May 2019 Submission to the UN Special Rapporteur on Extreme Poverty & Human Rights Regarding His Thematic Report on Digital Technology, Social Protection & Human Rights

Summary

This submission examines the human rights implications of Artificial Intelligence (AI) [1] and other data-driven technologies in welfare benefits programs, such as cash and food assistance programs. Through a series of case studies, this submission explains how States delegate key welfare functions, such as determinations of eligibility and benefits levels, to automated decision-making models, some of which rely on data mining, machine learning and other processes or technologies typically associated with the field of AI. It also assesses how automated decision-making interferes with the rights to privacy and social security, and the obligations of States to guarantee the exercise of these rights without discrimination and undue private interference.

Overview of Applicable International Human Rights Law

The right to privacy

Article 17 of the International Covenant on Civil and Political Rights (ICCPR), which derives from Article 12 of the Universal Declaration of Human Rights (UDHR), establishes the right to "the protection of the law" against "arbitrary or unlawful interference" with one's "privacy, family, home or correspondence."

---

[1] The UN Special Rapporteur on the right to freedom of opinion and expression has defined Artificial Intelligence (AI) as the "constellation of processes and technologies enabling computers to complement or replace specific tasks otherwise performed by humans, such as making decisions and solving problems." Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, David Kaye, Human Rights Council, A/73/348, Aug 29, 2018, ¶ 3 (internal quotation marks omitted), http://www.un.org/ga/search/view_doc.asp?symbol=A/73/348.
This is also the definition that Human Rights Watch uses in this submission.

The Human Rights Committee has concluded that the prohibition against "arbitrary or unlawful interference" establishes a two-part test. First, interferences with privacy can take place only "in cases envisaged by the law."[2] Under this requirement, States must "specify in detail" in relevant legislation "the precise circumstances in which such interferences may be permitted," and ensure that decisions "to make use of such authorized interference must be made only by the authority designated under the law, and on a case-by-case basis."[3] Second, for interferences to be non-arbitrary, the Committee has concluded that they must be "proportionate to the end sought, and … necessary in the circumstances of any given case."[4]

The right to social security

Article 9 of the International Covenant on Economic, Social and Cultural Rights (ICESCR) and Article 22 of the UDHR recognize the right of everyone to "social security, including social insurance." The Committee on Economic, Social and Cultural Rights (CESCR) has concluded that this right establishes the obligation of States to ensure that eligibility criteria for social security benefits are "reasonable, proportionate and transparent."[5] Furthermore, the "withdrawal, reduction or suspension of benefits should be "based on grounds that are reasonable, subject to due process, and provided for in national law."[6]

Access to information is also a precondition of the enjoyment of the right to social security. The CESCR has found that beneficiaries of social security schemes "must be able to participate in the administration of the social security system."[7] Accordingly, the system should "ensure the right of individuals and organizations to seek, receive and impart information on all social security entitlements in a clear and transparent manner."[8]

---

[2] Human Rights Committee, General Comment 16: Article 17 (Right to Privacy), ¶ 3, https://www.ohchr.org/Documents/HRBodies/TB/HRI-GEN-1-REV-9-VOL-I_en.doc.

[3] *Id.*, ¶ 8.

[4] *See e.g. Toonen v. Australia*, Communication No. 488/1992, U.N. Doc CCPR/C/50/D/488/1992 (1994), ¶ 8.4, http://hrlibrary.umn.edu/undocs/html/vws48Id.8.htm; *Antonius Cornelis Van Hulst v. Netherlands*, Communication No. 903/1999, U.N. Doc. CCPR/C/82/D/903/1999 (2004), ¶ 7.6, http://hrlibrary.umn.edu/undocs/html/903-1999.html.

[5] CESCR, General Comment 19: The right to social security (art. 9), Nov 23, 2007, ¶ 24, https://tbinternet.ohchr.org/_layouts/15/treatybodyexternal/Download.aspx?symbolno=E%2fC.12%2fGC%2f19&Lang=en.

[6] *Id.*, ¶ 24.

[7] *Id.*, ¶ 26.

[8] *Id.*

Non-Discrimination Obligations

Article 2(1) of the ICCPR and Article 2(2) the ICESCR require States to guarantee Covenant rights without discrimination of any kind based on "race, color, sex, language, religion, political or other opinion, national or social origin, property, birth or other status." Article 26 of the ICCPR additionally guarantees the right "to all persons equal and effective protection against discrimination on any ground." The Human Rights Committee has found that Article 26 establishes an "autonomous right" that "prohibits discrimination in law or in fact in any field regulated and protected by public authorities."[9]

In the context of social security, the CESCR has found that States should ensure that social security schemes "do not discriminate in law or in fact."[10] States should also "pay special attention" to those that disproportionately experience difficulties in accessing social security, such as women, people with disabilities, minorities and "casual" or "seasonal" workers.[11]

Obligation to Protect Against Third-Party Interference

States have a duty to protect individuals from undue third-party interferences with their rights to privacy and social security. The Human Rights Committee has concluded that Article 2(1) of the ICCPR, which establishes the State's duty to "ensure" the right to privacy and other Covenant rights, imposes a positive obligation to protect individuals against "acts committed by private persons or entities that would impair the enjoyment of [these] rights."[12] This duty requires the adoption of "appropriate measures" or "due diligence to prevent, punish, investigate or redress the harm caused by … private persons or entities."[13]

The CESCR has categorized State obligations under the ICESCR as obligations to respect, protect and fulfill Covenant rights.[14] The obligation to protect the right to social security requires States to

---

[9] Human Rights Committee, General Comment No. 18: Non-discrimination, Nov 10, 1989, ¶ 12, https://tbinternet.ohchr.org/Treaties/CCPR/Shared%20Documents/1_Global/INT_CCPR_GEC_6622_E.doc.

[10] CESCR, General Comment 19, *supra* n. 5, ¶ 30.

[11] *Id.*, ¶ 31.

[12] Human Rights Committee, General Comment No. 31 [80]: The Nature of the General Legal Obligation Imposed on States Parties to the Covenant, Mar 29, 2004, ¶ 8, http://docstore.ohchr.org/SelfServices/FilesHandler.ashx?enc=6QkG1d%2FPPRiCAghKb7yhsjYoiCfMKoIRv2FVaVzRkMjTnjRO%2Bfud3c PVrcM9YRoiW6Txaxgp3f9kUFpWoq%2FhW%2FTpKi2tPhZsbEJw%2FGeZRASjdFuuJQRnbJEaUhby31WiQPl2mLFDe6ZSwMMvmQGVHA%3D %3D.

[13] *Id.*

[14] CESCR, General Comment 19, *supra* n. 5, ¶ 43.

prevent corporations and "agents acting under their authority" from interfering with that right.[15] In the context of social security schemes that are "operated or controlled by third parties," States "retain the responsibility of administering the national social security system and ensuring that private actors do not compromise equal, adequate, affordable, and accessible social security."[16] To prevent abuses, "an effective regulatory system must be established which includes framework legislation, independent monitoring, genuine public participation and imposition of penalties for non-compliance."[17]

Human Rights Implications of Using Technology in Cash and Food Assistance Programs

States rely on AI and related technologies to automate two critical stages of the welfare distribution process: the verification of claimants' identity, and the assessment of eligibility and benefits levels. Throughout the entire welfare delivery cycle, States also employ these technologies to investigate, adjudicate and impose penalties for fraud.

Identity Verification

Aadhaar (India)

The use of AI to verify the identity of welfare recipients may be part of a broader push to establish national digital identity frameworks that manage individuals' access to government entitlements through a single, government issued identity. In 2009, India launched Aadhaar, a digital identity framework that assigns every citizen a unique twelve-digit identification number linked to the individual's biometric and demographic data. Under the Aadhaar Act of 2016, beneficiaries of various government welfare programs such as the Public Distribution System (PDS), which provides subsidized food grains to millions of households, are required to register and use Aadhaar to access their entitlements.[18]

---

[15] *Id.*, ¶ 45.

[16] *Id.*, ¶ 46.

[17] *Id.*

[18] "India: Identification Project Threatens Rights", Human Rights Watch news release, Jan 13, 2018, https://www.hrw.org/news/2018/01/13/india-identification-project-threatens-rights ("HRW Jan 13 release").

In 2018, the Indian Supreme Court upheld the government's authority to mandate Aadhaar as a precondition for accessing food rations and other welfare benefits.[19] However, the Court ruled that certain provisions of the Aadhaar Act were unconstitutional, and also barred the private sector from seeking access to Aadhaar data.[20] In February 2019, the government passed amendments to the Aadhaar Act that restored such access and bypassed the Court's ruling.[21]

Concerns

Human Rights Watch has found that eligible families have been denied access to subsidized food grains and other benefits because they did not have an Aadhaar number, had not linked it to their ration cards or experienced failures in authenticating their fingerprints.[22] Authentication failures disproportionately affect manual laborers, older persons and other individuals with worn fingerprints.[23] Since the Aadhaar machines installed in food distribution outlets require an internet connection, poor connectivity in rural areas has also led to disruptions in food distribution schedules.[24] Local activists have found that Aadhaar-related denials of food rations have led some to starve to death.[25]

Aadhaar also imposes invasive biometric identification and data collection requirements as conditions for accessing subsidized food grains and other essential public services. These requirements have created the world's largest database of biometric identity information, escalating the risk of unnecessary and disproportionate surveillance.[26] To mitigate this risk, the Supreme Court imposed several restrictions, including the requirement of judicial approval for law enforcement access to Aadhaar data, and a six-

---

[19] "India: Top Court OK's Biometric ID Program", Human Rights Watch news release, Sep 27, 2018, https://www.hrw.org/news/2018/09/27/india-top-court-oks-biometric-id-program.

[20] *Id.*

[21] Gautam Batia, *The Aadhaar ordinance raises serious constitutional concerns*, Hindustan Times, Mar 1, 2019, https://www.hindustantimes.com/columns/the-aadhaar-ordinance-raises-serious-constitutional-concerns/story-MbbAgChx8ao04DC1E3tLEI.html.

[22] HRW Jan 13 release, *supra* n. 18.

[23] *Id.*

[24] Jean Drèze, Nazar Khalid, Reetika Khera, Anmol Somanchi, *Aadhaar and Food Security in Jharkhand Pain without Gain?*, Economic & Political Weekly (Dec 16, 2017 Vol. LII No. 50), 55, http://www.im4change.org/siteadmin/tinymce/uploaded/Aadhaar%20and%20Food%20Security%20in%20Jharkhand%20Pain%20without%20Gain.pdf.

[25] "Of 42 'Hunger-Related' Deaths Since 2017, 25 'Linked to Aadhaar Issues'", The Wire, Sep 21, 2018, https://thewire.in/rights/of-42-hunger-related-deaths-since-2017-25-linked-to-aadhaar-issues.

[26] HRW Jan 13 release, *supra* n. 18.

month limit on the retention of authentication records and transaction logs.[27] However, these changes do not address the scope of biometric data and personal information collected under the program. Human Rights Watch has also raised concern about the multiple data breaches associated with Aadhaar since its implementation.[28]

These interferences with privacy disproportionately affect minorities: for example, local activists fear that transgender individuals are at greater risk of discrimination and persecution when they are forced to disclose their gender identity to the government, or if such information is leaked to the public.[29] These risks also raise the possibility that transgender individuals will be deterred from seeking access to essential public services linked to Aadhaar.

Knowledge Based Authentication System (California, United States)

Countries without national ID schemes also rely on automated decision-making to verify the identity of welfare claimants by comparing multiple sources of identity-related information drawn from a wide range of government and private databases. In the United States, the California state legislature in 2017 amended the Welfare and Institutions Code to replace fingerprint imaging with an "automated, nonbiometric" method for verifying the identity of applicants to the CalWORKs program, which provides cash assistance to needy families.[30] The Department of Social Services (DSS) selected US-based private company Pondera Solutions[31] to conduct a pilot of a cloud-based identity verification system known as the Knowledge Based Authentication system (KBA).[32] The pilot was conducted in six counties.[33]

---

[27] *Justice K.S. Puttaswamy (Retd.) et al. v. India et al.*, Writ Petition (Civil) No. 494 of 2012 & connected matters, majority opinion by A.K. Sikri, J, ¶¶ 205, 344, 345, https://www.supremecourtofindia.nic.in/supremecourt/2012/35071/35071_2012_Judgement_26-Sep-2018.pdf.

[28] HRW Jan 13 release, *supra* n. 18.

[29] "Aadhaar exposes transgenders to violence, discrimination and surveillance: SC told", India Today, May 20, 2018, https://www.indiatoday.in/india/story/aadhaar-exposes-transgenders-to-violence-discrimination-and-surveillance-sc-told-1193964-2018-03-20.

[30] Section 10831, Welfare and Institutions Code, https://leginfo.legislature.ca.gov/faces/codes_displayText.xhtml?lawCode=WIC&division=9.&title=&part=2.&chapter=4.6.&article=

[31] "Analytics With Context", Pondera Solutions, https://www.ponderasolutions.com/about-us/. Human Rights Watch contacted Pondera to obtain more information about the KBA but it did not respond at the time of the submission.

[32] California Department of Social Services, *Summary of Options for Replacing the Statewide Fingerprint Imaging System* (October 2017), http://www.cdss.ca.gov/Portals/9/Leg/SFIS%20Replacement%20Leg%20Report%20FINAL%2010.27.2017.pdf?ver=2017-11-21-150105-823, 5 ("CDSS 2017 report").

[33] *Id.*, 5.

KBA checks a CalWORKs application against "over 10,000 public sources" of data from "dozens of categories and hundreds of jurisdictions," including data from credit bureaus, government agencies and "utility and telephone companies."[34] This initial assessment is "designed to verify that the identity provided to the program is legitimate."[35] It also generates a multiple-choice quiz for applicants that "seeks to ensure that the applicant is in fact the individual that they are representing themselves to be."[36] Applicants are assigned a fraud risk code based on the system's initial assessment and their answers to the quiz.[37]

At the conclusion of the pilot in October 2017, DSS announced that it was intending to roll out KBA for phone or online benefits applications by the summer of 2018.[38] It is unclear, however, whether it has adhered to this timeline.

Concerns

In its report on the results of the pilot, DSS did not explain how the KBA analyzes the wide variety of data sources at its disposal to verify an applicant's identity and generate quiz questions. Although KBA standardizes the spelling of addresses "to avoid misspellings and other common mistakes," it is unclear how the system responds to other errors in the data it is provided, such as discrepancies in dates, phone numbers and demographic details.[39] The report also acknowledges KBA's potential to generate "false positives," but does not provide information about DSS's plans to prevent or mitigate such errors.[40]

This lack of transparency makes it difficult for welfare claimants and the broader public to assess the reliability, accuracy or fairness of KBA's risk assessment calculus. If incorrect information is associated with claimants and answers to the quiz questions are wrongly marked as errors, it will be difficult for them to identify the source of the error and hold the relevant authorities accountable. In addition, KBA's analysis of large datasets containing a wide range of sensitive and personal information raises questions about how the system safeguards applicants' privacy.

---

[34] CDSS 2017 report, *supra* n. 32, 18.

[35] *Id.*,

[36] *Id.*, 12 – 13.

[37] *Id.*, at 15; Kathleen Wilson, *Ventura County joins ID program for welfare benefits* (Mar 26, 2017), https://www.vcstar.com/story/news/local/2017/03/26/ventura-county-joins-id-program-welfare-benefits/99434068/.

[38] CDSS 2017 report, *supra* n. 32, 9.

[39] *Id.*, 18.

[40] *Id.*, 17.

The Coalition of California Welfare Rights Organizations has also raised concern that KBA's multiple-choice quiz creates additional obstacles for marginalized populations. Families that have been homeless for a long time may be unable to answer questions such as "How long have you lived in your current residence?" or "Which of the following streets have you ever lived or used as your address?"[41] Furthermore, questions regarding residential and relationship history in the United States assume that respondents have longstanding community ties, and are ill-suited to the needs and concerns of newly arrived immigrant families.

Assessment of Eligibility and Benefits Levels

Ontario Works (Ontario, Canada)

Governments are also replacing or supplementing case workers' assessments of eligibility and benefits levels with predictive analytics and other AI-based assessment tools. Since November 2014, Ontario Works, the financial assistance program of the Canadian province of Ontario, has been relying on the Social Assistance Management System (SAMS) to automatically generate decisions on eligibility for cash transfers and other benefits. Decisions are generated based on data that frontline workers collect from applicants and recipients and subsequently "fit into narrow-drop down menu categories."[42]

SAMS is based on Cúram, a customizable off-the-shelf software sold by IBM as a platform for "complete intake, eligibility determination and benefit calculation for social programs."[43] Latest versions of the software are also equipped with functions to monitor impermissible instances of "concurrent eligibility" in food and cash assistance programs, potentially indicating the system's

---

[41] "Advocate Response to DSS Options for Replacing the Statewide Fingerprint Imaging System (SFIS)", Coalition of California Welfare Rights Organizations, Inc., Dec 4, 2017, 12, https://www.ccwro.org/2012/1743-12-4-17-sfis-replacement-calworks-consumer-report-sb-89/file.

[42] Jennifer Marie Raso, Administrative Justice: Guiding Caseworker Discretion, 2018, 245 – 246, https://tspace.library.utoronto.ca/bitstream/1807/82936/1/Raso_Jennifer_201803_SJD_thesis.pdf ("Raso thesis").

[43] "Income Support", IBM Knowledge Center, https://www.ibm.com/support/knowledgecenter/en/SS8S5A_7.0.4/com.ibm.curam.nav.doc/common/ctr_income_support_solution.html

ability to perform both benefits assessments and fraud detection.[44] Cúram is also used to administer welfare programs in Alberta, North Carolina, Hamburg, Queensland and New Zealand.[45]

A 2015 audit of SAMS conducted by the state's Auditor General found that SAMS suffered from "serious defects and was not fully functional,"[46] leading to potential underpayments of benefits totaling $51 million CAD.[47] In one case, SAMS erroneously deducted $32 from a client's total benefit payments each month after it incorrectly determined that the client had been previously overpaid.[48]

The Auditor General also found that SAMS "automatically generated" letters to beneficiaries with "incorrect information" that caused "stress and confusion."[49] For example, a letter sent to two beneficiaries, whom owed $1,328 in overpayments, accused them of owing $8,736.[50] Another letter notified a beneficiary that their benefits would be withdrawn because they no longer lived in Ontario, but caseworkers found that the beneficiary "had never left Ontario."[51]

Universal Credit (United Kingdom)

The Special Rapporteur on extreme poverty and human rights has observed that the Universal Credit (UC), the UK's welfare benefits program, "is only possible because of the automated calculation of benefits."[52] The Real Time Information (RTI) system calculates UC payments based on earnings information reported by employers to Her Majesty's Revenue and Customs (HMRC), the country's tax authority.[53]

---

[44] "Monitoring Concurrent Eligibility", IBM Knowledge Center, https://www.ibm.com/support/knowledgecenter/en/SS8S5A_7.0.4/com.ibm.curam.content.doc/IncomeSupport/c_INCSPRT_MonitoringConcurrentEligibility.html

[45] "Ministry of Community and Social Services SAMS Transition Review", Apr 30, 2015, 60 – 62, https://www.mcss.gov.on.ca/documents/en/mcss/social/SAMS_Transition_Review_Final.pdf.

[46] "SAMS—Social Assistance Management System," 2015 Annual Report of the Office of the Auditor General of Ontario, Fall 2015, 479, http://www.auditor.on.ca/en/reports_en/en15/3.12en15.pdf.

[47] *Id.*, 485.

[48] *Id.*, 480.

[49] *Id.*, 481.

[50] *Id.*

[51] *Id.*

[52] "Statement on Visit to the United Kingdom, by Professor Philip Alston, United Nations Special Rapporteur on extreme poverty and human rights", OHCHR News and Events, Nov 16, 2018, https://www.ohchr.org/en/NewsEvents/Pages/DisplayNews.aspx?NewsID=23881&LangID=E.

[53] "Universal credit: RTI and Universal Credit", revenuebenefits.org.uk, https://revenuebenefits.org.uk/universal-credit/guidance/entitlement-to-uc/rti-and-universal-credit/#What%20is%20Real%20Time%20Information?

It appears that RTI's calculations are unable to correct for late, inaccurate or missing reports, and this can lead to delays and errors in UC payments.[54] In the 2016/2017 fiscal year, 5.7% of 590m UC payments were marred by late reporting.[55] To address RTI's inability to take into account reporting errors, HMRC and the Department of Work and Pensions, which oversees UC, have established a special joint initiative known as the Late, Missing and Incorrect RTI Project.[56]

Concerns

The failures of automated decision-making in Ontario Works and UC indicate that governments have overestimated the technology's capacity to conduct complex and context-sensitive assessments of eligibility and benefits levels. Benefits calculators are only as accurate as the data provided to them: unlike case workers, these systems are unable to investigate the reasons for data inaccuracies and other discrepancies and make necessary adjustments on a case-by-case basis.

Despite these limitations, both programs fail to supplement automated decision-making with human review that ensures benefits changes are reasonable and in accordance with due process and domestic law requirements. In the Ontario Works program, Professor Jennifer Raso of the University of Toronto Law School has found that the design of SAMS "obstructs frontline workers from challenging the substance of its decisions."[57] For example, SAMS does not permit frontline workers to challenge its assessments of whether Ontario Works recipients with a history of living in the same household are dependent on each other – a common ground for reducing the value of benefits.[58]

In the UK, service workers have not been provided with training that enables them to effectively troubleshoot RTI and other IT errors that may lead to the withdrawal, reduction or suspension of benefits. According to welfare researchers from the University of

---

[54] "We know that Universal credit is a mess, but what about HMRC Real Time Information system?", Disabled People Against Cuts, Oct 25, 2017, https://dpac.uk.net/2017/10/we-know-that-universal-credit-is-a-mess-but-what-about-hmrc-real-time-information-system/; Patrick Butler, *Universal credit IT system 'broken', whistleblowers say*, The Guardian, Jul 22, 2018, https://www.theguardian.com/society/2018/jul/22/universal-credit-it-system-broken-service-centre-whistleblowers-say.

[55] House of Commons Debate (Oct 24, 2017), vol. 630, col. 224, https://hansard.parliament.uk/Commons/2017-10-24/debates/81F68207-B696-4D36-92FB-1A00AFD3817F/UniversalCreditRoll-Out#contribution-E6990374-47F5-41EA-9012-0BE88794C0AB.

[56] Response to Sept 4, 2017 FOI request, DWP Central Freedom of Information Team, FOI Ref No. 3665, Sept 26, 2017, https://www.whatdotheyknow.com/request/429139/response/1043259/attach/2/FoI%203665%20reply.pdf?cookie_passthrough=1.

[57] Raso thesis, *supra* n. 42, 254.

[58] *Id.*, 254 – 255.

Birmingham and the University of Leeds, former staff of Jobcentre Plus, an agency which administers UC's benefits for jobseekers, "described being permanently on the 'back foot', in that digital services were rolled out without staff being given the relevant training."[59] A former UC call center worker in Grimsby also told The Guardian that there was "massive variation" in staff's understanding of the UC policies and systems, leading to contradictory responses to the same query.[60]

Detection, Investigation and Punishment of Welfare Fraud

Systeem Risico Inventarisatie (The Netherlands)

Governments increasingly rely on automated fraud detection systems to detect and flag risks of welfare fraud. In the Netherlands, the Ministry of Social Affairs and Employment operates the *Systeem Risico Inventarisatie* or System Risk Indication (SyRI), which is used by several municipal governments to detect benefits fraud. SyRI flags individuals as potential fraud risks through an algorithmic risk assessment tool that draws on multiple sources of data, including tax returns.[61] However, the government has offered few details on the specific types of data used and the criteria for determining risk. It has rejected calls for transparency about how the algorithm works, claiming that disclosing such information would reduce its effectiveness in detecting fraud.[62] It has also not explained the circumstances under which case workers or fraud investigators may deviate from these risk assessments, if at all.

Online Compliance Intervention (Australia)

Several governments do not only automate assessments of welfare fraud, but also the imposition of penalties such as fines or the reduction or withdrawal of benefits. In July 2016, Australia's Department of Human Services (DHS) launched Online Compliance Intervention (OCI), a fully automated income data verification system that generates debt notices based on differences

---

[59] Kayleigh Garthwaite, Jo Ingold, and Mark Monaghan, *Universal Credit and the perspectives of ex-Jobcentre Plus staff*, LSE British Politics and Policy blog, Jan 15, 2019, https://blogs.lse.ac.uk/politicsandpolicy/ex-jobcentre-plus-staff/

[60] Patrick Butler, *Universal credit staff: 'It was more about getting them off the phone'*, The Guardian, Jul 22, 2018, https://www.theguardian.com/society/2018/jul/22/universal-credit-whistleblowers-heartbreaking-impact-flawed-system-claimants.

[61] "Automating Society: Taking Stock of Automated Decision Making in the EU", AlgorithmWatch and Bertelsmann Stiftung, January 2019 (1st ed.), 101, https://algorithmwatch.org/wp-content/uploads/2019/02/Automating_Society_Report_2019.pdf.

[62] "Beantwoording Kamervragen over 'Algoritme voorspelt wie fraude pleegt bij bijstandsuitkering'", Apr 12, 2018, https://www.rijksoverheid.nl/documenten/kamerstukken/2018/05/28/beantwoording-kamervragen-over-algoritme-voorspelt-wie-fraude-pleegt-bij-bijstandsuitkering.

between fortnightly income figures reported by welfare beneficiaries and their employers.[63] A Parliamentary inquiry conducted by the Senate's Community Affairs Reference Committee found that OCI's income calculation formula is not programmed to take into account fluctuations in a beneficiary's income, leading to "inaccurate calculations of debt," particularly for casual or seasonal workers with irregular incomes.[64] OCI is also unable to make adjustments for employer error.[65]

DHS does not require manual review of the OCI's findings, instead placing the onus on beneficiaries to submit evidence rebutting debt notices.[66] However, affected beneficiaries and their representatives testified before the Senate that debt notices either provided inadequate information or were too complicated to understand, making them difficult to challenge.[67] Some complained that they had to submit Freedom of Information requests to compel DHS to release information about how their debts were calculated.[68]

The Senate Committee has urged DHS to put OCI "on hold" until these issues are resolved,[69] but the Department has rejected this recommendation and claimed that its implementation has gone "quite well."[70]

Michigan Integrated Data Automated System (Michigan, United States)

In October 2013, Michigan's Unemployment Insurance Agency (UIA) launched the Michigan Integrated Data Automated System (MiDAS) to adjudicate and impose penalties for unemployment benefits fraud. Between October 2013 and August 2015, MiDAS was programmed to automatically treat differences between income figures reported by beneficiaries and their employers as evidence of fraud.[71] The system was not capable of investigating whether there are legitimate

---

[63] The Senate, Community Affairs References Committee, "Design, scope, cost-benefit analysis, contracts awarded and implementation associated with the Better Management of the Social Welfare System initiative," 21 June 2017, 13, https://www.aph.gov.au/Parliamentary_Business/Committees/Senate/Community_Affairs/SocialWelfareSystem/~/media/Committees/clac_ctte/SocialWelfareSystem/report.pdf ("Senate Committee OIC Report").

[64] *Id.*, 32, 34.

[65] *Id.*, 32.

[66] *Id.*, 15.

[67] *Id.*, 57 – 58.

[68] *Id.*

[69] *Id.*, ix.

[70] Rae Johnston, *The Department Of Human Services Says RoboDebt 'Went Well' And 'Delivered Lots Of Savings'*, Gizmodo, Mar 27, 2018, https://www.gizmodo.com.au/2018/03/the-department-of-human-services-says-robodebt-went-well-and-delivered-lots-of-savings/.

[71] *Cahoo, et al. v. SAS Analytics Inc., et al.*, Nos. 18-1295/1296 (6th Cir. 2019), 3, https://law.justia.com/cases/federal/appellate-courts/ca6/18-1296/18-1296-2019-01-03.html.

reasons for these discrepancies, such as employer error or pay disputes.[72] Like OCI, MiDAS was also unable to determine whether these discrepancies are attributable to fluctuations in a beneficiary's income.[73]

Based on its initial assessments, MiDAS sent beneficiaries suspected of fraud online multiple-choice questionnaires asking whether they are "intentionally provid[ing] false information to obtain benefits you were not entitle[d] to receive," and "[w]hy … you believe you were entitled to benefits."[74] Failure to respond in ten days, or a response that MiDAS deemed unsatisfactory, would automatically trigger conclusive determinations of fraud.[75] Based on these determinations, MiDAS would terminate the benefits of affected beneficiaries and initiate proceedings to seize their tax refunds or garnish their wages.[76]

UIA subsequently found that, between October 2013 and August 2015, about 44,000 of the 62,784 determinations of fraud that MiDAS generated were in error.[77] In a lawsuit that a group of beneficiaries filed against UIA, the U.S. Federal Court of Appeals for the region concluded that MiDAS "did not allow for a fact-based adjudication or give the claimant the opportunity to present evidence to prove that he or she did not engage in disqualifying conduct."[78]

Despite these failures, UIA continues to operate MiDAS.[79] It has committed to additional data analysis to detect benefits payments needing "further review" and to enhance the appeals process.[80] However, it is unclear whether UIA has made any changes to the underlying data matching algorithm or incorporated meaningful human review into the system's fraud detection functions.

---

[72] *Id.*

[73] *Id.*

[74] *Id.*, 3 – 4.

[75] *Id.*, 4 – 5.

[76] *Id.*

[77] "Michigan's unemployment agency completes review of fraud determination cases; comprehensive changes underway to improve customer service and operations", Michigan.gov press release, https://www.michigan.gov/som/0,4669,7-192-47796-428651--,00.html; Paul Egan, *After Falsely Accusing Thousands of Unemployment Fraud and Wrongly Taking Their Money, Michigan Makes Amends. But Is It Enough?*, Governing, Aug 15, 2017, https://www.governing.com/topics/health-human-services/tns-michigan-unemployment-fraud-accusations.html.

[78] *Cahoo v. SAS Analytics*, *supra* n. 71, 3.

[79] "Final Agency Response for OAG Performance Audit of MiDAS System," State of Michigan Department of Talent and Economic Development, Jun 17, 2016, https://audgen.michigan.gov/wp-content/uploads/2016/07/ap641059315.pdf.

[80] *Id.*, 6.

Automated Verification of Job Activity Reports (Sweden)

The Swedish Employment Service (*Arbetsförmedlingen*) relies on automated decision-making to verify whether recipients of unemployment benefits have complied with job-seeking and other workfare obligations, and issue warnings, withhold payments and enforce other sanctions based on these assessments.[81] At the end of 2018, *Arbetsförmedlingen* discovered a 10 – 15% error rate in the automated verification of beneficiaries' job activity reports, potentially leading to 70,000 erroneous decisions to withhold benefit payments.[82] These errors have forced *Arbetsförmedlingen* to manually screen all job activity reports until the system can be repaired.[83]

    Concerns

    These cases reinforce the need for appropriate human review that corroborates fraud findings generated by automated systems, as well as clear, transparent and accessible appeals mechanisms that enable beneficiaries to meaningfully challenge these findings. Without these safeguards, limitations or errors in automated decision-making potentially lead to mass violations of the right to social security. The Swedish and Michigan examples illustrate that the automation of fraud determinations at scale replicates errors in data processing and analysis across the entire system, leading to incorrect benefits changes and penalties that affect thousands of beneficiaries. The lack of transparency compounds these failures, preventing beneficiaries from accessing information about their case or participating in its adjudication.

    These failures also illustrate the potential for welfare discrimination based on beneficiaries' socio-economic backgrounds. Flaws in OCI and MiDAS' income calculation formulae, for example, disproportionately affect workers with irregular incomes, whom the CESCR has designated as a protected category in the social security context.[84] The Michigan Law School Unemployment Insurance Clinic has also raised concern that beneficiaries experiencing financial hardship have extremely limited options to challenge MiDAS's determinations: Charges of fraud disqualify them from free representation under the state's pro bono program, and it is unlikely that they are able to afford private

---

[81] Tom Wills, *Sweden: Erroneous algorithm stops payments for over 70,000 unemployed*, AlgorithmWatch, Feb 28, 2019, https://algorithmwatch.org/story/rogue-algorithm-in-sweden-stops-welfare-payments/.

[82] "SVT avslöjar: Datafel kan ha skapat tiotusentals felaktiga beslut hos Arbetsförmedlingen", SVT Nyheter, Feb, 13, 2019, https://www.svt.se/nyheter/inrikes/svt-avslojar-stort-datafel-hos-arbetsformedlingen-tusentals-kan-ha-forlorat-ersattning.

[83] *Id.*

[84] CESCR, General Comment 19, *supra* n. 5, ¶ 31.

representation.[85] Under OCI, the Senate Committee heard evidence that beneficiaries with poor literacy or English language skills found it particularly difficult to understand the highly technical language used in debt notices.[86]

The Role of the Private Sector

The case studies outlined in this submission show that the private sector is key in developing and operating automated systems of welfare governance. Companies involved range from those providing specialized fraud detection services to large enterprise software manufacturers. However, it is unclear whether these companies have established policies or processes that meaningfully address their human rights impacts.

These public-private partnerships make it difficult to hold both State and non-State actors accountable for failures in the welfare delivery services that are outsourced. AI Now, an organization dedicated to examining AI's public and social impacts, has found that risk assessment models and other automated decision-making tools are typically hidden behind broad assertions of intellectual property and trade secrets, making it difficult for affected rights holders and the broader public to scrutinize their potential for discrimination and other human rights impacts.[87] There also does not appear to be much pressure on companies to conduct human rights impact assessments or consultations with welfare recipients during the design, customization and implementation of welfare delivery software, particularly since governments are not insisting on adherence to the UN Guiding Principles.

Implementation of the UN Guiding Principles in the Information and communications technology sector, which has hitherto focused on the responsibilities of internet and telecommunications companies to respect freedom of expression and privacy, offers general guidance and best practices that are adaptable to the commercial delivery of welfare-related services.

Human rights due diligence is a central component of these responsibilities, and requires impact assessments that address issues of privacy, discrimination and exclusion early on in the design

---

[85] Steven Gray, *The Future of UIA Claims After the MiDAS Fraud Scandal*, University of Michigan Law School Unemployment Insurance Clinic, Jan 15, 2015, 5, http://spb.mplp.org:8080/download/attachments/11239427/SPB+UI+Materials+1-18.pdf?version=1.

[86] Senate Committee OIC Report, *supra* n. 63, 51 – 54.

[87] Dillon Reisman, Jason Schultz, Kate Crawford, Meredith Whittaker, *Algorithmic Impact Assessments: A Practical Framework for Public Agency Accountability*, AI Now, April 2018, 13, https://ainowinstitute.org/aiareport2018.pdf.

and engineering phase, internal training, dialogue and collaboration on these issues, and regular consultations with civil society and affected rights holders.[88] Companies should also establish meaningful transparency measures, such as policies to disclose the outcomes of impact assessments and the concrete steps they have taken to prevent or mitigate human rights risks.[89] Furthermore, companies have a responsibility to provide access to effective remedies (such as financial restitution) when they have "caused or contributed to adverse [human rights] impacts."[90]

In the context of digital welfare, companies should, at a minimum, provide accessible explanations of how AI and other data-driven technologies are integrated into welfare decision-making, disclose and address automation errors in a timely fashion, submit to audits of algorithms and training data by external assessors, and develop processes for identifying, correcting and mitigating discrimination and bias in system inputs and outcomes.[91]

In accordance with their obligations to protect against private interference with Covenant rights, States should establish implementation of the UN Guiding Principles as a mandatory condition for the sale of identity verification, benefits assessment and fraud detection products and services to welfare agencies and other relevant authorities.

---

[88] Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, David Kaye, Human Rights Council, A/HRC/35/22, Mar 30, 2017, ¶¶ 52 – 60, https://undocs.org/A/HRC/35/22.

[89] *Id.*, ¶¶ 70 – 72.

[90] *Id.*, ¶¶ 73 – 75.

[91] A/73/348, *supra* n. 1, ¶¶ 65 – 70.