perpetrators accountable. What would it take for the relevant authorities to give the necessary credence to civilian investigations so as to ensure prosecution? As multiple panellists pointed out, these government authorities have a monopoly on force that Mr. Strick and other BellingCat contributors lack. New technologies like the eyeWitness to Atrocities app developed by WITNESS – which enables civilians to securely collect and store verifiable photos and videos of human rights violations – may offer part of the solution, and ultimately help to "democratize" international criminal investigations.[2] As open-source investigations become increasingly verifiable, governments and courts may need to consider ways in which they can better collaborate with organisations like BellingCat.

Second, are civilian investigations reliable, and if not, how might they affect the presumption of innocence? Professor van Eck's description of the non-neutral nature of the use of technology raises the question of how the individual biases of civilian investigators affect their use of open-source tools. It was clear from Mr. Strick's presentation that in investigating the Cameroon government's claims about the video, he was operating from the assumption that the crime had in fact been carried out in Cameroon by members of the Cameroonian military. Bias in civilian investigations has the potential to produce inaccurate findings which, left un-checked, could jeopardise the fairness of the response (whether by a justice institution or a vigilante mob).

Third, what impact might civilian investigations have on the right to privacy? Nowadays, police in the Netherlands use social media for criminal investigations.[3] Surprisingly, however, the Netherlands has no clear legislation governing the use of open-source information for this purpose. Though select organisations are working to reform the Dutch criminal code, *inter alia* by regulating open-source investigations, it may be a long time before meaningful changes are introduced.[4] This legal lacuna brings to light the need for safeguards against the abuse of open-source information by police in addition to civilian investigators.

In a recent report, "Harnessing the Power of Technology for Human Rights" Human Rights Watch highlights 'greater capacity to conduct open-source investigations using the vast material available online' as one of its priority investments for the future. While the power of open-source investigation to document and seek accountability for human rights violations is undeniable, it is worth considering what it might mean when these tools are used by relatively anonymous individual actors as opposed to an internationally recognised organisation with a clear human rights mandate. Overall, this event emphasised the need for additional research on the impact of AI and big data on human rights, as well as improved legislation to regulate the use of technology.

---

[2] R. Hamilton, 'The Hidden Danger of User-Generated Evidence for International Criminal Justice', at: https://www.justsecurity.org/62339/hidden-danger-user-generated-evidence-international-criminal-justice/?fbclid=IwAR21bjpzgTXA-Jah4SUu5fYQz29gXn94C8OiR24HEk7LGJ5GmTrtEJBNsf8 (last accessed on 8th of July 2019).

[3] E. Moerman, *supra* note 1, p. 2.

[4] *Idem,* p. 3.