Report on the Presidency of Germany CoE's Online Event: *"Human Rights in the Era of AI: Europe as International Standard Setter for Artificial Intelligence"*[1]

As courts and authorities around the world try to find the balance with Big Tech -say, the latest debate on Australia's legislative approach on particular tech companies,[2] The German Presidency of Council of Europe (CoE) held an online conference named "Human Rights in the Era of AI: Europe as International Standard Setter for Artificial Intelligence" with a very divergent and distinguished list of speakers on January 20, 2021. Under the captivating moderation of Deutsche Welle TV's Chief Political Correspondent Melinda Crane, the event consisted of two panel debates with experts and scholars, a high-level roundtable, vivid opening and closing remarks from both Germany CoE Presidency and CoE Commissionaires. The conference was fruitful with the exchange of thought-provoking ideas and had the core spirit of the hot debate with the addition of some deeply striking clips from the documentary iHuman, directed by Tonje Hessen Schei, through the flow of speeches. The availability of different language options increased the accessibility of the program with three options provided to the audience: English, German and French.

"What sort of society do we want to live in?", a quote from the documentary iHuman, echoed throughout the entire event, while the debate scrutinized the relationship between one of the emerging new technologies, namely, AI and human rights, democracy, and rule of law. After Melinda Crane's opening, the Federal Minister for Foreign Affairs for Germany, Heiko Maas, took the stage and directed the audience's attention to some use cases concerning which AI deployment can be both 'good' and 'bad'. He raised economic welfare, research to fight diseases -such as the global pandemic- and easier communication in cyberspace as beneficial uses, while he pointed out to autonomous weaponization in warfare or the weaponization of AI, authoritarian regimes' mass surveillance policies and the radicalization of people within the social networks on the other side. Continuing his words on the widely acknowledged need for public regulations when it comes to defining the thin line between freedom of speech and hate crime, he said those regulations should be made through multilateral cooperation and international collaboration that is free from any geopolitical bipolarity. His words were referring to the Chinese model of digitalization and Silicon Valley's profit-prioritizing approach, which cannot uphold any kind of value on monopolies. He called for expanding Europe's AI capability by highlighting the new EU budget for digitalization and concluded his speech on setting standards for human centered-AI by emphasizing the importance of establishing human control over algorithms, manifesting a good practice of privacy and non-discrimination, standing tall against the extremism and hate speech.

Right after the first opening speech, the Council of Europe Commissioner for Human Rights Dunja Mijatović continued with her remarks on human rights and AI relations. Talking about unlocking potential benefits of AI, she underlined how unwanted outcomes could be created when relying too heavily on calculations. Giving examples of errors and failures, like decisions made by algorithms and those about social payments in the area of public services, tracking and tracing devices which are used to enforce lockdowns and to measure the spread of infections, she raised concerns about discrimination and data protection risks. She upheld that the authorities should consider the safeguards to be set up for the digital welfare context, the contingencies to be covered while also ensuring the technological solutions do not infringe the right to privacy, data protection, non-discrimination and dignity.

Mentioning the current tools, she pointed out the Recommendation on Artificial Intelligence by the Committee of Ministers of CoE in 2019. Among ten areas of action, she deemed transparency rules and businesses' compliance with human rights standards to be the most important ones. To our view, the need for raising awareness within the tech industry to incorporate human rights into the design of AI is one of the challenging topics in today's discussion. Dunja Mijatović crowned her speech with a call for more inclusive and interdisciplinary cooperation between all stakeholders including states, academia, private sector, NGOs, the media and civil society to enhance AI capabilities for a prosperous future.

Following the opening remarks, a mini clips section from iHuman was presented. After the mini exhibition of the documentary, director Tonje Hessen Schei and producer Danielle Turkov Wilson shared their thoughts and experiences, as well as their perspective on the extraordinary effort spent on the film. The concentration of power, the governance challenges, and the blind-eyed trust in calculations in the public sphere -especially in predictive policing- were the most salient points they raised. Overall, it can be said that the film has a stimulating effect on not only policymakers but citizens too and manages to raise awareness among a variety of people and interest groups.

In the course of the conference, three panel debates with the participation of highly dedicated speakers and also an insightful presentation were eloquently delivered. In the first panel debate, AI's impact on human rights, democracy and rule of law and possible ways out for problematic cases were scrutinized with the contributions of AI Law & Ethics specialist Nathalie Smuha from KU Leuven; Cornelia Cutterer, Senior Director and Lead of the Rule of Law & Responsible Tech team at Microsoft; Francesca Fanucci, Member of the CAHAI, Conference of INGOs of the Council of Europe and David Leslie, Ethics Theme Lead at the Alan Turing Institute and CAHAI Bureau member.

Following the presentation of CAHAI Chair Gregor Strojin on CAHAI's feasibility study and prospective CoE legal framework for AI, the next panel debate, where the likelihood of new international frameworks as well as the concerning uses and consequences of AI were examined, was comprised Professor of Law & Technology Peggy Valcke from KU Leuven, the Deputy Director-General for G7 and G20 Relations Yoichi Iida from the Ministry of Internal Affairs and Communications of Japan, Europe Policy Analyst Daniel Leufer from Access Now, the Deputy Director of the Department of Strategic Development and Innovation Egor Shipitsyn from the Ministry of Economic Development of the Russian Federation and the Co-Founder and Executive Director of AlgorithmWatch Matthias Spielkamp.

The last high-level panel debate was committed to questioning the how's and why's of the further enhancement and development of the proceedings performed by the CoE and the EU, with the participation of the State Secretary Christian Kastrop from the Federal Ministry of Justice and Consumer Protection, the Assistant Director-General for the Social and Human Sciences Gabriela Ramos from UNESCO, the Director of Information Society and Action against Crime Jan Kleissen from the CoE, the Director of the EU Agency for Fundamental Rights (FRA) Michael O'Flaherty, the Acting Director-General, DG Justice and Consumers Salla Saastamoinen from the European Commission.

Keeping the human rights, democracy and rule of law as indicators at the heart of the debate, the most remarkable takeaways from the debate follows as:

- Public use of AI -mostly automated decision-making systems- can result in a potential denial of services. It can cause people to lose out on salary promotions, to be denied

loans and to have low rates for obtaining insurance packages. More importantly, it can even be used to measure whether citizens are prone to commit a crime or not and result in unfair outcomes.[3]

- AI use in social welfare programs can make disadvantaged and underrepresented groups of people increasingly more excluded. For instance, youth, indigenous communities, people of color and vulnerable portions of society can be impacted by unfair and unchecked decisions even if they cannot access the Internet. At this point, two concrete examples can be given, which are the UK's A-level algorithm scandal in 2020[4] and an algorithm named 'Amigo-boras' used in Dutch migration control.[5]

- From a broader angle and in the long run, deep societal infrastructures such as human rights, rule of law and democracy can be damaged. These negative consequences might have different appearances on human dignity -by physical and mental means-, human autonomy and freedom.

- The use of AI in social networks, with the political and social debates shifted into cyber spaces, can leave the entire society radicalised and polarized by "being a safe haven" to those who exacerbate hate speech and extremism. The latest and the saddest example of this is the attack on the Capitol of Washington, on January 6, 2021. The polarisation can easily amplify its effects in such circumstances by scale and scope. This, to our view, also proves that there is a subtle border between real life and cyberspace which can have reciprocal and irreversible consequences regardless of the time and space.

- One of the frequently raised dilemmas was the broad proliferation of smart cities and smart 'everything'. However, all these developments can lead to function creep, just like in San Diego[6] or in a more explicit way like Chinese surveillance panopticon[7].

- Another point of conflict is the contest between innovation and the laws. While the question "Do you want innovation or all these laws and ethical rules?" might be too reductive , the scariest scenario is hampering technological development with binding and granular rules which might result in a sort of *"innovation migration"*. However, it should not necessarily turn out that way, especially with the help of some effective 'socio-technical' business models (STMs)[8] and socio-technical system design (STSD)[9].

- With regard to FRTs and automated gender recognition, i.e. gender attribution, the use of AI in smart public billboards, public toilets and in the rest of vast applications like Giggles,[10] can cause many discriminative problems by excluding trans people and violate the very right to equality and non-discrimination.

---

[3] https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3357874.
[4] https://www.bbc.com/news/technology-53836453.
[5] https://doi.org/10.1080/1369183X.2019.1674134.
[6] https://www.bloomberg.com/news/articles/2020-08-06/a-surveillance-standoff-over-smart-streetlights.
[7] https://www.theatlantic.com/magazine/archive/2020/09/china-ai-surveillance/614197/.
[8] https://www.researchgate.net/publication/263065156_The_Three_Roles_of_Business_Models_for_Socio-Technical_Transitions.
[9] https://academic.oup.com/iwc/article/23/1/4/693091.
[10] https://www.theverge.com/2020/2/7/21128236/gender-app-giggle-women-ai-screen-trans-social.

- An issue that was raised many times during the conference is the transparency considerations. Transparency has two faces, one of them is more innate as it relates to the algorithm itself while the other one is more extrinsic as it concerns proceedings of the use and the deployment of AI. Therefore to solve transparency issues, a participatory co-design with a multidisciplinary approach is necessary.

- Accountability was another frequently addressed topic through the event. As questions like "Who do you hold responsible when an algorithm causes harm and maybe even someone's death?" are being asked, it may be time to embed the responsible deployment of these technologies into innovation policies. At this point, one of the big tech companies can be a good example when it comes to incorporating ethical considerations and human rights sensitivity into engineering processes. With an inhouse department to develop responsible and explainable technology, their commitment to respect the rule of law shall be appreciated. On the flip side, language use can make a difference. In order to prevent legal gaps in the responsibility context, it is crucial for policymakers to diligently regulate the accountability aspect within jurisprudence, and for the society to have a more demanding attitude about human responsibility by being fully aware that AI is not magic but human-made.

- The explainability was another critical issue through the discussion. Thinking of simplifying the complexity, one can deem that there is no one size fits all solution regarding the non-linearities and high-dimensional correlations within the high-profile calculations in the algorithms. Effective and practical solutions for the so-called black-box problem can remain in socio-technical approaches and good practice by diminishing unconscious bias among developers, deployers and recruiters. Therefore, the explainability can be provided by education and awareness training right from the start of the processes.

Grey lines, red lines and hybrid legal solutions

In order to figure out what type of rules should be relied on and what sort of policy should be followed, the work done so far by the CoE's Ad Hoc Committee of Artificial Intelligence (i.e., CAHAI) has an important role. The role of CoE in standard-setting with its cross-cultural consultation custom was repeatedly underlined with examples of fruitful conclusions of the Budapest Convention on Cybercrime[11] and the Convention on Data Protection (i.e. Convention 108)[12] which are ratified by a broad range of countries.

Most of the speakers were in consensus on whether there is a shift from a rights approach to an obligations approach. There should be some 'concrete' regulations that prevent any formation of legal vacuums. Hence, as the citizens of the biosphere we already have some settings in place with self-regulation efforts, standards by civil organisations, NGOs, scientific authorities, roadmaps by international and national bodies. Also, on a more legal ground, legislations such as European Convention of Human Rights, the EU Charter of Fundamental Human Rights; the above-mentioned

---

[11] https://www.coe.int/en/web/conventions/full-list/-/conventions/treaty/185.
[12] https://www.coe.int/en/web/portal/-/convention-on-data-protection-turns-40-committee-of-ministers-declaration.

Convention 108 and the Convention on Cybercrime all are applicable to some extent. However, these frameworks are not enough to protect individuals when an incident occurs and to preserve democratic societal infrastructures as a whole in the long term.

As frequently perceived in the discussion, it is more preferable to take a step back and to scrutinize with a broad vision rather than strongly advocating for soft law or hard law institutions. Having AI everywhere in our lives requires a diligent inspection and maybe a new and different legal solution with an hybrid ecosystem. In accordance with the feasibility study of the CAHAI in December[13], a mixture of binding and non-binding rules might provide an acceptable and reasonable solution. However, having such an entangled setting of AI use in real life practice can also require semi-public and semi-private solutions in place. For instance, the governments' reliance on the data and the algorithms provided by private companies, from which they purchase and deploy these algorithmic systems into the public sphere, can be an issue. It is one of the arguments when it comes to regulations because of the grey lines of these two poles that have their very own dynamics.

Also, one of the questions at stake was whether there is a need to define AI or not. At this point, one of the good examples given about the definition problem was the Oviedo Convention on Human Rights and Biomedicine[14]. It was stated in the conference, that there is no need for thorough definitions to prohibit human cloning and design babies. As such, the same attitude can be adopted in the AI practice. Anyone can agree that there could be some risks and injustices within the outputs when an algorithm is trained by biased and imbalanced data sets[15] or data sets that do not fit for the purposes of the application of the system (e.g. data sets collected by selective or purposeful bias[16], poor data visualization[17] so on) without agreeing on a single definition.

The majority of the stakeholders iterate the 'risk-based approach'[18] within legislative processes. This, to our view, is the rightfully inevitable conclusion of the ubiquity of different types, levels and functions of AI. Regarding this view, binding rules should be proportionate, minimal and sector-specific or maybe cross-sectorial so that the future innovation will not become hampered by introducing a set of strict regulations. Nevertheless, this approach does not necessarily target only high-risk applications but all levels of risks. That is why a broad mapping of AI applications has an important role to unveil even the least risky uses of AI, with the aim of keeping the technology more transparent, traceable, verifiable and under human control.

By uniquely focusing on human rights, rule of law and democracy, the Council of Europe's complementary inner setting brings about associations with lots of different kinds of stakeholders. In this regard, the consultation with other international bodies and civil society was also one of the points emphasized. In this instance, the observatory seats and involvement of parties such as the UN and the UNESCO are valuable. Additionally, with the inauguration of the new US president Joe Biden, a hopeful statement was made for a future transatlantic cooperation in the legislative efforts during this crucial phase on AI's fate. In respect to the complementary inclusive forum setting, one of the points worth appreciating was the CoE's global consultation with youth representatives, in other words, the

13 https://www.coe.int/en/web/artificial-intelligence/-/the-feasibility-study-on-ai-legal-standards-adopted-by-cahai

14 https://www.coe.int/t/dg3/healthbioethic/Activities/Bioethics%20in%20CoE/

15 https://towardsdatascience.com/handling-imbalanced-datasets-in-machine-learning-7a0e84220f28

16 https://www.datapine.com/blog/misleading-statistics-and-data/

17 http://blog.analytics-toolkit.com/2020/the-perils-of-poor-data-visualization-in-cro-a-b-testing/

18 https://www.ferma.eu/european-commission-considers-a-risk-based-approach-to-artificial-intelligence/

future generations of the innovative society. Additionally, it was suggested that the national action plans should aim to address the problems concerning the better inclusion of the marginalised and disadvantaged groups of people. The representation of the Global South is one of the issues that should be carefully carried out by working groups and expert committees.

Overall, the entirety of the concerns raised in the talks should be understood not only through the lens of European values of modern democratic society, but also from an international humanitarian angle, considering how the technology affects every one of us as citizens of the one and only habitable planet within our "solar street". Stressing that the ethical sensitivity is embedded into human rights, it would be exciting to witness the creation of new frameworks which put the human at the center and respect the universal values of the humanist hive mind, with collaborative multi stakeholder co-design that involves all divergent and valuable entities, grassroot movements and ground citizens of the planet Earth, we hope and believe that the policymaking processes will have fruitful conclusions in the trust and excellence ecosystem.