Human Rights and eService Innovation: An Unlikely Crossbreed or an Essential Need?

Kyiv, 2020

The "Human Rights and eService Innovation: An Unlikely Crossbreed or an Essential Need?" discussion paper looks at the crossroads of digital transformation, human rights and development assistance in the twenty-first century. It was originally intended as an internal document for UNDP Ukraine circulation. At the same time, subsequent discussions have revealed interest of external stakeholders in this topic, and the publication became an open document. The proposed piece opens the space for a wider and more participatory dialogue on the noted topics.

Authors: Maksym Klyuchar, UNDP Digital Governance and Policy Development Consultant and Leila Haccius, UNDP Human Centered Design and Research Consultant.

The Digital, Inclusive, Accessible: Support to Digitalisation of State Services in Ukraine (DIA Support) Project Inception Phase is implemented by UNDP as a preparatory step for the bigger digital transformation initiative.

UNITED NATIONS DEVELOPMENT PROGRAMME (UNDP) partners with authorities, civil society, businesses and communities to help build nations that can withstand crisis, drive and sustain the kind of growth that improves the quality of life for everyone. In Ukraine, UNDP work focuses on three core areas: peace-building and recovery, democratic governance and energy efficiency. Ukraine currently has one of the largest UNDP programmes in the Europe and CIS region.

www.ua.undp.org

Contents

'Across Europe and the globe algorithms, Artificial Intelligence (AI), machine learning (ML) and automated decision-making (ADM) […] are increasingly, and sometimes stealthily, encroaching upon the space ordinarily inhabited by human actors.'[1]

'While efficiency may be valuable, those responsible for human lives should not pursue efficiency at the expense of fairness – fundamental human rights must hold a central place in this discussion.'[2]

'[…] The well-off might have instant access to up-to-date and easy-to-use computers and other hardware as well as fast and efficient broadband speeds, the least well-off are far more likely to be severely disadvantaged by out-of-date equipment and time-consuming and unreliable digital connections.'[3]

1   'Regulating for an equal AI: a new role for equality bodies. Meeting the new challenges to equality and non-discrimination from increased digitisation and the use of Artificial Intelligence'. https://equineteurope.org/wp-content/uploads/2020/06/ai_report_digital.pdf, p. 27.

2   'Bots at the gate. A human rights analysis of automated decision-making in Canada's immigration and refugee system'. https://citizenlab.ca/wp-content/uploads/2018/09/IHRP-Automated-Systems-Report-Web-V2.pdf, p. 62.

3   Report of the UN Special Rapporteur on extreme poverty and human rights. https://undocs.org/A/74/493, p. 15.

Executive summary

Ukraine is determined to introduce an increasing number of electronic public services (eServices) for its citizens. The overall narrative until this point has been focused on *efficiency*, *cost-saving* and the *benefits of innovation* for Ukraine as a state rather than on the interaction between rights-holders (citizens-clients) and formal and informal duty-bearers (state agencies and the corporate sector). Meanwhile, the intervention of the Ukrainian Ombudsperson and systematic application of the Human Rights Based Approach (HRBA) at all stages of the service re-design, including the re-engineering of the business-process, the creation of software and deployment of systems to full use, could facilitate a shift in this paradigm.

This discussion paper originally aimed to document cases where ombudspersons in other, mainly European, jurisdictions were able to successfully incorporate human rights considerations into the design of eServices. The review has not found evidence of such practices in the relevant annual reports or on the institutional websites reviewed. The paper, however, found a rich body of data, research and current developments (almost all from 2019 and 2020) surrounding human rights and the deployment of artificial intelligence (AI) systems, which are briefly described. The analytical note concludes with a set of recommendations for the United Nations Development Program in Ukraine and includes a deeper examination into the practices of the National Human Rights Institutions (NHRIs) of Denmark, Estonia and Finland as well as the line agencies of Sweden.

Introduction and contextual background

The United Nations Development Programme (UNDP) in Ukraine is currently working on the launch of a three-year initiative to assist the Government of Ukraine (GoU) in better delivering on their promise of *more accessible, convenient and secure electronic services at the national level.* The anticipated project focuses primarily on state-provided public services (usually handled by Ministries or central government bodies) rather than on the functions of local or regional-level authorities. As such, the planned intervention will work on re-engineering several large-scale and complex sets of services to bring them online and to mobile phones through the unified national Diia ('State and myself')[4] eService platform while ensuring that those who do not have access to such devices or internet service are not left behind.

Previously, the e-governance focus in Ukraine was largely aimed at efficiency, cost-saving and the innovative benefits for Ukraine as a state rather than on the interaction between rights-holders (citizens-clients) and the formal and informal duty-bearers (state agencies and the corporate sector). The intervention of the Ukrainian Ombudsperson and systematic application of the Human Rights Based Approach (HRBA) at all stages of the service re-design, including the business-process re-engineering, creation of software and deployment of systems to full use, could facilitate a shift in this paradigm.

Overall, UNDP intends to collaborate with Ukraine's Ministry for Digital Transformation (MDT) to apply a Human Rights Based Approach lens to this task. The HRBA is a conceptual framework for policy design and execution that is based on universal standards and international human rights law and principles. The core human rights principles include 1) the express application of a human rights framework (ensuring that all aspects of planning reflect the human rights legal framework); 2) empowerment (ensuring that policies, strategies and

---

4   See https://diia.gov.ua

programmes are based on empowerment not charity); 3) participation (ensuring that citizens (as rights-holders) are not just consulted but are considered active, free and meaningful participants represented by CSOs); 4) non-discrimination (ensuring that vulnerable groups are not excluded from the design or project cycle); 5) accountability (ensuring that formal and informal duty bearers can be held accountable, that there is clarity about rights and duties and that baselines, indicators and targets are measurable).

By collaborating with key institutions (such as the MDT, the Ombudspersons Office (OO) and/or other agencies), UNDP will strive to identify and analyse the significant factors that may constrain rights-holders (citizens) from actively, meaningfully and effectively claiming their rights and prevent duty-bearers from meeting their obligations in terms of access to the benefits, documents and services through electronic means. Barriers may include both human and material capacity gaps at both the duty-bearer and rights-holder levels, such as awareness about rights, lack of access to information, lack of transparency, lack of material resources etc.

Specifically, the project should ensure that the most vulnerable, underprivileged populations are not left *behind* or *excluded* from discussions when services are de-composed, re-assembled and turned into code, applications and IT systems. Ukraine has many marginalised communities, including internally displaced persons, retirees, people with disabilities, people who lack identification documents or proper registration/residency permits (so-called 'propiska'), women who

may have to balance child-rearing with work or running a business, persons with low income and those who have recently become unemployed, rural dwellers, veterans returning to civilian life and others.

The second guiding rail for the project is to guarantee that considerations regarding human rights impacts will always accompany the standard features of service efficiency, speed, deregulation and appealing design. In other words, the unique contribution of UNDP will focus attention on the genuine involvement of target groups throughout the entire service re-design and apply the Human Rights Based Approach (HRBA) lens and principles, as explained above, at all stages.

Hence, human rights and the HRBA are central to both the project and e-governance reform and should guide the overall process, from scope, design and implementation to monitoring, adaptation and evaluation. For this to succeed, engagement is critical on the part of the duty bearers to ensure an active and meaningful dialogue with rights-holders and oversight bodies (for example, Parliament, CSOs and the OO) and the integration of the most vulnerable groups into the service design.

In the case of digital service creation, this approach allows for the co-creation of digital and mobile services, encompassing everything from the mapping of necessary regulatory changes to the creation of new user journeys according to the rules and the identification of accessibility features. As noted above, the project will also emphasise compensation for the possible lack of a compatible and connected devices among some of the

rights-holders (citizen-clients). In this journey, the overall objective is to empower citizens (as rights-holders) and make their interaction with the state (as the duty-bearer) in a particular life situation easier rather than construct systems that are sleek and stylish but benefit some while disenfranchising others.

The brief desk review[5] and informal conversations with subject-matter experts presented here were initially introduced to closer examine the ways in which National Human Rights Institutions (NHRIs) were handling the growing wave of public service digitalisation and working to identify and minimise rights violations inevitable to the deployment of any new policy tool. The review team was particularly interested in discovering concrete examples where NHRIs, despite their highly diverse mandates and the peculiarities of national regulations, were able to join a policy cycle early, maintain a human-rights–based focus in the executive branch and co-design a product that would have a lower risk of human rights violations.

We can summarise the original hypothesis of the review as follows: *'In today's world of rapidly-changing technologies, it is likely that NHRIs somewhere in Europe have already partnered with executive branch authorities that review and develop new policies for the creation of electronic services for citizens/rights-holders. In such a case, NHRI specialists could serve as the HRBA lens for the policy process and would be able to make sure that vulnerable citizen groups are not left behind, low-income citizens would be able to use these services on par with those who can afford the latest unlimited internet phone models, that personal data*

security is given paramount priority and that no one is discriminated through the application of the new electronic service mechanism.'

The scope of the review initially focused on the broad spectrum of human rights issues that arise when a new digital service is designed (related to regulations, user-journey mapping, software and the completed and launched product) rather than the now 'traditional' approaches to website or electronic instrument accessibility and user-friendliness. However, as the desk-research progressed, it became increasingly clear that the original hypothesis was insufficient.

Instead, as our query has demonstrated, the primary focus of NHRIs today appears to concentrate largely on the prevention of possible human rights violations that stem from the exponential growth of artificial intelligence systems and their application in situations of fully automated or assisted decision-making. These considerations will be presented in a dedicated section of the review paper.

This thought-piece intends to serve two practical purposes. First, a preliminary review of Ukraine's Office of the Ombudsperson website and consultations held with line experts suggested that the NHRI could benefit from peer-to-peer experience sharing with its colleagues regarding the ways in which HRBA considerations are embedded into the blueprint, design, rollout and full-scale implementation of electronic services elsewhere. Indeed, Ukraine's NHRI is indirectly confronting themes

---

5   The research was conducted between 1 May 2020 and 1 August 2020.

that should be included in a full-scale HRBA-based policy review (for instance, protecting the right to access information or controlling the personal data protection regime) but has not yet worked with state authorities on a complete policy cycle through this paradigm. Accordingly, one of the sub-goals of the exercise was to discover good cases and build preliminary connections to serve as a foundation for possible experience cross-pollination. In addition, as the team of the new project comes on-board in late 2020, we hope to better inform their activities by offering recommendations and distilled knowledge processed from multiple sources and consultations with sector experts. This will facilitate a smoother point-of-departure for the project team and faster deployment of core operations.

The paper has been structured into sections to help readers search for relevant information faster and dwell on various themes that were brought to the research team's attention throughout the desk research and expert consultations. **Section one** summarises the findings from a review of NHRI reports, website materials and selected expert discussions in search of case studies exemplifying the participation of human rights institutions in the policy-development process for electronic services or in strategic cases of rights-violations due to a lack of service accessibility or inability to enjoy a service due to socio-economic constraints or other systematic faults. In **section two**, we look at the broader spectrum of players in the IT and human rights nexus and present the topical issue of human rights and systems for algorithmic decision-making and artificial intelligence.

The paper concludes with an **annotated bibliography** that suggests further readings for practitioners who would like to dive deeper into the issues of human rights and IT deployment by governments worldwide.

The research team would like to thank all of the peer-reviewers and line experts who have provided their valuable expertise and input to make this piece more comprehensive and well-rounded.

Section one:

NHRIs and their participation in electronic service design and implementation

These days, the digitalisation of state services is gaining momentum in countries around the world as an instrumental means of reducing petty or administrative-level corruption, improving transparency and strengthe-ning democratic governance. The seminal 2019 thematic report of the UN Special Rapporteur on extreme poverty and human rights notes that the digitalisation of state services has the potential to resolve many human-rights–related issues, including financial inclusion, reduced gender inequalities through the empowerment of women and girls, increased access to health services and social safety nets for the poor.[6]

At the same time, the report cautions against a 'rosy' view of digital transformations and outlines three high-level challenges that may be present throughout the design of digital public services. First, the report argues that the digitalisation realm, out of all possible economic activity areas, has remained a laissez-faire field given to the market to self-adjudicate or self-adjust. As such, while there is lip service from major technological companies on human rights and the ethics of proposed solutions, the reality stands in stark contrast to the given statements. Second, most states in the world lack detailed rules for the operation of such systems, and, when launched, some of them operate illegally. Third, the mandatory introduction and rapid expansion of such systems evoke questions about entirely new digital gaps (as old ones are, arguably, closed by 'technology in your pocket'). In the digital era, most services that once took place over the phone or in person are being replaced by online services. Citizens will, therefore, experience the digitalisation of state services in different ways depending on their socio-economic status, internet accessibility in the region where they live and digital literacy skills. This will particularly impact citizens who do not possess an internet-enabled device

---

6   See https://undocs.org/A/74/493, p. 6.
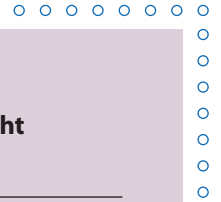
as well as the elderly who did not grow up in a digital world. This poses a severe threat to the enjoyment of social rights.[7] More generally, the possible violation of social and economic rights is linked to a paradigm shift; the holders of rights become applicants, and the humans once governing the process of recognising the applicants' rights have been replaced by machines.[8] This, the report argues, makes it more challenging to maintain a human-centric process that is sensitive to specific vulnerabilities.[9]

The states hold the primary responsibility to respect, protect and fulfil human rights. Consequently, one could expect the National Human Rights Institutions (NHRIs) to act and demonstrate how new policies for digitalised services need to mainstream human rights and ensure that the 'applicant – decision-maker' logic is replaced by, or at least softened with, the 'rights-holder – duty-bearer' vision. Indeed, NHRIs are supposed to serve as the watchdogs of government activity while simultaneously advising them[10] and acting as a bridge between the international and national level as well as between the CSO and state entities. Their mandate consists of various activities that include advising state institutions on current practices and national laws as well as their compliance with international law; monitoring the human rights situation in the field; reporting the human rights situation in their respective countries to national and international bodies; handling complaints of individuals and providing legal assistance; promoting international human rights through the publication of research, recommendations and opinions;

Consequences of machine-made decision-making on a citizen's right to an effective solution

The delegation of decision-making processes to machines poses three key challenges. First, people are less likely to challenge decisions made by machines, either because they lack the necessary information to challenge them or because those decisions are considered more objective and neutral. This expectation of neutrality must be nuanced. As the following section will explain in further detail, machine-made decisions risk repeating existing inequalities and discrimination (so-called biases). Second, governments do not necessarily understand how this decision-making process unfolds, nor the grounds on which decisions are made. Especially in the case of AI, it is often unclear who to hold accountable in the case of error and/or it might not be possible to get the necessary information from a private supplier. Third, machines are not able to exercise discretion in the same way that people are, which is important for the application of public policies.

---

7   See https://undocs.org/A/74/493, pp. 15–16.
8   Ibid., p. 16.
9   Ibid., pp. 16–17.
10   For more information, see https://www.humanrights.dk/learning-hub/challenges-nhri-effectiveness

cooperating with NGOs, CSOs and other regional actors; promoting the culture of rights through training and awareness-raising.[11]

Thus, this paper posits that, faced with ever-proliferating digital services, the NHRIs in question (institutions existing within the EU) were expected to, at a bare minimum, deliver three core functions: 1) assist states in the design of proper policies and development of new systems based on HRBA; 2) monitor the deployment and continued use of such systems while defending the rights of users; 3) distribute information about new tools and the digital literacy skills necessary to enjoy them. In particular, one could question the existence of an obligation of due diligence toward NHRIs to identify and mitigate harm during the development, planning and purchase of new technology prior to the rollout. Limiting their activity to monitoring the impact of new technologies on human rights after they have already been launched may be too late for reversal. With reference to the UN Guiding Principles on Business and Human Rights, NHRIs may also play a role in states' obligation to develop a National Action Plan on Business and Human Rights.

To assess whether NHRIs[12] possess experience in the design and development of eServices in accordance with HRBA, the following paper examined the annual reports of NHRIs over the past five years (2015–2020), conducted a website search and considered publications from the European network for NHRIs (ENNHRI). The initial research was limited to EU member states.[13]

NHRIs that have not mentioned digitalisation and e-governance

The desk review has not found any evidence of digitalisation-related engagement from the following NHRIs: Greece, Ireland, Latvia, Liechtenstein, Lithuania, Norway, Poland, Romania, Slovakia, Slovenia, Switzerland. This may indicate that the NHRIs of these states have yet to look into the matter because it is considered 'too recent', or it may be due to a failure to update their own web-resources accordingly. This would imply that they are aware of the importance of adhering to HRBA in the development of digital tools, and that it is only a matter of time before they eventually delve into the subject. This may be the case, for example, for the NHRIs in Lithuania and Slovenia as both attended a conference in Tallinn, Estonia, regarding the implementation of the Estonian e-government programme and the emerging challenges and opportunities created by these new technologies that intend to improve the implementation of the functions of state institutions.[14]

Yet another reason could be that NHRIs may not have sufficient in-house expertise in this field and may not see it as an area in which they should build the relevant knowledge. That is, of course, until citizen complaints 'come in droves'. It is also possible that these states are slow-acting on the uptake of eServices, and there are still few electronic systems in place. As such, rights-holders may not yet be complaining of violations to the NHRIs. This final explanation is highly unlikely, seeing as most

---

11   See http://ennhri.org/about-nhris/
12   We looked into the NHRI that are official members of the ENNHRI, see: http://ennhri.org/our-members/
13   Including EFTA countries: Iceland, Liechtenstein, Norway and Switzerland.
14   See http://www.lrski.lt/images/dokumentai/Seimo-kontrolieriai-ataskaita-2018_EN_WEB.pdf

EU member states are currently undergoing a process of digitalisation.[15] The research team believes that it is only a matter of time before these NHRIs begin to address the nexus of human rights and the digitalisation of state services. Inevitably, such questions and issues will become more frequent and attract increasing attention.

No information or annual reports could be found on the websites of the NHRIs in Cyprus, the Czech Republic and Hungary. Moreover, the states of Italy[16], Malta[17], Iceland[18] and Sweden[19] do not have an NHRI. Annual reports were not always available for every year between 2015 and 2020.[20]

NHRIs that have started looking into digitalisation and e-governance

The NHRIs of the following states mentioned digitalisation and e-governance in their annual reports or on websites: Belgium, Denmark, Estonia, Finland, France, Germany, Great Britain, Luxembourg, the Netherlands, Portugal, Scotland, Spain. Among those, four of the NHRIs discussed the issue of digitalisation in relation to the COVID-19 response measures. Considering that such measures are temporary, the human rights concerns related to digitalisation (mainly tracking apps and accessibility of the internet) are also limited to narrow and transitory state regulations. This is the case for the NHRIs of France (which explored digital tracking),[21] of Great Britain (which highlighted the differences in terms of access to justice),[22] of Luxembourg (which analysed restrictions to the right to education during COVID-19 for children in refugee centres)[23] and of Scotland as an independent human rights institution (which focused on the human rights implications of digital contact tracing technology).[24] These NHRIs concentrated primarily on the results of these new technologies and their potential for violating human rights or excluding certain groups of the population from the enjoyment of their rights. However, there was no mention of HRBA compliance in the design of these new technologies.

The remaining NHRIs have been reporting on the topic of the digitalisation independently of the response to COVID-19. Two of them explored the question of digitalisation through the lens of the private sector. The German NHRI mainly emphasised corporate responsibility in the design of new digital tools, the potential impact of algorithms on human rights and, more generally, the risks associated with biased analytical software.[25] The Dutch NHRI has been questioning the accessibility of private

---

15   See https://ec.europa.eu/digital-single-market/en/countries-performance-digitisation
16   See http://ennhri.org/news-and-blog/the-case-for-an-nhri-in-italy-presented-before-italian-law-makers/
17   See http://ennhri.org/news-and-blog/ennhri-advised-on-maltas-plan-to-establish-nhri/
18   See http://ennhri.org/rule-of-law-report/iceland/
19   See http://ennhri.org/news-and-blog/ennhri-comments-on-the-proposal-for-the-establishment-of-an-nhri-in-sweden/
20   For more information, see excel table (annex).
21   See http://ennhri.org/rule-of-law-report/france/
22   See https://www.equalityhumanrights.com/en/our-work/news/preventing-health-crisis-becoming-justice-crisis
23   See http://ennhri.org/rule-of-law-report/luxembourg/
24   See https://www.scottishhumanrights.com/media/2028/contact-tracing-briefing-180520-final.pdf
25   See https://www.institut-fuer-menschenrechte.de/fileadmin/user_upload/Publikationen/ANALYSE/Analysis_Business_and_Human_Rights_in_the_Data_Economy.pdf

websites for individuals with disabilities.[26] These questions are crucial and must be considered when designing new digital tools, precisely because state institutions tend to cooperate with big technology companies for the design of state services.[27] The relationship between algorithms and human rights will be discussed in further detail in the following section.

The NHRIs in Belgium[28] and Finland[29] looked at the issues surrounding the digital divide (between people of different ages and socio-economic backgrounds and people with disabilities). The Portuguese NHRI has been addressing the issue of eService access through a digital signature as well as access to bank-services.[30]

Estonia[31] and Denmark[32] are among the leading EU member states in terms of the digitalisation of public services. The Estonian NHRI has been confronting the issue of eService accessibility, particularly for people with visual impairments.[33] The Danish NHRI is behind some of the core principles and approaches to HRBA, advocating for its use in the implementation of the state's digitalisation agenda.[34] It has also been promoting the creation of an action plan to improve the public's digital

skills through targeted public awareness campaigns for various groups regarding digital behaviour and rights as well as for instruction at primary and secondary schools.[35] The contributions of the Estonian NHRI in relation to online voting for the 2019 elections is also worth mentioning. Indeed, the Chancellor for Justice (Estonian NHRI) received multiple requests from citizens to explain whether online voting in the country respected the principle of secrecy. The Chancellor for Justice explained that the procedure was encrypted, which would ensure that voters would remain anonymous. Moreover, there was a question whether some people had been unjustly denied their right to vote due to outdated residence data.[36] This is a good example of how an NHRI can manage individual complaints and inform citizens of their rights in relation to digital public services.[37]

Finally, in its decision on 21 August 2019 the EU Ombudsperson issued a ruling on the failure of the European anti-fraud office (OLAF) to make the fraud notification form on its website accessible for individuals with visual impairments.[38] Following complaints to the EU Ombudsperson from a Hungarian citizen who was unable to report a case of fraud due to his visual impairment, the

---

26   Annual Report for the year 2018, p. 25.
27   For more information about the undergoing challenge in international law to hold private actors responsible for human rights violations, see https://conflictoflaws.net/2020/corporate-responsibility-and-private-international-law/
28   See https://www.luttepauvrete.be/wp-content/uploads/sites/2/2019/12/Durabilite-et-Pauvrete-Rapport-bisannuel.pdf
29   See https://humanrightscentre-fi-bin.directo.fi/@Bin/d0c0cc8d2f2b5262be0fb29bb31002c0/1595364222/application/pdf/751647/IOK%20TOKE%202019%20EN.pdf
30   See https://www.provedor-jus.pt/site/public/archive/doc/Relat2019_web.pdf
31   See https://ec.europa.eu/digital-single-market/en/scoreboard/estonia
32   See https://ec.europa.eu/digital-single-market/en/scoreboard/denmark
33   See http://ennhri.org/rule-of-law-report/estonia/
34   See https://www.humanrights.dk/sites/humanrights.dk/files/media/annual_report_2019_web.pdf
35   See https://www.humanrights.dk/sites/humanrights.dk/files/media/dokumenter/udgivelser/annual_report_imr_beretning_uk_web.pdf
36   Voting lists are made based on registration. The Chancellor didn't find any violation of the right to vote, see http://ennhri.org/rule-of-law-report/estonia/
37   Already in 2005, the Constitutional Review Chamber of the Supreme Court of Estonia was approached to rule on the online voting system. It ruled that such a system did not violate the principle of uniformity and respected democratic thresholds.
38   See https://www.ombudsman.europa.eu/en/decision/en/118138

Ombudsperson contacted OLAF. The office answered that it was *'fully committed to improving the accessibility of its website and online tools to users with disabilities'* and that it was currently '*in the process of testing an alternative technical solution. (…) If this solution proves successful, OLAF will undertake to implement it on all online forms in the coming months.'*[39]

> *Values such as dignity, choice, self-respect, autonomy, self-determination and privacy are all traded off without being factored into the overall equation, all but guaranteeing that insufficient steps will be taken to ensure their role in the new digital systems.*
>
> Special Rapporteur on poverty and human rights (Philip Alston), p. 19, available at: https://undocs.org/a/74/493

As this review suggests, several NHRIs have been working on issues related to the digitalisation of state services. Most have been reporting on risks or actual human rights violations. There have been a few examples of NHRIs addressing individual complaints that concern the accessibility of digital state services. However, there has been no mention in annual reports of the involvement of NHRIs in the design or development of the digital tools necessary for a state's transition to e-governance. Collaboration with an NHRI from an early stage is crucial as it promotes the inclusion of all groups of citizens, including the most vulnerable, in the design of new digital tools. It ensures that states respect, protect and fulfil their international, national and regional human rights obligations – namely, the realisation of their citizens' socio-economic rights.

---

39   See https://www.ombudsman.europa.eu/en/decision/en/118138

Section two:

A wide-angle lens. Artificial intelligence, algorithmic systems and human rights

One of the signs that an area of human relations is experiencing rapid change is the emergence of academic studies and, at later stages, state-generated regulations. When governments initiate the development of new laws and standards, one can safely assume that they are doing so in an area where many innovative ideas have already hit the ground running. In such cases, initial progress has likely already been made, and authorities are rushing to catch up with regulations. An explosion in the number of academic studies and government policy instruments (both prescriptive and visionary-strategic) over the past 3–5 years regarding the intersection of human rights and information technology, with a considerable focus on systems for fully-automated or assisted decision-making, suggests that the area is booming with innovation and requires further investigation. The last CAHAI meeting (ad hoc Committee for Artificial Intelligence) particularly emphasised the importance of developing a co-

regulatory approach to AI regulation. The evolving nature of AI requires flexible regulations – a strictly binding instrument to establish a general legal framework should be combined with detailed rules, in addition to non-binding sectoral instruments.[40]

While we will be primarily focusing on the European experience, vital developments are emerging in all corners of the world. These considerations will be presented where relevant and possible to enrich the picture and present possible ideas for further UNDP programming in Ukraine or the ECIS region.

The following section will explore the ways in which governments, international institutions and CSOs are approaching the latest technological developments. It will also propose forecasts and likely developments in government-backed technology for the years to come.

---

40 See CAHAI meeting report, 6–8 July 2020 at: https://rm.coe.int/cahai-2020-12final-pv1-plen2-en/16809f490a

Core issues with AI and human rights.
A summary

There is a large body of research and analysis on the issues that could or already have gone wrong with the application of AI systems without the proper application of HRBA. While detailed resources for further reading are presented in the annotated bibliography, here we present some of the most pertinent concerns about human rights and AI as enumerated in the processed studies and reports. The details of this list as well as the information in the reports should not be regarded as a recommendation to refrain from useful features of AI but rather as a call to strengthen oversight and remain critical when weighing the potential tradeoffs linked to the deployment of such systems. Worldwide, AI systems have been found to reproduce the discrimination and biases inherent in the original data that was initially 'fed' to the system. Negative outcomes were also highly prevalent when datasets lacked information regarding certain demographics – people with disabilities, undocumented people and underprivileged groups have been found to be especially vulnerable. Biases could also be inherent in the process of selecting the parameters on which a system should deliver. Even 'innocent' variables, such as postal codes, have been known to produce discriminatory results (as this indicator is a proxy for the financial well-being of a person – that is, the home they can afford).

Many complex AI systems are self-taught and, therefore, operate on models that are inexplicable by human logic. In addition, particular judgements may be difficult to retrace to determine the reasoning behind those decisions. This violates the principle of decision 'explainability' and fair trial. Moreover, without the proper adjustments, AI makes projections for the future based on data from the past, thereby perpetuating the social situation as it is today. While the percentage rates of erroneous false positives or false negatives may seem negligible in well-calibrated systems, the absolute numbers mean that thousands of people will feel the impact. This is especially grave in situations such as 'AI assisted' court verdicts for essential welfare allocation or asylum. Finally, with the proliferation of AI into so many spheres of life, there are potentials to violate almost every possible human right, from the right to privacy, non-discrimination and a fair trial to the right to life in situations of automated combat weapons, such as military drones.

Other core issues include impunity toward private actors involved in the design of these new technologies and violations of the rights to privacy in relation to the use of private data.

However, it is important to first make note of the terminology that will be used throughout this section. The broad subject area that encompasses machine-operated or machine-assisted systems, in addition to the fields of knowledge that make such operations possible, is not uniform. Moreover, there are no universally adopted terms for any of these areas – that is, 'artificial intelligence'. A classification made by the Access Now CSO[41] is useful in this respect to demonstrate that electronic services delivered by the state may be founded on a plethora of technological tools.

One important terminological distinction is necessary before proceeding with the findings. *Algorithmic decision-making* may or may not be based on the foundations of *artificial intelligence* (or AI). For example, decisions regarding pension sizes or the approval of small business registrations (as electronic public services) need not employ complicated and costly AI products. These systems may operate based on human-created multi-variable decision tree algorithms (de facto, equations where all of the elements are known). As such, the verdict of any case may be retraced to understand, through human logic, why a particular conclusion was made and, if necessary, identify the errors in the mathematical model and alter the outcome for the appealing rights-holder.
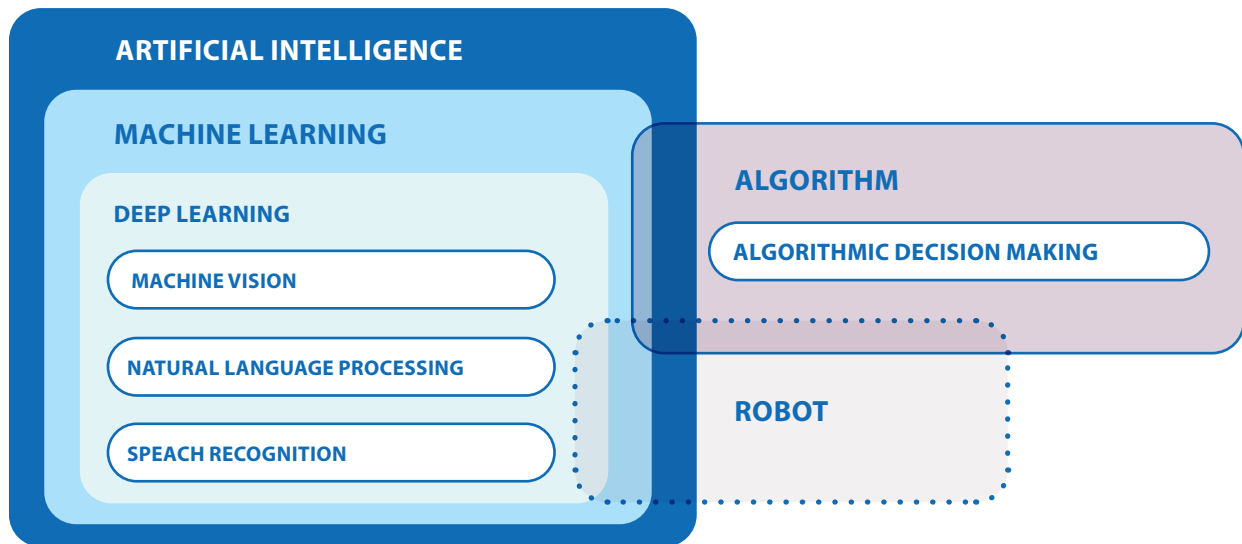

Figure. **The relationships of the concepts in artificial intelligence**

---

41   Please see Human Rights in the Age of Artificial Intelligence, a 2018 report at: https://www.accessnow.org/cms/assets/uploads/2018/11/AI-and-Human-Rights.pdf

By contrast, algorithmic decision-making based on AI is already being implemented in the commercial sector to decide on loans as well as by some law enforcement agencies for so-called predictive policing or related themes.[42] However, these technologies operate quite differently. The AI-based systems that either generate a full decision or use a system of flags to allow human operators to make the ultimate move are based on a technology known as 'deep learning'. One of the principles of this new and potentially revolutionising instrument is that after initial 'training' the machine can generate models *itself* based on identifying patterns in large massifs of data. The algorithms resulting from such an iterative analysis may *not be fully understood by engineers* (as the machine produces complex models itself, without human intervention).[43] Therefore, even if a rights-holder is successful in bringing a case to a human to appeal the result, there is no guarantee that one would be able to retrace the decision tree and understand why the machine came to that particular conclusion.[44]

The distinction between human-generated (algorithmic but not AI) and machine-generated (through deep learning) systems for decision-making is essential for contextualising the Ukrainian situation and offering insight into possible developments.

International and national governance actors reacting to the latest IT developments

One of the first global-level instruments to address the dilemma of human rights and the progress of IT systems was the 2012 summer Human Rights Council Resolution 'The promotion, protection and enjoyment of human rights on the Internet'. This landmark document affirmed that '[…] the same rights that people have offline must also be protected online […].'[45] While primarily aimed at the right to freedom of expression, the document opened the floodgate for transposing 'real-world' rights into the realm of IT system management. An intellectual follow-up to this document may be found in the essays posted on the UN Research Institute for Social Development (UNRISD) in 2018.[46] It should be noted that these rights do not only equate to online and offline protection, but that human rights should also be considered in the twilight zone of electronic instrument *creation, testing* and *access* on the part of rights-holders.

In the EU, the cornerstone regulation in this regard is Regulation (EU) 2016/679 'On the protection of natural persons with regard to the processing of personal data and on the free movement of such data', commonly

---

42 Please see comprehensive, in-depth critique in Bots at the Gate. A Human Rights Analysis of Automated Decision-Making in Canada's Immigration and Refugee System, 2018 by Toronto's International Human Rights Program (IHRP) at the Faculty of Law and the Citizen Lab at the Munk School of Global Affairs and Public Policy, with support from the IT3 Lab at the University of Toronto at: https://citizenlab.ca/wp-content/uploads/2018/09/IHRP-Automated-Systems-Report-Web-V2.pdf

43 Note the following consideration 'ML systems differ in a few key ways. First, whereas traditional statistical modelling is about creating a simple model in the form of an equation, machine learning is much more fine-tuned. It captures a multitude of patterns that cannot be expressed in a single equation. Second, unlike deterministic algorithms, machine learning algorithms calibrate themselves. Because they identify so many patterns, they are too complex for humans to understand.' Source, Human Rights in the Age of Artificial Intelligence, a 2018 report at: https://www.accessnow.org/cms/assets/uploads/2018/11/AI-and-Human-Rights.pdf p. 13.

44 See a more detailed and comprehensive discussion here https://www.wired.com/story/our-machines-now-have-knowledge-well-never-understand/

45 See https://digitallibrary.un.org/record/731540

46 See http://www.unrisd.org/unrisd/website/projects.nsf/(httpProjects)/21155E25524058A7C125824200305278?OpenDocument
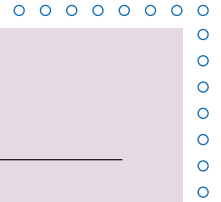
known as the 'General Data Protection Regulation', or GDPR, introduced in spring of 2016.[47] This foundational regulation determines how the personal data of rights-holders may be handled by duty-bearers (both state and private-sector) and introduces in Article 17 the 'right to erasure ('right to be forgotten')' – wherein rights-holders may request that their data be wiped from systems both public and private upon the wish of that person.

A separate group of recently introduced regulations confront issues of accessibility and regard to the rights of persons with disabilities. These are the dedicated Directive (EU) 2016/2102 'On the accessibility of the websites and mobile applications of public sector bodies'[48] and more general-application Directive (EU) 2019/882 'On the accessibility requirements for products and services'[49] with the accompanying 2018 Harmonised European Standard 'Accessibility requirements for ICT products and services'.[50]

The emergence and greater proliferation of AI additionally prompted the EU to produce the February 2020 White Paper 'On Artificial Intelligence – A European approach to excellence and trust'[51] that has, in spite of the COVID-19 pandemic, generated an enormous amount of interest in the CSO sector and caught the attention of national human rights institutions and equality bodies. Partially in response to this, the Council of Europe (CoE) released its April 2020 Recommendation CM/Rec(2020)1 'On the

Automated systems at work – an example from Sweden

Social benefits – the Trelleborg model

Since 2017 Trelleborg (city in the South of Sweden) has automated parts of its decision-making when it comes to social benefits. New applications are automatically checked and cross-checked with other related databases (for example, the tax agency and unit for housing support). A decision is automatically issued by the system. The number of caseworkers has been reduced from 11 to 3, and the municipality argues that they have considerably reduced the number of people receiving social benefits. They have been heading a pilot project to export their automation model to 14 additional municipalities and have received several innovation prizes. However, applicants and citizens have not been explicitly informed about the automation process. During the implementation process in another municipality, more than half of the caseworkers left their jobs in protest.

** Quoted from 'Report: Automating Society.*
Sweden' at https://bit.ly/2DZMnP7

---

47   See https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32016R0679
48   See https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32016L2102
49   See https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32019L0882
50   See https://www.etsi.org/deliver/etsi_en/301500_301599/301549/02.01.02_60/en_301549v020102p.pdf
51   See https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf
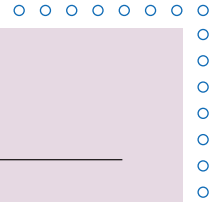
human rights impacts of algorithmic systems'[52] informed by and linked to the CoE thematic report 'Discrimination, artificial intelligence and algorithmic decision-making'.[53] It is common practice for the CoE to issue practical guides for the implementation of its recommendations, and such a publication should be anticipated in the near future. Additionally, the OECD has been developing principles for the responsible stewardship of trustworthy AI, with a focus on inclusive growth, sustainable development and well-being, human-centred values and fairness, transparency and explainability, robustness, security and safety, accountability. National policies and international cooperation for trustworthy AI are currently receiving substantial support, particularly where it involves investment in AI research and development, fostering a digital ecosystem for AI, providing an enabling policy environment for AI, building human capacity and preparing for labour transition. This continued cooperation on policies being developed by Council of Europe member states is critical in moving forward.

The United Nations High Level Panel on Digital Cooperation[54] and UNESCO are continuing to work on the issue of artificial intelligence and look for possible synergies with the work of CAHAI.[55] Following a decision by the UNESCO General Conference at its 40th session in November 2019, UNESCO is currently in the process of preparing the first global standard-setting instrument on the ethics of artificial intelligence.[56] This inclusive and

Automated systems at work – an example from Sweden

Large algorithm failure in 2018

In 2018 Sweden was forced to reverse a complex digital system used by the Employment Service to communicate with jobseekers due to problems that resulted in as many as 15% of the system's decisions to likely be incorrect.

* Prooflink 'Sweden: Rogue algorithm stops welfare payments for up to 70,000 unemployed' at https://bit.ly/32TdZjo

multidisciplinary process will incorporate consultations with a wide range of stakeholders, including the scientific community, people of different cultural backgrounds and ethical perspectives, minority groups, civil society, government and the private sector.

Additionally, the Secretary General plans to establish a Global Advisory Body on AI[57] aimed at creating a diverse, inclusive and informed platform to highlight and support ongoing work as well as connect the dots for global cooperation on a standard of AI that is trustworthy, human rights-based, safe, sustainable and promotes

---

52 See https://search.coe.int/cm/pages/result_details.aspx?objectid=09000016809e1154
53 See https://rm.coe.int/discrimination-artificial-intelligence-and-algorithmic-decision-making/1680925d73
54 See https://www.un.org/en/digital-cooperation-panel/
55 See https://www.un.org/en/content/digital-cooperation-roadmap/assets/pdf/Roadmap_for_Digital_Cooperation_EN.pdf
56 See https://en.unesco.org/generalconference/40/results
57 See https://dig.watch/updates/un-secretary-general-establish-ai-advisory-body

peace. This non-normative advisory body will strive to construct, connect, exchange and share expertise and make it available to those who need it most.

Individual states, both in Europe and across the world, have also made progress in the development and adoption of their own strategies for harnessing the power of AI in the government sector and Public-private partnerships. There are currently close to 30 national strategic policy documents internationally, with the world's first AI National Strategy appearing in Canada in 2017.[58] Moreover, that number is expected to grow exponentially.[59] Sweden[60] and Denmark[61] adopted similar documents in 2018 and 2019, respectively. Across the countries of the Eastern Partnership (EaP), only Ukraine and Azerbaijan have initiated policy work on AI. The former has recently finished a series of relevant public policy consultations.[62]

In 2017, Oxford Insights created the world's first Government AI Readiness Index to answer the question: How well-placed are national governments to take advantage of the benefits of AI in their operations and delivery of public services? The results sought to capture the current capacity of governments to utilise the innovative potential of AI.[63]

NHRIs and equality bodies uniting over concerns about artificial intelligence

While not yet reacting individually, NHRIs and European Equality Bodies initiated the joint June 2020 appeal to the European Commission[64] (ENNHRI) and authored the 2020 report 'Regulating for an equal AI: a new role for equality bodies. Meeting the new challenges to equality and non-discrimination from increased digitisation and the use of Artificial Intelligence' (EQUINET).[65]

Civil society and academia presenting their considerations regarding both accessibility and artificial intelligence/ algorithmic systems

Civil society and academia have not been observing these developments quietly. Instead, they been researching relevant themes and advocating their opinions more intensively in the past several years than ever. The 2018 report 'Plug and pray? A disability perspective on artificial intelligence, automated decision-making and emerging technologies' by the European Disability Forum[66] is of

---

58   See https://www.investcanada.ca/programs-incentives/pan-canadian-ai-strategy

59   See https://www.gp-digital.org/wp-content/uploads/2020/04/National-Artifical-Intelligence-Strategies-and-Human-Rights—A-Review_April2020.pdf for the April 2020 review of existing today National AI strategies and their human rights components.

60   See https://www.government.se/491fa7/contentassets/fe2ba005fb49433587574c513a837fac/national-approach-to-artificial-intelligence.pdf

61   See https://en.digst.dk/media/19337/305755_gb_version_final-a.pdf

62   See https://thedigital.gov.ua/news/doluchaytes-do-obgovorennya-kontseptsii-rozvitku-sferi-shtuchnogo-intelektu-v-ukraini

63   See https://www.oxfordinsights.com/ai-readiness2019

64   See http://ennhri.org/wp-content/uploads/2020/06/ENNHRI-letter_White-Paper-AI.pdf as both the reaction to the white paper and summary of main risks with AI systems as seen by the European Network of National Human Rights Institutions

65   See https://equineteurope.org/wp-content/uploads/2020/06/ai_report_digital.pdf by the European Network of Equality Bodies

66   See http://www.edf-feph.org/sites/default/files/edf-emerging-tech-report-accessible.pdf

particular note as it illustrates the reaction of groups of CSOs regarding issues of technology and the related challenges faced by people with disabilities. Moreover, the June 2020 report 'Recommendations for a Fundamental Rights-based Artificial Intelligence Regulation Addressing collective harms, democratic oversight and impermissible use' by the European Digital Rights CSO[67] highlights important considerations and warnings against the uncritical adoption of AI as technology.

In terms of digital welfare policy, several conclusions emerge. First, there should always be a genuine, non-digital option available. Second, programmes aimed at digitizing welfare arrangements should be accompanied by programmes designed to promote and teach the digital skills needed and to ensure reasonable access to the necessary equipment as well as effective online access. Third, to reduce the harm caused by incorrect assumptions and mistaken design choices, digital welfare systems should be co-designed by their intended users and evaluated in a participatory manner.

Special Rapporteur on poverty and human rights (Philip Alston), p. 16, available at: https://undocs.org/a/74/493

In summary, the area of artificial intelligence for use by public authorities, including for prediction-making, analysis or even the preparation of decisions for public officers on the individual cases of rights-holders, is on the rise. This trend is going to continue and has been especially exacerbated by the COVID-19 pandemic. In this respect, it is essential to continue contributing to the global dialogue and trace developments in the ways that academia, civil society and government watchdogs, including NHRIs and equality commissions, play a role in the unfolding digital transformations and human rights debates.

---

67   See https://edri.org/wp-content/uploads/2020/06/AI_EDRiRecommendations.pdf

Section three:

Ukraine's contextual gaps – suggestions and recommendations

The Ukrainian governance reform agenda has experienced radical change between late 2013 and 2020. Nevertheless, one premise has remained consistent. Previous and current administrations have been keen on offering digital (including mobile-based) solutions as a means to reduce direct contact between citizens and authorities. One of the more enduring traits has also been the 'everyone sees everything' model of digital transparency and data disclosure for reducing corruption at all levels.

The post-Maidan government took a significant step in this regard by launching the State Agency for Electronic Governance of Ukraine in the summer of 2014. Later, in 2019, the Agency morphed into the Ministry for Digital Transformation of Ukraine (MDT) with a broader mandate. The Cabinet of Ministers programmes shaped by both the Honcharuk and Shmyhal Governments have heavily emphasised digital transformation. Notably, both programs have outlined almost identical priorities for this area of concern: the development of high-speed internet infrastructure, a rapid expansion in the number and thematic scope of public services offered to citizen-clients, the promotion of digital skills across the nation and the further enhancement of Ukraine's digital economy.

Ukraine does possess a number of very advanced systems, including its ProZorro model and software for electronic procurement as well as its DoZorro platform, which has initiated the application of AI in the form of 'supervised learning' algorithms. Transparency International (TI) Ukraine claims that the use of such technologies for public procurement monitoring may be internationally unprecedented.[68] However, it should be noted that DoZorro is not a public initiative but a civic initiative run by TI Ukraine.

---

68   See https://backend.dozorro.org/storage/app/media/story%20eng%20print%201.pdf

Other systems worth mentioning include the national data-exchange hub Trembita, as well as – more recently – the Diia mobile application and portal.

A majority of systems that have been designed to date, including those that provide services to the rights-holders, have been created on the basis of engineer-specified multi-variable logic and decision trees based on user actions in highly restrictive environments. For their part, on 21 May 2020 MDT announced consultations over a draft National AI strategy. As such, rapid action is required to ensure that the ideas and considerations already researched by the global community are taken into account as Ukraine charts its way forward.

Below are recommendations that have been enumerated for consideration both by UNDP Ukraine and by the future team of the Digital, Inclusive, Accessible: Support to Digitalisation of Public Services in Ukraine (DIA Support) project. These recommendations are based on the findings of multiple reports, which are provided in the annotated bibliography.

For the UNDP Ukraine Country Office team

- Explore the feasibility of designating a team or focal point within the organisation (not a consultant but a long-term employee) to monitor developments within the rapidly changing area of digitalisation and human rights.

This person or team could also assist the organisation in the implementation of the current UNDP Digital Strategy and its follow-on plans after 2021 and serve as the knowledge hub for issues associated with the latest technological developments and their impact on the developing world.

- Consider identifying and partnering with Ukraine's most advanced technical universities in the areas of computer science and mathematics to secure access to subject-matter knowledge from both the private sector and academia. Consider piloting a curriculum with any of these academic institutions to promote the longer-term digitalisation agenda as pursued by UNDP.

- Look into implementing dedicated initiatives (either as part of the current project portfolio or as quick-response ad hoc assistance) to amend Ukraine's equality, non-discrimination and human rights legislation to guarantee that issues arising from digital transformation are covered by relevant regulation.

- Explore opportunities to integrate issues of digital transformation, gender and human rights into projects, where possible. For this, consider developing an internal office manual for 'Planning, design and launch of IT systems in projects – Ukrainian regulations, essential considerations and the human rights framework'. Supplement the manual with an HRBA checklist.

- Suggest the creation of guidelines for government contracting with private sector vendors for software and service development to ensure the accountability and transparency of relations and the resulting product.

- Conduct an internal exercise to catalogue and extract lessons from previously implemented IT instruments, as assisted by UNDP.

- Offer to establish deeper and practical partnerships with other bodies that work on issues of human rights and digitalisation. These bodies may include the Council of Europe, its dedicated Ad Hoc Committee on AI (CAHAI) and experts featured on the CoE roster.

For the DIA Support Project team

- Proactively seek to incorporate human rights considerations into the regulations required to launch the project's dedicated services as well as any policies that do not directly relate to the project target services but may need to be enhanced with HRBA considerations.

- In this respect, consider running a full-scale human rights impact assessment (HRIA) of the MDT-proposed National Strategy on Artificial Intelligence, strengthened with human rights

considerations. If developments with the draft Strategy start to gain momentum prior to the Project launch, UNDP should consider deploying fast-response expertise to prevent the draft from being adopted without any human rights screening. A template screening mechanism may be found here: https://bit.ly/3jmifxC from page 33 and onward.

- More widely, and as resources allow, consider analysing the wider Ukrainian legislative context to ensure that the development of digital tools for state services are consistent with human rights principles. Consider exploring the theme more widely by incorporating situations where businesses come into play and may influence the way that services are designed. In this case, the UN Guiding Principles on Business and Human Rights should be incorporated in the regulatory framework.[69]

- Conduct a human rights and gender analysis to assess potential risks for human rights violations and vulnerable groups. It is crucial to include a gender analysis, given that women and girls do not always receive equal access to digital tools.[70]

- In accordance with the more general recommendation for UNDP Ukraine, establish a wide network of connections to academia and civil society organisations in the areas of digital transformation and human rights. Establish a

---

69   See https://www.ohchr.org/Documents/Publications/GuidingPrinciplesBusinessHR_EN.pdf

70   For example, in Roma communities, A study conducted by UNDP in 2018 showed that in some Roma communities most women do not possess a cell phone.
      See, p. 10: https://www.eurasia.undp.org/content/rbec/en/home/library/roma/nowhere-to-turn-gbv-against-roma-women.html

practice of regular, even if unofficial, brown-bag lunch meetings to discuss the agenda, share news and assess approaches.

- A cursory overview revealed that the Swedish human rights actors did not possess significant experience with the crossover between IT and human rights. Nevertheless, other government actors may be beneficial for exchanging experience, such as the Committee on Coordinated and accelerated policy development linked to fourth industrial revolution technologies (https://bit.ly/3fJRUr1) and the Ministry of Enterprise and Innovation, which was responsible for the national approach to artificial intelligence. The Raoul Wallenberg Institute of Human Rights and Humanitarian Law may also offer a number of important connections and provide access to valuable expertise (https://rwi.lu.se).

- Considering the strong human rights component of the Danish National Strategy for Artificial Intelligence (https://bit.ly/2OETsqn), an effort should be made to strengthen connections with the Danish Institute for Human Rights.

- The overview also suggests that Finland has made great strides in digital transformation, arguably on the same level as Estonia, considering that 'According to the fifth "digital barometer" (2018), Finland is one of the leading countries in utilising digital technology at three main levels (preconditions, usage and effects) and sectors (business, citizens and public). The Finnish government also actively supports digitalisation with so-called key projects that are based on the government programme. An example of the aims linked to digitalisation is experimentation and deregulation to digitalise public services.'71 The project should further 'diversify' their sources of good practices and inspiration when designing eServices.

- Where possible, the project should serve as an internal UNDP Ukraine leader for championing HRBA IT assessments by enhancing its own in-house capacity or collaborating with sister-projects.

- The project team should iterate anticipated products with target groups as often as possible. Consider establishing a consultative unpaid oversight board that represents different rights-holder groups to ensure that the project's products are externally reviewed and validated.

- Build a robust connection with the Office of the Ukrainian Ombudsperson and involve them as a partner wherever possible to help the institution stay on top of the combined digital transformation and human rights agenda. Work on a memorandum of understanding between the Ombudsperson in Ukraine and MDT to include the NHRI in the early stages of designing/developing digital tools—not only under the DIA Support project.

---

71 See https://vm.fi/en/public-services-will-be-digitalised

Annotated bibliography

Reports and publications

European Digital Rights (June 2020).
Recommendations for a Fundamental Rights-based Artificial Intelligence Regulation.
Addressing collective harms, democratic oversight and impermissible use.
https://edri.org/wp-content/uploads/2020/06/AI_EDRiRecommendations.pdf

This paper provides a critique of the earlier-published EU whitepaper on Artificial Intelligence. While EDR does not oppose the use of AI, it points out that governments in the new AI arms race often rush to develop and implement AI solutions into practice. The paper decries the lack of transparency of governments that decide to use such systems without public consultation and the opaque approaches of technology companies that produce AI systems. EDR goes on to describe the various issues that AI implementation must confront to serve the common good: ensure proper data protection, address concerns of built-in or resulting inequality or discrimination, prevent the misuse of AI for stifling democratic controls, prevent the abuse of AI against the rights to freedom of expression, guarantee proper procedural rights and true access to justice (including in cases related to migration control). The authors also promote the Human Rights Based Approach as the foundation for the EU's update of the whitepaper and further drafting of regulations. The paper concludes in listing recommendations for EU policymakers, including the development of clear criteria for the legality of AI, a ban on uses of AI that could encroach upon human rights (for example, indiscriminate biometric surveillance, predictive policing, autonomous lethal weapons), the avoidance of any exemptions to human rights vis-à-vis AI, the introduction of mandatory human rights impact assessments, implementation of a clear oversight and enforcement model, including the launch of National Centres of Expertise on AI with strong engagement and involvement with civil society, equality bodies and human rights.

Equinet European Network of Equality Bodies (June 2020).
Regulating for an equal AI: a new role for equality bodies.
https://equineteurope.org/wp-content/uploads/2020/06/ai_report_digital.pdf

This report focuses on the challenges that European Equality Bodies face to ensure that the principles of equality and non-discrimination are respected in their respective states. In particular, it addresses the role of the above-mentioned institutions and the consequences of digitalisation on inequality in relation to the development of new technologies using AI, machine learning (ML) and automated decision-making (ADM). The report seeks to answer six main questions:

◆ What tasks should Equinet's Members undertake to ensure that AI, ML and ADM advance do not facilitate inequality and discrimination?

◆ What capacity do Equinet's Members have for this?

◆ How can they be assisted to gain better capacity?

◆ Who or what are the other actors in this field with whom Equinet's Members should be working?

◆ Does the current discourse on the ethical approach to AI support legal rights to equality?

◆ How well do the other regulatory tools available to states work with equality rights in the context of AI?

The report stresses that NEB should serve as key points of reference about AI systems and their impact on equality and non-discrimination for all actors and stakeholders involved. It explains the response of Equinet's Members to new challenges that arise from AI systems and refers to legal resources to emphasise how the existing equalities framework can be applied to tackle discrimination resulting from such AI-related challenges.

European Commission (February 2020).

White Paper on Artificial Intelligence: A European approach to excellence and trust.

https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf

The White Paper on Artificial Intelligence was launched by the European Commission as part of its digital strategy for the next five years. This paper recognises the opportunities that AI represents for Europe and shares the Commission's view on how to coordinate and promote the use of AI. The paper qualifies AI as risky if it meets two cumulative criteria: if it is used in a high-risk sector and if that use could potentially raise significant risks for individuals or companies, in particular, from the viewpoint of safety, consumer rights or fundamental rights. However, some sectors would still be qualified as high-risk, even if they do not meet the above-mentioned criteria. This would, for example, include the use of AI in a recruitment/employment context, and in remote biometric identification systems.
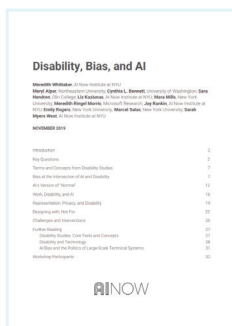
AINow (November 2019). **Disability, bias and AI.**

https://ainowinstitute.org/disabilitybiasai-2019.pdf

This report explores the intersection between AI and disability. It raises awareness about the increasingly important role of AI in core social domains, such as eligibility for social benefits. The report argues that AI is not intended to take marginalised groups into account – namely, people with disabilities, and warns of the dangers surrounding an 'AI version of normal'. If facts that are open to interpretation (nuanced or complex matters) are decided by AI, there will no longer be space for flexibility and adaptability. For example, the notion of disability is not clearly defined and will vary depending on other personal factors and context. If not carefully designed, AI runs the risk of reproducing pre-existing societal inequalities, stereotypes and discrimination, or so-called bias. The report concludes that in order to tackle the core justice issues that relate to the proliferation of AI systems people with disabilities and other vulnerable groups need to be at the centre of any approach and define the terms of engagement and priorities.
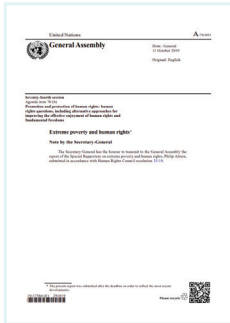
Special Rapporteur on poverty and human rights (October 2019).
Thematic report on digital welfare states and human rights.
https://undocs.org/A/74/493

This report seeks to raise awareness about the risks of emerging digital welfare states to human rights. Despite shortly acknowledging that digitalisation offers numerous advantages, it argues that many issues must be addressed to avoid a 'digital welfare dystopia'. The report criticises leading technology companies, which are often in charge of designing, constructing and operating important aspects of the digital welfare state, for not following international human rights law. The report also notes the absence of a national legal framework and the exacerbation of major disparities among different groups of the population with regard to the access and skills necessary to benefit from digital services. Finally, it addresses the shifting paradigm that arises from the fact that when requesting the realisation of their socio-economic rights individuals are no longer regarded as human-rights-holders but as 'applicants' in a rigid process that is governed by machines instead of human-beings.
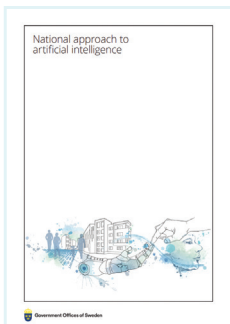
Government Offices of Sweden (February 2019).
National approach to Artificial Intelligence.
https://www.government.se/491fa7/contentassets/fe2ba005fb49433587574c513a837fac/national-approach-to-artificial-intelligence.pdf

The National approach to artificial intelligence is a report that aims to identify an overall direction for AI-related work in Sweden and lay down the foundations for future priorities. Sweden seeks to be a world leader in the field of digitalisation. For this to occur, the government of Sweden has enumerated a number of recommendations, including educating and researching; innovating and using AI; and regulating AI.
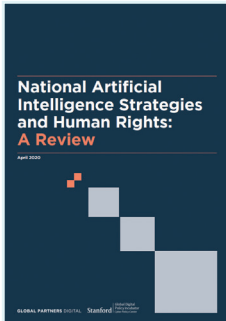
Global Partners Digital/Stanford (April 2020).
National Artificial Intelligence strategies and human rights: a review.
https://fsi-live.s3.us-west-1.amazonaws.com/s3fs-public/national_artifical_intelligence_strategies_and_human_rights-a_review1.pdf

This report seeks to understand how/if human rights have been incorporated in National Artificial Intelligence Strategy documents. It then proposes recommendations on how to include such considerations in the future. Although most of the strategies studied make explicit reference to human rights in their texts, the depth of these references varies. Many strategies refer to the need for ethical frameworks. Some insist on specific human rights (often the right to privacy) and suggest a form of prioritisation. Others do not mention human rights but still engage with issues of human rights, such as the right to work. Most strategies were quite vague on how to protect human rights. All in all, the report states that very few NAS engage with the human rights aspect of AI. This lack of engagement may be because some states do not prioritise the protection of human rights in their policymaking. In other cases, governments may find it difficult to conciliate human rights standards with other government goals for AI, such as economic or geopolitical competitiveness, despite acknowledging the importance of human rights. The question of new ethical frameworks for the governance of AI can sometimes act as an attempt to circumvent human rights frameworks. However, it can also be an attempt to exceed the human rights frameworks. Finally, the report also mentions that even if the strategy refers to the importance of human rights the actual implementation of that strategy must be considered to assess a state's commitment to respecting human rights in practice.
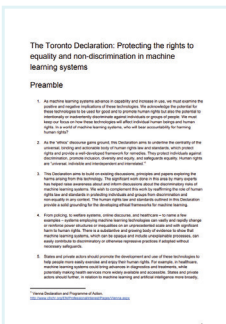
Access Now/Amnesty International (May 2018).
The Toronto Declaration: protecting the rights to equality and non-discrimination in machine learning systems. http://www.intgovforum.org/multilingual/sites/default/files/webform/toronto-declaration-final.pdf

The Toronto Declaration is a statement led by Access Now and Amnesty International that urges governments and companies to respect and protect human rights and, in particular, the right to protection from discrimination in reference to machine learning, artificial intelligence and advanced computing. Building on the existing dialogue surrounding artificial intelligence and ethics, this statement reiterates the applicability of human rights law, the necessity of human-centred technology and the responsibility of private actors to respect human rights. The declaration has been endorsed by a variety of civil society organisations, non-governmental organisations and research groups.
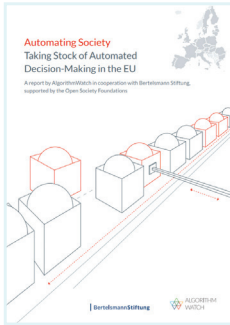
AlgoruthmWatch/BertelsmannStiftung (January 2019).
Automating society, taking stock of automated decision-making in the EU.
https://algorithmwatch.org/wp-content/uploads/2019/01/Automating_Society_Report_2019.pdf

This report is oriented around four main questions. First, it looks at societal discussions surrounding automated decision-making. For this purpose, it compiles findings from 12 EU members states and the EU. It presents regulations, legislation and various organisations and institutions that are addressing this issue. Second, it examines the existing regulatory proposal, discussing the full range of possible governance measures in addition to laws. The third stage focuses on the existing oversight bodies and mechanisms that consider the question of which sectors and processes they are responsible for and various approaches to the task. Finally, the report considers existing automated decision-making systems and poses the question of who oversees the development and use of those systems – the public sector or private companies?
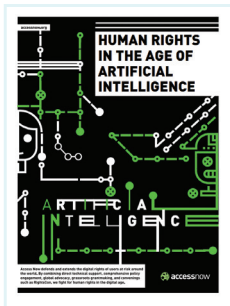
Accessnow (November 2018). **Human rights in the age of Artificial Intelligence.**
https://www.accessnow.org/cms/assets/uploads/2018/11/AI-and-Human-Rights.pdf

This report provides a distinction between 'helpful' AI and 'harmful' AI. Examples of helpful AI functions include improving access to healthcare and predicting disease outbreaks; making life easier for the visually impaired; optimising agriculture; mitigating climate change; making government services more efficient and accessible. Among the exploitative uses of AI on the part of governments or the private sector, the report mentions the perpetration of bias in criminal justice; the facilitation of mass surveillance; discriminatory profiling; the spread of disinformation; the perpetration of bias in the job market; the exacerbation of financial discrimination against marginalised populations. It stresses that two new trends distinguish AI from previously-implemented technology: big data and the rise of algorithmic decisions. The report then focuses on the nexus between human rights and AI and seeks to understand the importance of human rights and their impact on AI. It reiterates that because all human rights are interconnected all human rights are affected – not just the right to be protected from discrimination. The report then individually examines a series of human rights and highlights the possible impacts of AI on each of them and makes recommendations for preventing the unjust restriction of human rights for individuals and marginalised groups.
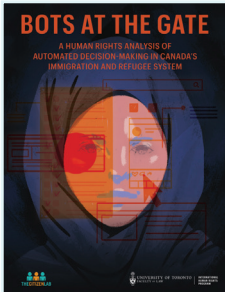
The Citizen Lab at the Munk School of Global Affairs and Public Policy/International Human Rights Programme at the University of Toronto/The information Technology, Transparency and Transformation Lab. (September 2018).
Bots at the gate – a human rights analysis of automated decision-making in Canada's immigration and refugee system.
https://citizenlab.ca/wp-content/uploads/2018/09/IHRP-Automated-Systems-Report-Web-V2.pdf

This report concentrates on the impact of automated decision-making on Canada's immigration and refugee system through a human rights lens. It argues that the application of AI to decision-making in the field of refugee law threatens to create 'a laboratory of high-risk experiments within an already highly discretionary system'. An automated decision-making process would reduce the complexity and nuance necessary when assessing the delicate subject of refugee claims. After providing insight into the immigration and refugee process in Canada, the report develops a human rights analysis of the use of automated decision systems from a national and international perspective. Without a human rights analysis, AI risks violating a number of human rights, including the right to be protected from discrimination, freedom of movement, expression, religion and association, privacy rights, the right to life, liberty and security. The possible impact of such systems on procedural fairness and standard of review requires consideration as well. Finally, the report addresses the question of access to justice and the question of public confidence in the legal system as well as the accountability of the private sector.
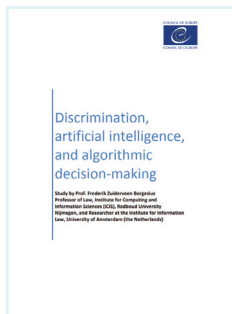
Council of Europe (January 2018).
Discrimination, artificial intelligence and algorithmic decision-making.
https://rm.coe.int/discrimination-artificial-intelligence-and-algorithmic-decision-making/1680925d73

This report, written for the Anti-discrimination Department of the Council of Europe, concerns discrimination caused by algorithmic decision-making and other types of artificial intelligence (AI). Despite recognising the positive aspects of AI (efficiency, health and economic growth), the report warns against potential negative effects, such as when AI systems adhere to discriminatory or mistaken human decisions.

AI can be utilised by both the public and private sectors. In the public sector, it can be used in predictive policing or for deciding the eligibility of individuals for payments, housing assistance or unemployment benefits. In the private sector, it can be used to select job applicants, while banks can use it to allocate consumer credits and set interest rates. The most effective tool for preventing AI discrimination is the adoption of a non-discrimination law and a data protection law. Cooperation between Council of Europe member states, human rights monitoring bodies (for example European Commission against Racism and Intolerance) and Equality Bodies is also essential for the better enforcement of current non-discrimination norms.

However, it is important to consider that, legally speaking, discrimination refers to unequal treatment based on a protected characteristic. Certain forms of differentiation could be unfair, yet legal, if they target particular groups. Hence, the report suggests creating additional regulations to protect fairness and human rights in AI. Nevertheless, the use of AI is so varied that there is little point regulating AI in general. Given that the particular values at stake are largely sector-dependent, sector-specific rules should be considered when applying regulations.
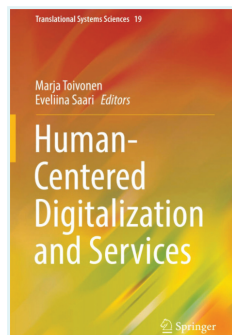
Marja Toivonen/Eveliina Saari (eds.). Springer (2019).
Human-Centered Digitalisation and Services.

This book is divided into four sections that are then divided into four chapters.

The first section (Theoretical perspectives on digitalisation and service innovation) proposes three topical approaches to analysing human-centred service intervention. The first chapter investigates the historical development of the service economy. The second chapter discusses human-centric service innovations in public services and highlights the experimental nature of these innovations. The third chapter questions the relationship between human and technological resources. Finally, the fourth chapter presents a human-centred co-evaluation method for the evaluation of service innovation in the context of digital services.

The second section (Approaches and case studies on human interaction in the service context) presents issues of interaction related to digitalisation. The first section studies the impact and change of digitalisation in everyday life (chapter 5). The second section tackles the digitalisation of services for refugees and immigrants (chapter 6).

The third section focuses on digital services for patients and the healthcare sector. It studies the organisation of healthcare in Lithuania and the way the patients' value co-creation is perceived (chapter 7). The last chapter of the second section analyses technological breakdowns and their potential for promoting learning and technological literacy (chapter 8). Section three of the book (Analysis of the new opportunities provided by digital solutions) provides examples of the application of digital tools. The first chapter of the third part presents the findings of a study that compares the Japanese and the Finnish elderly care service systems (chapter 9). The second part also focuses on elderly care but engages more specifically with the use of care robots and discusses their acceptability in this sector (chapter 10). The third part looks at experimental developments and analyses its application in a public sector case in a mid-sized Finnish city, more specifically a new integrated model of well-being that seeks to promote multi-professional collaboration and citizen empowerment in child and family services (chapter 11). The last chapter of the third section presents the digitalisation of healthcare services from the perspective of nursing in large hospitals (chapter 12).

Section four of the book (Understanding the interaction between digital and human resources) is about the human-centred approach to digitalisation. The first chapter of this section tackles data activism and argues that an efficient alternative in the current data economy, individualistic and humancentric data activism needs to become increasingly intertwined with social science perspectives (chapter 13). The second part explores digitalisation and social sustainability in the case of elderly care systems (chapter 14). The third chapter of this section discusses the transition to service digitalisation. It focuses on the future of employment and provides an analysis based on interviews with employees, management and individuals in human resource development (chapter 15). Finally, the last chapter of the book confronts issues surrounding artificial intelligence. The authors insist on the necessity of human-centric artificial intelligence for achieving long-term societal stability. They argue that it is crucial to maintain human control over artificial intelligence. Although AI may be considered one of the biggest challenges of our time, the authors still believe that it may be utilised to the benefit of society through sufficient political and social will.

Legal and Soft Regulation (advisory) Instruments

Council of Europe (8 April 2020).
Recommendation CM/Rec(2020)1 of the Committee of ministers on the human rights impacts of algorithmic systems. https://search.coe.int/cm/pages/result_details.aspx?objectid=09000016809e1154

The Committee of Ministers issued recommendations on the impact of algorithmic systems on human rights. These recommendations include a set of guidelines for addressing (A) the human rights impacts of algorithmic systems, (B) the obligations of states in respect of the protection and promotion of human rights and fundamental freedoms in the context of algorithmic systems and (C) the responsibilities of private sector actors in respect of human rights and fundamental freedoms in the context of algorithmic systems.

Each section (A, B, C) is then subdivided into a variety of topics. State obligations to address the impact of algorithmic systems on human rights include the principles of general application; data management; analysis and modelling; transparency, accountability and effective remedies; precautionary measures; research, innovation and public awareness.

Council of Europe (14 June 2017).
Recommendation CM/Rec (2017)5 of the Committee of ministers on standards for e-voting.
 https://search.coe.int/cm/Pages/result_details.aspx?ObjectId=0900001680726f6f

This recommendation targets the core aspects of e-voting and addresses election technology. It primarily discusses the use of electronic means to cast and count votes (systems such as Direct Recording Electronic (DRE), voting machines, ballot scanners, digital pens and internet voting systems). Perhaps of most interest, it contains recommendations for ensuring inclusion (Appendix I).

EU (27 April 2016).
Directive (EU) 2016/679 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data. https://eur-lex.europa.eu/eli/reg/2016/679/oj

Commonly known as the General Data Protection Regulation, or GDPR, this regulation determines the ways in which the personal data of rights-holders will be handled by duty-bearers (both state and private-sector) and introduces in article 17 the 'right to erasure ('right to be forgotten')' – wherein rights-holders may request that their data be wiped out from systems both public and private.

EU (26 October 2016).
Directive (EU) 2016/2102 on the accessibility of the websites and mobile applications of public sector bodies. https://eur-lex.europa.eu/eli/dir/2016/2102/oj

This directive seeks to ensure better access to websites and mobile apps for public services. In particular, it refers to specific standards to increase the accessibility of websites and apps, requires the publication of an accessibility statement that assesses the level of accessibility for each app and website, expects feedback mechanisms for users and the regular monitoring of public sector websites by Member States.

EU (17 April 2019).
Directive (EU) 2019/882 on the accessibility requirements for products and services. https://eur-lex.europa.eu/eli/dir/2019/882/oj

Known as the European Accessibility Act, this directive seeks to standardise accessibility requirements for certain products and services and clarifies existing accessibility obligations in EU law, particularly, in the field of procurement and structural funds.