Rathenau Instituut

Input for the OHCHR consultation on the practical application of the Guiding Principles on Business and Human Rights to the activities of technology companies

By Rathenau Instituut, February 2022

The Rathenau Instituut is the Dutch technology assessment institute positioned between science, policy en society. Our role is to research and debate the impact of science, innovation, and technology on society, and to support the Dutch Parliament in its decision-making on these issues. From our 35 years of experience researching the impact of technology on society, and our recent lessons in using a Business and Human Rights lens for the technology sector, we want to share the following observations and insights, along the lines of the session topics described in your call for input.

Addressing human rights risks in business models
We observe different human rights risks when it comes to business models applied by technology companies.

In our recent research report on Online Harms, we distinguish a number of mechanisms that facilitate and inspire online harmful behavior. One of these mechanisms is the *Attention Economy*. The Attention Economy entails the ads-based business models run by many technology firms. By using algorithms to display content or personalized ads, companies try to capture our attention and keep us active on their platforms. The more users are engaging on the platforms, the more lucrative ads-based business models become. Sensational, indiscriminate and extreme content tends to elicit engagement of users, which means this content is often prioritised. Recent reporting in the Wall Street Journal based on Facebook whistleblower Francis Haugen points in this direction.

The social media platforms provide a platform for the dissemination of private opinions and news. This *amplification* function has clear advantages: everyone has access to a public stage. However, amplification also has disadvantages: quality and fact-checking are no criteria to reach the masses. Rather, indiscriminate and sensational content generally has a greater reach than nuanced and balanced information on platforms. It is useful in this context to distinguish between Freedom of Speech and Freedom of Reach (of content). Currently, the social media platforms use algorithms to determine which content is widely distributed and which less. Platforms have the ability to make or break the reach of content, with possible impact on the functioning of the public debate and freedom of speech, essential elements of deliberative democracies.

Data collection (tracking people), data analysis (profiling people), and application of data analyses (influencing people's behaviour) are fundamental elements in various business models in the technology sector. This so-called *cybernetic loop* provokes inherent human rights risks. This is the case for business models dependent on online advertisement through social media platforms or search engines, but also for business

models evolved around immersive technologies such as Virtual Reality (VR), Augmented Reality (AR) and Speech Technology.

What these business models have in common, is that they often use artificial intelligence (AI) systems in order to improve outputs. AI can be seen as the "brain" behind various of these technologies, whether it may be for more personalized ads for users on social media platforms, or for the rendering of more realistic, real-time virtual worlds in VR. AI is a general-purpose technology and can be found in many digital technologies, such as the Internet of Things, digital platforms, or virtual and augmented reality.

Considering that the application of AI-driven technologies should take place within human rights-frameworks, we want to highlight two human rights risks in business models in particular. Firstly, it has an impact on **the right to privacy**, since the quality of AI based systems (e.g. virtual worlds) rendered to users depends heavily on collected personal data. Secondly, the application of collected data for the creation of personalized experiences (e.g. personalized news feeds or virtual objects in public spaces only visible to users of VR/AR-technology), may lead to fragmented worldviews. It may significantly affect our shared sense of reality with detrimental effects for e.g. marginalized groups and social cohesion. As described in our report Digital threats to democracy, examples of information that may lead to fragmented worldviews are deepfakes and disinformation, which are often disseminated through group chats, search engines, virtual assistants or by making use of micro-targeting. The selection and filtering by AI-driven technologies can pose a threat to access to information. Access to information is an integral part of the **right to freedom of opinion and expression.**

Human Rights Due Diligence (HRDD) and end-use
Combining knowledge of Business and Human Rights frameworks with the tech sector, we observe a disconnect between the two: HRDD frameworks are not yet commonplace in discussions on the corporate accountability of technology companies, even though these frameworks seem very applicable and useful for addressing the human rights challenges and dilemmas. Unlike unfair distributions of benefits and risks between the global north and south along more traditional supply chains (such as food supply chains or the garments industry), many of the current human rights challenges in the technology sector affect users of in the global north and south simultaneously. The challenges have to do with balancing rights of individuals within the same society: among others privacy versus accountability and freedom of speech versus safety. To tackle these dilemmas, outcome based policies are less useful than risk and process based policies such as due diligence. This relates to the following question of accountability and remedy.

Accountability and remedy
Accountability in the tech sector is in its infancy and tech optimism has led to great accountability gaps. It is only recently that we start to recognize the societal impacts of online platforms, social media and AI based systems. Examples are the use of AI-

systems with a disproportional discrimination bias towards minority groups, or news manipulation. Currently, very few actors in technology supply chains recognize their responsibility for adverse human rights impacts. The due diligence process described in the UNGPs and OECD Guidelines in particular can be useful to clarify responsibilities in diffuse actor networks, including from a supply chain perspective. Of particular importance is the distinction between different types of responsibility depending on the type of involvement in caused harm: that companies have a responsibility to contribute to remedy for rights holders when they cause or contribute to adverse human rights impacts, and that the 'directly linked' scenario also comes with responsibilities to address the adverse impacts.

Therefore, it is encouraging to notice that the upcoming European Digital Services Act (DSA) is expected to adopt this risk and process based approach. The DSA as proposed by the European Commission puts forward the obligation for very large online platforms to carry out risk assessments on the systemic risks associated with their services (e.g. in relation to their content moderation policies and the use of their algorithms), to mitigate those risks and to have them independently verified. Several adopted amendments by the European Parliament aim to strengthen the due diligence obligations further.

The State's duty to protect, or regulatory and policy responses
In the European context, we see due diligence frameworks entering the policy debate on digitalisation. Examples are the upcoming European AI Framework and the aforementioned DSA. Here state actors and supranational bodies try to steer the development and effects of new technologies with respect for human rights. While basic elements of HRDD are built into these legislative proposals, there is potential to build more strongly and explicitly on the human rights due diligence frameworks, guidance and best practices that have been developed by the OHCHR, the OECD and in other business sectors from 2011 onwards. At the minimum, the legislative proposals could explicitly refer to UNGP and the OECD Guidelines. Furthermore, the governance of the enforcement mechanisms that will be developed, will be crucial for the effectiveness of the proposed policies. An overreliance on auditing by third party firms does not seem recommendable because of the demonstrated limited success of this approach to detect human rights risks and prevent human rights impacts in other sectors (e.g. garments, electronics, oil and gas). Robust transparency and due diligence obligations combined with public oversight and sufficient capacity to perform this massive task seem to be required to further guide digitalisation within the boundaries of human rights frameworks.