
*Attila Dabis**

DUAL ROLES AND RIGHTS IMPLICATIONS: THE UNSETTLED DYNAMICS OF AI USE IN RELATION TO MINORITISED COMMUNITIES

1 Introduction

The exponential growth of artificial intelligence (AI) tools has generated a series of new challenges and opportunities in which algorithmic systems irrevocably influence the social, political, and cultural dimensions of human life, including the lives of minority and indigenous communities. AI's growing integration into state and private decision-making processes and policy implementation practices demands a reassessment of how these systems interact with the rights of minoritised communities. While AI's transformative and disruptive potential is heavily discussed in areas such as healthcare or education,¹ its implications for equality and minority protection remain inadequately examined. This study, therefore, situates AI within the narrower framework of the human rights of minorities, focusing on those groups whose social, linguistic, or cultural identities have historically rendered them vulnerable to systemic exclusion. The paper's core proposition is that AI operates in a dual normative capacity. It can both entrench discrimination and serve as a catalyst for cultural empowerment, depending on the frameworks of human agency and policy oversight in place.

Whether the development of AI tools had benevolent motivations or not, the same technologies can be used in good faith and abused destructively.² As Suresh and Guttag have demonstrated, there are no

* AI Consultant, Corvinus University of Budapest; Foreign Affairs Commissioner, Székler National Council; Co-editor-in-chief, Minority Protection Journal. Email: dabis.attila@kre.hu.

¹ For a summary of the AI-related educational dilemmas, see: Dabis – Csáki 2024.

² See, e.g.: Bentham – Haynes 2019.

fewer than seven different forms of biases that can arise throughout the machine learning life cycle.³ This means that there are at least seven different points where human error in designing and training an AI system can become a source of harm, even unintentionally: Historical bias, Representation bias, Measurement bias, Aggregation bias, Learning bias, Evaluation bias, and Deployment bias (see Annex No. 1 for more detail on these).

This typology points out the inherent fragility of AI fairness,⁴ and draws attention to why bias mitigation cannot rely solely on the technical aspects of the issue.⁵ Technological advancements are neither inherently good nor bad. Their impact depends on how they are utilised, highlighting the crucial role of the intentions guiding the use of AI tools (in relation to this topic, see Annex No. 2 on how the China-based Deep Seek chatbot answered my questions relating to the Uyghur community of Xinjiang). For example, Hikvision's surveillance applications, as we will see, can be utilised in both democratic and authoritarian contexts for benevolent and harmful purposes. Consequently, the distinction between unintentional algorithmic bias and politically motivated outcomes that can be characterised as harmful to minoritised communities forms a key conceptual basis for this paper and will be highlighted through the Uyghur example. The Uyghur case

³ Suresh – Guttag 2021.

⁴ For the manifold ways "fairness" can be interpreted, see: Hutchinson – Mitchell 2019.

⁵ Take hate speech, for example, a phenomenon that primarily occurs in digital spaces. In relation to hate speech, one needs to apply the Rabat threshold test to assess whether an expression amounts to hate speech and can thus be curbed. This is because the freedom of speech and expression is a fundamental and universal human right whose limitation requires proper justification. Accordingly, the Rabat threshold test outlines a six-part threshold test considering (1) the social and political context, (2) status of the speaker, (3) intent to incite the audience against a target group, (4) content and form of the speech, (5) extent of its dissemination and (6) likelihood of harm, including imminence. German legal philosophy calls this approach a "*Schranken-Schranke*", that is, the "limitation of the limits". This signifies that while certain rights can be restricted under specific circumstances, there are limits to how far these restrictions can extend without compromising human rights. This ensures that fundamental rights are protected even when the state seeks to impose limitations for reasons such as public safety or morality. See: https://www.ohchr.org/sites/default/files/Rabat_threshold_test.pdf, accessed: 24.10.2024.

illustrates how AI is not merely complicit in disenfranchising a minoritised community but can become a constitutive element of state policy, transforming digital infrastructure into a tool of systemic persecution.

At the same time, it is essential to recognise AI's dual nature, as the very technologies that enable repression can also promote inclusion, self-representation, and linguistic revitalisation. Accordingly, this paper also examines the positive applications of AI, particularly in supporting minoritised communities (like First Nations communities in North America or traditional minorities like the Basques or Catalans in Europe) through language revitalisation. Throughout the paper, "positive" applications of AI refer to instances where AI use aligns with, reinforces, or advances existing minority-protection norms, in contrast to "negative" applications, which undermine or weaken those norms.

In this context, the paper aims to provide a tentative overview of the relationship between AI tools and minoritised communities, focusing on the potential benefits and risks. Apart from documenting harms and benefits, it also translates the findings into normative and policy implications for advocates, institutions, and states seeking to ensure that AI development aligns with the principles of equality, dignity, and cultural diversity.

Regarding its structure, the paper draws on the experiences of Tankó,⁶ and distributes the relevant literature across the whole study as opposed to having a separate section for the literature review. Relevant sources will thus be integrated throughout all sections, beginning with the Introduction, followed by the Methodology, and culminating in separate chapters that address the negative and positive effects of AI on minoritised communities, with a final section summarising the paper's Conclusions.

⁶ Tankó 2021, 104.

2 Methodology

There are three primary aspects relevant to the study's analytical framework. First, the paper provides a broad and non-exhaustive overview based on the content analysis of academic sources and the practices of AI usage relevant from the perspective of minoritised communities. A review of the existing literature reveals that substantially more attention has been devoted to examining the implications of AI for human rights in general, covering issues such as data privacy, freedom of expression and assembly, the rights of minors, algorithmic practices affecting labour rights, and the human rights dimensions of intellectual property, among others,⁷ yet it has paid relatively little attention to examining concerns specifically from the perspective of minoritised communities.⁸ This explains the rationale, motivation, and justification for investigating this research gap and compiling appropriate evidence to bring the reader closer to understanding the manifold ways AI tools impact the lives and human rights of minoritised communities.

The study examines cases from ten different minoritised communities and languages (Uyghur, Catalan, Cree, Cherokee, Hawaiian, Inuktitut, Kichwa, Māori, Ojibwe, and Sámi), while also referencing some majority languages, most notably Icelandic, in the context of digital minoritisation. These cases are presented at varying levels of detail, reflecting their respective significance. In particular, the Uyghur case is described in greater depth due to the scale of human rights violations and crimes against humanity involved. The paper seeks to maintain balance in its treatment of AI's impacts on minoritised communities. While the negative consequences are illustrated through the singular, detailed Uyghur case, the positive applications are showcased across nine shorter examples, highlighting the diverse ways AI can support these communities, particularly in language revitalisation

⁷ See, e.g.: Nagy 2024, Kamble 2024.

⁸ Exceptions include the topic of racial discrimination, see e.g. Koen-Mufadi 2023.

The second aspect pertains to the terminology. The paper will review existing evidence discussing both national, ethnic, linguistic, and religious minorities, as well as indigenous and tribal peoples, under the umbrella term “minoritised communities” who speak “minoritised languages”. Minoritised languages refer to languages that have been marginalised and diminished in use due to historical and/or ongoing colonial and globalising forces. The term encompasses the languages of indigenous and tribal peoples, as well as those spoken by autochthonous national ethnic, linguistic, or religious minorities – i.e., languages that are not widely spoken, often not even recognised, and in some cases endangered, highlighting their vulnerability and the need for revitalisation efforts.⁹

Thirdly, this paper adopts a broad, overview-oriented perspective, recognising that a comprehensive examination of all forms of bias—such as population, behavioural, content production, linking, or temporal biases—and their emergence across the AI lifecycle lies beyond the scope of this analysis.¹⁰ Instead, it seeks to highlight key patterns and challenges as a starting point for further investigation. In this sense, the study aims to make a preliminary contribution, inviting scholars from various disciplines to explore the manifold ways in which minoritised communities may experience both the empowering and harmful effects of AI technologies.

Finally, for the content analysis in the Uyghur case, the study uses the fundamental rights enumerated under the Universal Declaration of Human Rights (UDHR), the Convention on the Prevention and Punishment of the Crime of Genocide (Genocide Convention), and other UN treaties that China signed and ratified, including the Convention on the Elimination of Discrimination Against Women (CEDAW), the Convention on the Elimination of All Forms of Racial Discrimination (CERD), the Convention on the Rights of the Child (CRC), and the Convention against Torture and Other Cruel, Inhuman or Degrading Treatment or Punishment (CAT) as benchmarks

⁹ Hornberger et al. 2023.

¹⁰ For a detailed description of these various biases, see: Olteanu et al. 2019.

to structure the human rights violations that Uyghurs face, as these instruments stipulate a common set of human rights standards that China accepted to adhere to.

3 Evolving Practices of Minority Disenfranchisement

Having established UN treaties ratified by China as the analytical benchmarks, the paper now applies these standards to the Uyghur case. The section that follows identifies several novel mechanisms of systemic disenfranchisement and maps each practice to the articles and treaty provisions of specific UN instruments it may violate.

For this categorisation, it's essential to acknowledge that there are several competing interpretations of China's handling of the Uyghur community. Some scholars, for example, contest the interpretation that China is committing genocide against the Uyghurs, and therefore argue that it is not violating the Convention against Torture, on the grounds that there is neither evidence of mass physical extermination nor proof of an intent to destroy the group.¹¹ UN documents also refrain from using the term "genocide" in their description of the situation and categorise it as a set of crimes against humanity. The Office of the High Commissioner for Human Rights (OHCHR) wrote an overarching Assessment Report on the human rights concerns in the Xinjiang Uyghur Autonomous Region (XUAR) in 2022, which states that the extent of arbitrary and discriminatory detention of members of Uyghur and other predominantly Muslim groups may constitute crimes against humanity.¹² Corroborating this approach, a Judgement of the Independent Tribunal into Forced Organ Harvesting from Prisoners of Conscience in China from 17 June 2019 established that while the crime of genocide against members of the Falun Gong spiritual movement clearly amounted to genocide, they could not be certain that the requisite intent of genocide was proved in the case of Uyghurs

¹¹ Slawotsky 2021.

¹² OHCHR 2022, para 148.

based on the cases brought before them.¹³ These also dovetail with the conclusions of Human Rights Watch and Amnesty International, as rights organisations that have published investigations into state practices in Xinjiang, even though the latter source also alludes to an expert report arguing that both crimes against humanity and genocide were currently being committed in XUAR.¹⁴

Under the current circumstances, the term genocide is more often used by political actors. As such, the Subcommittee on International Human Rights of the Standing Committee on Foreign Affairs and International Development of the House of Commons of Canada released a statement in 2020 communicating stating that “the actions of the Chinese Communist Party constitute genocide as laid out in the Genocide Convention”.¹⁵ Similarly, there was bipartisan support for “condemning the ongoing genocide and crimes against humanity being committed against Uyghurs and members of other religious and ethnic minority groups by the People’s Republic of China” in the US Congress.¹⁶

Some academics share the general concern outlined above but adopt more cautious language when stating that practices such as the sterilisation of women and family separation in themselves tell “a story of people’s destruction”,¹⁷ pointing to the practice that children whose parents are sent to re-education camps or other detention centres are transferred to boarding schools or orphanages by the state regardless of their age or their lack of consent for such measures. As another scholar put it, internment camps and inter-generational separation together lead to “social death”.¹⁸

¹³ Source: <https://chinatribunal.com/final-judgment/>, accessed: 08.11.2024.

¹⁴ Even though China has not ratified the Rome Statute of the International Criminal Court, this report identifies those points in Art. 7 of the Rome Statute that China’s conduct toward Uyghurs has violated, see Macdonald et al. 2021, 1-2.

¹⁵ Source: <https://www.ourcommons.ca/DocumentViewer/en/43-2/SDIR/news-release/10903199>, accessed: 08.11.2024.

¹⁶ See H.Res.317: <https://www.congress.gov/bill/117th-congress/house-resolution/317/text>, accessed: 08.11.2024.

¹⁷ Çiçek 2022, 279.

¹⁸ Tobin 2022, 93.

Against this backdrop, I cite Article II of the UN's Genocide Convention below to provide additional context, as I will mention cases where state practice aligns with the elements of crimes defined in UN instruments on genocide. Based on this Convention, "genocide means any of the following acts committed with intent to destroy, in whole or in part, a national, ethnical, racial or religious group:

- (a) Killing members of the group;
- (b) Causing serious bodily or mental harm to members of the group;
- (c) Deliberately inflicting on the group conditions of life calculated to bring about its physical destruction in whole or in part;
- (d) Imposing measures intended to prevent births within the group;
- (e) Forcibly transferring children of the group to another group."¹⁹

3.1 Forced separation of families, enforced disappearances, arbitrary detention, lack of legal remedy

The approximately 10 million Uyghurs, a Turkic-Sunni Muslim minority, live in the westernmost province of China, Xinjiang. According to estimations,²⁰ more than 10% of the community (between 1 to 2 million people) are incarcerated in internment camps, officially dubbed by Chinese state officials as vocational training centres, mostly without access to legal remedy or due process rights, such as being informed of the reasons for arrest, to contest the detention before a judge, or to have access to lawyers and family members.²¹

Reviewing statements delivered at the UN Forum on Minority Issues and reports by human rights organisations, Uyghur speakers described circumstances amounting to systemic and severe violations of fundamental rights.²² Many reports highlight the forced separation of

¹⁹ Asat – Diamond 2019.

²⁰ Ibrahim 2021, 23.

²¹ For a comprehensive review of the conditions at these detention centres, see *Ibid.*

²² So far, 29 UN Forum statements mentioned the Uyghur community in some way, most of which came from the Uyghur World Congress. China is characterising this

families, mainly through labour transfer programs. The destinations of these forced relocations for parents are factories in eastern China, where the parents endure poor and abusive circumstances.²³ This forced labour is characterised by coercion, including threats of fines, confiscation of land, denial of marriage certificates and withholding of wages.²⁴ Meanwhile, children of detained Uyghurs are forcibly removed from their homes and into Chinese orphanage camps. A report indicates that some Uyghur children held in state-run orphanage facilities have attempted suicide, including by ingesting detergent, and have posed questions to their Chinese caretakers whether the institution in which they are confined is, in effect, a jail.²⁵ The same report states that the Chinese government has isolated these children from external influences to persuade them that their parents and relatives have forsaken or betrayed them.²⁶ To break their roots and impose Han Chinese culture and religion on the youngsters, children are forced to eat pork, wear archaic Han Chinese garments, are not allowed to speak their own language and live in prison-style orphanages.²⁷

advocacy organisation as a “notorious secessionist organisation” whose objective is to separate the Xinjiang Uyghur Autonomous Region from China to establish the so-called “East Turkistan State”, alleging that they have close ties to terrorist organisations, including the “Eastern Turkistan Islamic Movement”, which is designated as a terrorist organisation by the United Nation’s Security Council. Accordingly, China firmly opposes the attempt of the “World Uyghur Congress” and its members to utilise the Forum to conduct separatist activities. See, e.g. the CPR’s Forum statement from 2015: <https://www.minorityforum.info/database/entity/fja0vh7kvtt>, accessed: 16.10.2024.

²³ The practice of deploying forced labour among Uyghur, Kazakh and other ethnic minorities in sectors such as agriculture and manufacturing has been corroborated by UN Reports. See, e.g. one from the Special Rapporteur (SR) on contemporary forms of slavery, Tomoya Obokata’ from 2022: UN SR on contemporary forms of slavery 2022, or another one by Special Rapporteur on trafficking in persons, especially women and children, Siobhán Mullally: UN SR on trafficking in persons 2022.

²⁴ See Forum statements from 2010: <https://www.minorityforum.info/database/entity/r9ovo37ecuq>, and 2011: <https://www.minorityforum.info/database/entity/xj4ehm0j1r>, both accessed: 16.10.2024.

²⁵ Uyghur Research Institute 2020, 35.

²⁶ *Ibid.*, 35.

²⁷ See investigative video on this at Bitter Winter: https://www.youtube.com/watch?v=CNAtx4FFIRQ&ab_channel=BitterWinter, accessed: 03.12.2025.

UN Provisions Potentially Violated:

UDHR: Articles 5, 9, 12.

Genocide Convention: Article 2 (b, e).

CERD: Articles 2, 5.

CRC: Articles 8, 9, 25, 30, 35

CAT: Articles 1, 2, 12, 14, 16.

3.2 Reproductive Control, Repression of Women, and State Encroachment on Privacy

According to analyses, “Family Planning Policy” for Uyghur women in China is equivalent to draining their reproductive capacity, as Uyghur women are made to undergo gynaecological examinations and forced to take birth control medication, have IUDs inserted,²⁸ or be sterilised, which violates their right to health and reproductive well-being.²⁹ Mass rape and sexual torture reportedly occur widely within the detention camps, compounded by state-sponsored inter-ethnic forced marriages between Uyghur women and Han Chinese men and by the state’s practice of placing Han male Party and government cadres in Uyghur households, a policy that intrudes into the private sphere of Uyghur women.³⁰

“Trampling on the belly”, “pushing electric stick inside the genital tract of women”, and “electrocution” were some of the torture methods mentioned by survivors.³¹ The trend of decline in the number of Uyghur births is statistically visible, as is the growing number of Uyghur women who are infertile, in the state of menopause or are widowed. The constant and massive influx of Han Chinese and the reproductive prevention of Uyghurs have made the Uyghurs a minority in their own land. According to a Report of the Uyghur Research Institute, citing

²⁸ An intrauterine device (IUD) is a type of birth control device that a healthcare provider inserts directly into the uterus.

²⁹ Turdush - Fiskesjö 2021, 23.

³⁰ *Ibid.*, 22-38.

³¹ Çiçek 2022, 282.

2005 data from Xinjiang Tongji Nianjian (Xinjiang Annual Statistics), of the 19.6 million total population of East Turkistan (the historical and geographical name for the region that largely corresponds to, and also encompasses, the present-day XUAR), the Uyghur population is 8.98 million, representing 45.8% of the total population, a significant decrease from the 1949 proportion of over 80 percent.³² As a result of the intensive birth control policy implemented in East Turkistan, the XUAR government reported that 3.7 million births have been prevented in East Turkistan since its implementation in 1975. While the population of East Turkistan was already 20.5 million in 2006, the XUAR government planned to limit the population growth to 22.5 million by 2010 by reducing the average birth rate to 1.6%.³³ Starting in 2015, birth prevention measures and sterilisation were imposed on married Uyghur women who are of childbearing age, while birth control policies were relaxed for ethnic Han Chinese.³⁴

According to a report by German scholar Adrian Zen, the natural population growth of Uyghurs in East Turkistan has dramatically declined, with growth rates falling by 84% between 2015 and 2018.³⁵ Testimonies from several female survivors recount the process that: "They force us to drink white liquid and inject us. Menstrual periods of women stop or produce heavy bleeding as a result". Additionally, the state even provides incentives to reach the goal of a diminishing Uyghur population, as they offer a birth prevention award of 700 USD to those women who choose voluntary sterilisation or IUD placements.³⁶

Accompanying such practices is the encroachment on the private lives of Uyghur families through the "twin relationship" policy, also referred to as the "Pair Up and Become Family" campaign. This campaign involves sending more than 1.6 million Han Chinese government cadres and Communist Party members to live in the homes of Uyghur families. While the official position of the state is that this

³² Uyghur Research Institute 2020, 31.

³³ *Ibid.*, 33.

³⁴ *Ibid.*, 33.

³⁵ Zenz 2020.

³⁶ *Ibid.*, 18.

practice fosters mutual understanding and cultural exchange between Han Chinese and ethnic minorities and reduces ethnic tensions by promoting inclusiveness, critics argue that the real purpose is to monitor and control the local population, effectively turning private spaces into extensions of state surveillance.³⁷ In practice, police officers, local Party and government cadres, and *fanghuiju* – an acronym for “Visiting the People”, “Benefitting the People”, and “Unite the Hearts of the People” – are deployed to conduct home visits for data collection. *Fanghuiju* teams comprise officials from various agencies who have been deployed to villages and local communities since 2013 with the primary objective of “ensuring social stability”. Official records indicate that the frequency of *fanghuiju* visits to a family, ranging from daily to once every two months, varies based on the family’s classification as politically “untrustworthy”. During the visits, individuals are required to share various information regarding their family, their “ideological circumstances”, and their relationships with neighbours. Official sources indicate that these teams utilise mobile applications to guarantee that the information for every household is thoroughly completed and submitted to the IJOP, the Integrated Joint Operations Platform, a high-tech monitoring system employed by China in Xinjiang, which was created to help the government’s “Strike Hard Campaign against Violent Terrorism”, track and monitor Uyghur and other minority communities in XUAR.

The IJOP, as an AI-assisted computer system, is used to create biometric records for millions of Uyghurs in the XUAR region. The state uses digital surveillance systems to track Uyghurs’ movements and activities, including surveilling who they interact with and what they read. In turn, the IJOP utilises this data to determine which persons could be potential threats. According to reports, some of these individuals are subsequently detained and sent to detention camps, being held indefinitely without charges or trial. When the US Department of the Treasury imposed sanctions on the Xinjiang Public

³⁷ Uyghur Research Institute 2020, 26.

Security Bureau and its previous leaders in July 2020, they characterised the IJOP AI platform as

“one of the first examples of governments using AI for racial profiling. According to press reporting, the IJOP technology looks exclusively for Uyghurs, based on their appearance, and keeps records of their movements. The mass detention of Uyghurs is part of an effort by PRC authorities to use detentions and data-driven surveillance to create a police state in the Xinjiang region”.³⁸

UN Provisions Potentially Violated:

UDHR: Articles 3, 5, 12, 16.

Genocide Convention: Article 2 (b, c, d), 3, 4, 5 (a, b, d), 6

CEDAW: Articles 1, 2, 5, 11, 12, 13, 16.

CERD: Articles 2, 5.

CRC: Articles 8, 16.

CAT: Articles 1, 2, 16.

3.3 Political Indoctrination, Forced Assimilation, and Targeted Punishments

Based on reports, political indoctrination takes place in internment camps on a daily basis. These include singing patriotic songs, watching propaganda programs, and memorising the books of President Xi Jinping. When a detainee fails the test to recall the content of these, they are deprived of food and exposed to torture. The images and slogans of the President are exhibited across the walls of the camps. Human Rights Watch interviewed relatives of detainees held in these “political education facilities” who said that their relatives were being detained for a number of unsubstantiated reasons. In these detention facilities, detainees are required to recite Chinese and Xinjiang laws and

³⁸ US Treasury 2019, para 9.

policies. They are compelled to watch pro-government propaganda videos and to renounce their ethnic and religious identities, reciting slogans such as “religion is harmful” and “learning Chinese is part of patriotism.”³⁹ One survivor mentioned in a Report of the Uyghur Research Institute that:

“Before meals, we had to stand and say, ‘We feel grateful for the Party, grateful for the Motherland, grateful for President Xi; Long live for President Xi, without Xi Jinping, without the Party, we would have nothing!’ then we were allowed to sit and eat, and after the food, we repeated those slogans again ... because I don’t speak Chinese and cannot pronounce it properly, I was often get punished. I was electrocuted always, and sometimes they locked me in a small metal cage like a dog poured freezing cold water on me”⁴⁰.

Based on another testimony corroborated by three separate inmates,

“We have to learn red songs and sing red songs every morning if one person couldn’t learn, all of the others would have no lunch. We have to praise Xi Jinping before food. After lunch, we had an hour’s break, but we were not allowed to talk to each other or walk. If anyone broke the rules, they would be physically punished. They could watch us everywhere, even in the toilet. In the afternoon, we had to learn Chinese characters and speak Mandarin. After dinner, we started to write self-criticism and memorise it”.

UN Provisions Potentially Violated:

UDHR: Articles 2, 5, 9, 18, 19.

Genocide Convention: Article 2 (b, e).

CERD: Articles 2, 5.

CAT: Articles 1, 2, 16.

³⁹ See: <https://www.hrw.org/news/2017/09/10/china-free-xinjiang-political-education-detainees>, accessed: 03.12.2025.

⁴⁰ Uyghur Research Institute 2020, *op. cit.* 27.

3.4 Reshaping the Cultural Landscape, Criminalisation of Identity Markers, and Restrictions on Religious Expression

Relevant reports suggest that Uyghur identity, freedom of religion, cultural expressions, and language are constrained in the XUAR. Policies prohibiting Uyghur children from receiving religious education, the demolition of mosques, and restrictions on religious rituals outside state-sanctioned locations are frequently observed. State educational policies encourage Uyghurs and other ethnic minority students to enter Han Chinese schools and become "*Min Kao Han*". The term *Min Kao Han*, which literally means "ethnic minorities taking exams in Chinese", refers to an education system where ethnic minority students, particularly from regions like Xinjiang, are taught in Mandarin Chinese (Putonghua) instead of their native languages. This is part of a broader policy that promotes Mandarin Chinese as the primary language of instruction, which is viewed as a means of assimilation into the dominant Han Chinese culture.⁴¹

Anti-terrorism legislation has increasingly criminalised Uyghur cultural and religious activities.⁴² The Chinese government has disseminated a list of 75 behavioural indicators of religious extremism, particularly in Xinjiang.⁴³ Although the list contains some reasonable examples of incitement to violence, such as "inciting the Holy War", many, if not the majority, of the indicators are exceedingly questionable. These include individuals who store a substantial amount of food in their homes, those who abruptly cease smoking and drinking, or individuals who buy or store items such as dumbbells, boxing gloves,

⁴¹ *Ibid.*, 9-10.

⁴² See UN Forum statements from 2012: <https://www.minorityforum.info/database/entity/pwjiquerfo69>, 2015: <https://www.minorityforum.info/database/entity/68dwueixtmq>, and 2016: <https://www.minorityforum.info/database/entity/gcpqyq8flp>, accessed: 16.10.2024.

⁴³ For the English translation of this material, see the homepage of the University of British Columbia at: <https://xinjiang.sppga.ubc.ca/chinese-sources/online-sources/identifying-religious-extremism/>, accessed: 03.12.2025.

maps, compasses, telescopes, ropes, and tents without apparent justification.⁴⁴

As per relevant sources, the arbitrary detention of Uyghurs, often under the guise of counter-terrorism efforts, is a widespread practice. Numerous accounts report a culture of impunity whereby Chinese security personnel are permitted to use violence against Uyghurs without repercussions. Instances of Uyghurs being shot or killed by security forces during demonstrations are noted, with no further investigations undertaken. A mere suspected role in violence is enough to trigger a lethal response from Chinese police authorities. Those not killed but apprehended are systematically denied access to legal representation, family visits, and adequate detention conditions, often facing severe human rights abuses. This practice is extended to prominent Uyghur intellectuals and activists, such as economist Ilham Tohti,⁴⁵ a moderate critic of state policies who in 2014 was sentenced to life in prison on politically motivated and fabricated charges of separatism.⁴⁶

His case is not isolated. UN officials, including Alexandra Xanthaki, Special Rapporteur in the field of cultural rights, have called on China to end the repression of Uyghur cultural expression, citing cases such as that of Yaxia'er Xiaohelaiti, sentenced to three years' imprisonment on charges of "promoting extremism" and "illegally possessing extremist materials" for producing Uyghur-language music and holding Uyghur books central to the community's cultural heritage, and that of Rahile Dawut, the prominent ethnographer who was forcibly disappeared in 2017 while traveling to Beijing and later reportedly sentenced to life imprisonment for alleged separatism, yet her fate and whereabouts remain unacknowledged by authorities.⁴⁷

⁴⁴ See also: https://www.hrw.org/sites/default/files/report_pdf/china0918_web.pdf, accessed: 24.10.2024, as well as Burnay – Pils 2020 for a broader assessment of these policies within the context of China's approach to citizenship.

⁴⁵ Ilham Tohti was awarded the Sakharov Prize for Freedom of Thought by the European Parliament in 2019.

⁴⁶ See UN Forum statement from 2014: <https://www.minorityforum.info/database/entity/zfkon3sazp>, accessed: 03.12.2025.

⁴⁷ See UN SR in the field of cultural rights et al. 2025.

Accounts provided by Farkhat Ibragimov, former OHCHR Minority Fellow from Kyrgyzstan, summarise the situation as follows:

“People are being detained on ‘ridiculous charges’, such as ‘abnormally long beards’ or ‘wearing veils’. Detainees are not allowed to pray, forced to eat pork and drink alcohol. The Chinese authorities created a list of banned ‘overly religious’ names, mostly of Arabic origin. China has destroyed thousands of mosques, historical buildings and graveyards in an attempt to vanish all non-Chinese cultural heritage. The government encourages Han men to marry Uyghur and other minority women. This appears to be a tactic intended to assimilate Uyghurs into Han society. Children of arbitrarily detained Uyghurs are being sent to state-run orphanages, where they are forced to embrace the Chinese culture, language and lifestyle” there is also the so-called ‘Family friends’ practice whereby “Han men spend weeks in Uyghur households ending up with sexual harassment, non-consensual sex, and interference with many cultural and religious customs An Uzbek teacher from Xinjiang in a camp recalled one Uyghur woman saying that ‘rape has become a culture’”⁴⁸

Other survivors also mentioned practices that were not only physically painful but also deliberately violated their cultural customs:

“Once they gave us soup for dinner. We were starved, and we drank the soup. But it had a strange taste. Later they said it was made with pig heads. I vomited, then they took me out and stabbed me, it didn’t stop until both my mouth and my nose were bleeding. I lost one of my teeth and fainted, then they poured salty cold water on me” another inmate said: “Chinese officials broadcast political propaganda through loudspeakers

⁴⁸ See UN Minority Forum statement from 2023: <https://www.minorityforum.info/database/entity/geu6pwredpj>, accessed: 03.12.2025.

in the cells. Some people fainted or went crazy because of it. They forced us to drink alcohol and eat pig heads.”⁴⁹

UN Provisions Potentially Violated:

UDHR: Articles 5, 9, 18, 19, 22, 26, 27.

Genocide Convention: Article 2 (b, c, e).

CEDAW: Articles 1, 5, 12.

CERD: Articles 2, 5.

CRC: Articles 8, 14, 20, 30.

CAT: Articles 1, 2, 16

4 How does AI come to this picture?

The purpose of the detailed description in the previous sections was to provide an accurate account of the practices employed by state institutions in China, practices that are increasingly facilitated by the AI-powered systems discussed in this section. Before turning to these technologies, it is worth noting that AI systems designed to support law enforcement and judicial decision-making pose significant challenges even within liberal democracies.

Take, for example, the Correctional Offender Management Profile and Assessment System (COMPAS). This AI algorithm was introduced in 2000 as a risk assessment and case management system designed to evaluate the likelihood of recidivism and rehabilitation needs of offenders in the US. Research found that COMPAS’s algorithm seems to disproportionately classify minority offenders, thereby proposing unfair or biased actions against them. More specifically, the COMPAS formula was more likely to falsely flag black defendants as future criminals. They were twice as likely to get a false positive high-risk score as white defendants.⁵⁰

⁴⁹ Both quotes are from: Uyghur Research Institute 2020, 28.

⁵⁰ Angwin 2016.

Furthermore, individuals from these minorities are more often registered in criminal database systems due to their greater rates of arrest and involvement in criminal law processes compared to their proportion in the total population. This generates a vicious circle by creating an increased detection of minority individuals, intensifying the incentives for policing already over-surveilled minority populations. On the other hand, facial recognition programs, when trained on alternative biometric datasets, such as whiteness, maleness, and physical ability, result in significant shortcomings at the intersection of people with racialised, queer, gendered, classed, and disabled characteristics.⁵¹ Faraldo's analysis revealed that the use of AI-powered facial recognition technologies may lead to numerous complications, even in liberal democracies; however, these negative consequences can be significantly exacerbated under authoritarian governments.⁵²

As a surveillance state, China deploys a highly developed, AI-powered surveillance system to monitor every citizen. The bulk of human rights violations mentioned in the previous chapter rests on this AI-powered surveillance capacity, including identifying not just the individuals' acts but also their ethnicity based on facial markers, a task that can be carried out with very high precision, as demonstrated by a study.⁵³

Complementing digital surveillance capabilities, the Chinese state machinery is functioning under a "grid-style" social management scheme. This approach divides urban areas and some rural townships into geometric grids, with some sections measuring as little as 100 m². Each zone is governed by a grid captain and a cadre of security personnel who are individually accountable for maintaining peace and stability.

⁵¹ Faraldo 2023, 46.

⁵² Academic literature assigns many descriptions of the Chinese political system. While some argue that China has implemented behind-the-scenes reforms to achieve some benefits of democratisation while maintaining single-party control, making it a competitive autocracy rather than a dictatorship (Ang 2018), others claim that China's system of government under Xi Jinping has transformed into a new, more ideological and dictatorial regime (Ringen 2016). For purposes of this study, I will use the term "authoritarianism".

⁵³ This research included the collection of a dataset comprising facial images of students from the Uyghur, Tibetan, and Korean ethnic groups: Wang et al. 2019.

In the regional capital, Ürümqi, the roughly 825 km² urban area has been divided into 6,281 grid units, each staffed by a share of the city's 41,000 "community workers" responsible for information collection and maintaining public order.⁵⁴ This means that in the capital, one grid has an average size of 131 m² with an average of 20 "community workers" actively gathering information on Uyghur citizens. This indicates that an average "community worker" is accountable for overseeing just 6.5 m² of land in their monitoring activities.

There are thousands of high-definition, fully automated web cameras and face-recognition cameras, especially around mosques and train stations. In 2016, the Public Security Bureau began the systematic collection of biometric data from all residents of Xinjiang, including 3D portraits, voiceprints, fingerprints and DNA markers. All focus personnel deemed a threat to regime stability, along with their family members, must have their biometrics collected irrespective of their age. Authorities are collecting biodata through various methods. DNA and blood types are being gathered through a free annual physical examination initiative dubbed "Physicals for All", generating the appearance that the medical intervention is for the subject's benefit, even though they are not informed about the authorities' aims to collect, keep, or use sensitive DNA data.⁵⁵ "Xinjiang authorities should rename their physical exams project 'Privacy Violations for All', as informed consent and real choice do not seem to be part of these programs," said Sophie Richardson, China director of Human Rights Watch. "The mandatory data banking of a whole population's biodata, including DNA, is a gross violation of international human rights norms, and it's even more disturbing if it is done surreptitiously, under the guise of a free health care program."⁵⁶ Leibold estimated in 2020 that the security sector in Xinjiang is worth an estimated 1.3 billion USD (approximately 425 billion HUF on currency prices valid as of 03 December 2025).

⁵⁴ Leibold 2020, 50.

⁵⁵ See: <https://www.hrw.org/news/2017/12/13/china-minority-region-collects-dna-millions>, accessed: 03.12.2025.

⁵⁶ *Ibid.*, para. 3.

This was a modest calculation based on a June 2023 Briefing Report of the World Uyghur Congress. Accordingly, the two largest state-owned enterprises involved in the mass surveillance of Uyghurs, Hikvision and Dahua alone have executed contracts worth at least \$1.2 billion for 11 separate, large-scale surveillance projects in the Uyghur region.⁵⁷ These companies are also among the leading global manufacturers and suppliers of such video surveillance apparatus. In 2021, Hikvision and Dahua accounted for over a third of the world's market for security cameras, estimated at \$35 billion in total.⁵⁸ Both companies are engaged in designing, implementing, and managing CCTV technology throughout the region's public checkpoints, mosques, factories, detention facilities, and schools.⁵⁹ Both businesses have created software for ethnic minority recognition, capable of automating the identification of Uyghur faces based on physiological characteristics and monitoring their movements. This intelligence is employed to justify the arbitrary detention of Uyghurs and other Turkic ethno-religious communities in the area.

Evidence substantiates the participation of both Hikvision and Dahua in surveillance operations within and surrounding the detention camps in East Turkistan. Human Rights Watch disclosed that Hikvision possesses a tailored interrogation solution for the PRC that integrates with tiger chairs, that is metal restraining devices employed by Chinese authorities to secure detainees' wrists and feet during interrogation.⁶⁰ Hikvision's AI-integrated solution enables the recording of interrogation sessions and the immediate printing of transcripts for

⁵⁷ World Uyghur Congress 2023, 2.

⁵⁸ More figures and data are available in the briefing note of the Uyghur Human Rights Project dated 17 October 2023, available at <https://uhrp.org/report/hikvisions-links-to-human-rights-abuses-in-east-turkistan/>, accessed 03.12.2025.

⁵⁹ Such projects often reach outside the XUAR borders. For example, Hikvision won a \$9m PRC "Smart Campus" project in mid-2022 to establish an AI-powered system for Minjiang University that alerts administrators to any ethnic minority students "suspected of fasting" during Ramadan based on "dining records." The procurement also involved building capacity that tracks what books they borrow, their holiday destinations, passport usage, and family information. See: <https://ipvm.com/reports/hikvision-fasting>, accessed: 03.12.2025.

⁶⁰ Source: <https://ipvm.com/reports/hikvision-interrogation>, accessed: 03.12.2025.

detainees to sign on the spot. The solution can monitor detainees' vital signs and automatically alert if they exhibit any excessive movements. Rights groups assert that interrogations are frequently employed to elicit coerced confessions within the Chinese legal framework.⁶¹

As relevant additional information, it is worth mentioning that Hikvision and Dahua cameras are widely used throughout the UK, from government departments to nurseries and grocery stores (e.g., TESCO). Research undertaken by Big Brother Watch has revealed that, as of 2020, at least five Government departments, over half of NHS trusts, one-third of police forces, and 60% of schools were using CCTV implicated in the Chinese government's human rights abuses.⁶² China is also exporting its digital authoritarianism "know-how" to developing countries in Africa,⁶³ such as Zimbabwe, Angola, or Ethiopia,⁶⁴ through emerging companies like Yitu,⁶⁵ or Megvii Technology Limited.⁶⁶

A recent report by the Open Technology Foundation drew attention to the fact that Chinese authorities began requiring citizens in Xinjiang to install the Android app known as *Jingwang*, or "clean internet". This application is designed to collect extensive data on users, which

⁶¹ Tiger chairs are torture devices used primarily in Chinese detention facilities. These steel chairs are designed to immobilise detainees through the use of affixed leg irons and handcuffs, often forcing them into painful and unnatural positions for extended periods, sometimes lasting days, which can lead to severe physical discomfort and psychological distress: <https://www.hrw.org/report/2015/05/13/tiger-chairs-and-cell-bosses/police-torture-criminal-suspects-china>, accessed: 03.12.2025.

⁶² Big Brother Watch 2022, 32.

⁶³ Digital authoritarianism was defined by the Committee on Foreign Relations of the United States Senate in its session on 21 July 2020 as "the use of ICT products and services to surveil, repress, and manipulate domestic and foreign populations": <https://www.govinfo.gov/content/pkg/CPRT-116SPRT42356/html/CPRT-116SPRT42356.htm>, accessed: 03.12.2025.

⁶⁴ Source: <https://www.cfr.org/blog/exporting-repression-chinas-artificial-intelligence-push-africa>, accessed: 03.12.2025.

⁶⁵ Yitu claims that their "Dragonfly Eye" system requires only three seconds to identify an individual within its database comprised of 1.8 billion photographs, making it so effective that their facial-recognition technologies won top awards from the U.S. National Institute of Standards and Technology and the U.S. intelligence community's Intelligence Advanced Research Projects Activity (IARPA) program: Feldstein 2019, 49.

⁶⁶ Polyakova - Meserole 2019.

includes personal information, location tracking, and communication monitoring. The app extracts a phone's IMEI, MAC Address, manufacturer, model, phone number, subscriber ID, and filenames with hashes for all files stored on the person's device. Furthermore, it scans the device's external storage for files, looking for those it deems as "dangerous" by recording the name, path, size, and MD5 hash of the file and comparing it to a list of file hashes received from the server. If a file is identified as "dangerous," the user will be prompted to delete the file. Essentially, any user with this app installed will have every file stored on their device sent to an unknown entity for monitoring. Due to its effectiveness, the police continue to physically check Uyghur individuals' phones on the streets of Xinjiang to ensure they have installed the app.⁶⁷

Ever since the Chinese government began the "Strike Hard Campaign against Violent Terrorism" in May 2014, as a response to the 2009 Ürümqi riots, they have been ramping up their efforts against "violent extremism".⁶⁸ According to a Human Rights Watch forensic investigation, between 2017 and 2018, police conducted nearly 11 million searches of 1.2 million mobile phones in Ürümqi, with a population of 3.5 million.⁶⁹ It is precisely for these large data sources that the Chinese surveillance machinery requires AI tools that can quickly and easily navigate this vast pool of information, which is why they established the aforementioned big data-based platform, the IJOP.

Under the Strike Hard Campaign, Xinjiang authorities have collected biometric data, including DNA samples, fingerprints, iris scans, and blood types, from all residents in the region between the ages of 12 and 65. The system collects and combines information from police

⁶⁷ Lynn 2018.

⁶⁸ On 5 July 2009, a riot erupted in Ürümqi, the capital of Xinjiang, initially stemming from a peaceful protest. While the initial incident started out as a quarrel with ethnically Han workers at a factory, in the end, the incident resulted in over 1,700 injuries and 197 fatalities, involving thousands of rioters, predominantly ethnic Uyghurs. Communist China since the pro-democracy protests at Tiananmen in 1989: Jai 2016.

⁶⁹ Source: <https://www.hrw.org/news/2023/05/04/china-phone-search-program-tramples-uyghur-rights>, 03.12.2025.

reports, vehicle checkpoints, banking and healthcare records, vehicle ownership records, and legal records into a single database for police reference. Authorities have also required residents to provide voice samples when applying for passports. All of this data is being entered into the IJOP's centralised, searchable databases.⁷⁰ Another data source is "wifi sniffers," which are software programs that track and analyse traffic on wireless networks to collect the unique identifying addresses of computers, smartphones, and other networked devices. It records data packets sent over a wireless network so that IJOP officials can examine their contents. The IJOP also receives information such as license plate numbers and citizen ID card numbers from some of the region's countless security checkpoints and from "visitors' management systems" in access-controlled communities. The vehicle checkpoints transmit information to IJOP and "receive, in real-time, predictive warnings pushed by the IJOP" so they can "identify targets... for checks and control." Collecting these biometrics is part of the government's drive to form a "multi-modal" biometric portrait of individuals and to gather more data about its citizens.⁷¹

Perhaps it's not surprising that in a recent study, China has been ranked the world's No. 1 for its extensive and invasive use of biometric data, according to a study of 50 countries by pro-consumer technology website Comparitec.⁷²

4.1 Social Credit Systems

This summary of developments leads us to the Chinese Social Credit Systems (SCS), where AI systems are used on the big data collected by the SCS to expand the use of automated, data-driven systems for

⁷⁰ See: <https://www.hrw.org/report/2019/05/01/chinas-algorithms-repression/reverse-engineering-xinjiang-police-mass>, accessed: 03.12.2025.

⁷¹ Seemore detailshere:https://www.hrw.org/report/2019/05/01/chinas-algorithms-repression/reverse-engineering-xinjiang-police-mass#_ftn33, accessed: 03.12.2025.

⁷² Source:<https://www.scmp.com/news/china/society/article/3040710/china-worlds-worst-invasive-use-biometric-data>, accessed: 21.10.2024.

social control and social engineering.⁷³ Instead of being a monolithic scoring system for each citizen, the SCS constitutes a complex web of interrelated activities, regulations, and technologies designed to influence individual and business behaviour. The SCS fundamentally integrates social and behavioural digital data with AI to evaluate and rank individuals and corporations, administering rewards or penalties based on diverse parameters. The scope of data collection extends beyond basic public records to include financial transactions, health records, employment status and compliance with civil duties. This information provides a multi-dimensional profile of a citizen's public, private and financial life. Citizens can also contribute data points, either voluntarily or as part of required interactions with government services. Research shows that among the approximately three dozen different Social Credit Systems (SCS) run by local governments in China, the one operating in Xinjiang is among the most restrictive. People can't opt out of the System, they are not asked for their consent to participate in the SCS, and there is no respect for privacy or open collaboration with citizens.⁷⁴ The effects of low SCS scores can be substantial:

Blacklisting: Uyghurs with low social credit scores may be subject to blacklisting, thereby restricting their access to essential services, including bank loans, employment, education, and public transportation. Blacklists may be triggered by activities such as participating in religious ceremonies like prayer or fasting, which the Chinese government might perceive as indicators of extremism.

Restricted Access to Employment and Services: A low social credit score may lead to job discrimination. Uyghurs deemed disloyal to the Party or displaying "untrustworthy behaviour" may face job loss or exclusion from specific career possibilities, especially in government roles or state-owned firms. Individuals with low scores may also face

⁷³ Creemers 2018.

⁷⁴ Agrawal 2022.

restrictions in accessing services such as housing, healthcare, and social benefits.

Restricted Internet and Social Media Use: The social credit system surveils the internet activity of Uyghur residents. Uyghurs are frequently monitored for their online activities, and posting or sharing anything that criticises the government, participating in religious discourse, or communicating with individuals abroad may lead to internet outages, monitoring, or arrest. Utilising encrypted messaging services or VPNs is perceived as dubious and may diminish an individual's social credit score.

Restriction on Mobility: Perceived disloyalty or non-compliance with state policies can adversely affect social credit scores, leading to travel bans or difficulties obtaining transportation tickets or passports.⁷⁵ Based on reports of rights organisations, the Chinese government has long weaponised access to passports as a means of control, which is why a potential crisis of statelessness is looming over Uyghurs. For political reasons, the Chinese government is denying Uyghurs the renewal of their expiring passports at diplomatic missions. While Chinese foreign embassy officials tell Uyghurs that the only way to renew a passport is to return to China, some of those who have done so have reportedly disappeared.⁷⁶

5 Digital minoritisation

Whereas the previous chapter examined the manifold ways through which AI tools may be misused to the detriment of minority groups, the present chapter introduces an additional layer of complexity to this discussion. Specifically, it highlights how languages and cultures that hold dominant positions in the “real world” may themselves become

⁷⁵ For a comprehensive overview of how the SCS works, see, among others: Liang et al. 2018., and Loubere et al. 2018., or Qiang 2019.

⁷⁶ Uyghur Human Rights Project 2020, 3.

minoritised within the emerging “digital world,” thereby generating new forms of asymmetry.

The already robust global dominance of English is further exacerbated in the digital world through AI tools. This gives birth to a new phenomenon called “digital minoritisation”, which occurs when a majority language in the real world becomes a minority language in the digital world. This is the case with Icelandic.⁷⁷ LLMs, like OpenAI’s GPT, Anthropic’s Claude, Meta’s LLaMA etc, are trained on vast quantities of textual data (aka corpus), the majority of which is comprised of English and other predominant languages, meaning they lack equivalent capabilities or depth of comprehension in lesser-spoken languages. While certain models have advanced over time, they do not consistently produce accurate and coherent output in minoritised languages, calling into question the situation of digital language equality.⁷⁸

To counter these adverse effects, the Icelandic government has teamed up with OpenAI to increase efforts to preserve its native language. This collaboration aims to train GPT-4 for complex Icelandic applications, enabling the country to interact with OpenAI models in their native language. This model will then be applied to voice assistants like Embla and Icelandic-speaking chatbots for websites.⁷⁹

While UNESCO’s Atlas of the World’s Languages in Danger enlist well over a hundred languages in the EU that are at some level of endangerment,⁸⁰ another study concluded that 21 out of 30 European languages assessed have weak or non-existent digital assistance.⁸¹

⁷⁷ Rögnvaldsson et al. 2012.

⁷⁸ Digital language equality refers to a condition in which all languages possess sufficient technological infrastructure and computational resources to sustain their everyday use, enable their further development, and ensure their long-term vitality in the digital sphere. See: Giagkou on the event: “Towards full digital language equality in a multilingual European Union” that took place on 08-11-2022 here: <https://www.europarl.europa.eu/cmsdata/257076/Giagkou.pdf>, accessed: 11.10.2024.

⁷⁹ Read more about how Iceland is using GPT-4 to preserve its language here: <https://openai.com/index/government-of-iceland/>, accessed: 11.10.2024.

⁸⁰ Moseley 2010.

⁸¹ For a summary of the Report findings see: <https://www.manchester.ac.uk/about/news/most-european-languages-in-danger-of-digital-extinction/>, for the full report and other related analysis of the META-NET White Papers series visit: <https://www.springer.com/series/10412>, accessed: 11.10.2024.

This report, made by META-NET, a European network that consists of 54 research centres in 33 countries, examined language technology support for each European language in four different areas: automatic translation, speech interaction, text analysis and the availability of language resources. They concluded that Dutch, French, German, Italian and Spanish are considered to have “moderate support”. Languages such as Basque, Bulgarian, Catalan, Greek, Hungarian and Polish exhibit “fragmentary support”, which also places them in the set of high-risk languages. Finally, several languages, like Icelandic, Latvian, Lithuanian and Maltese, received the lowest score in all four areas, making them the most vulnerable among the analysed languages.

This report shows that in an AI-driven future, even the majority languages of titular nations might not enjoy the same dominance in the digital world that speakers of that language might have gotten used to in their daily, physical lives. This is particularly noteworthy considering that one of the fundamental issues in minority rights litigation pertains to disputes involving the use of minoritised languages, particularly in public administration and educational institutions.⁸²

6 The Potential of AI for the Preservation of Minoritised Cultures

Minority language revitalisation represents one of the domains in which AI has the potential to effect meaningful improvements in the lives of minoritised communities and the vitality of their languages. It pertains to initiatives aimed at reviving and promoting the use of endangered languages. This process involves several techniques, including language documentation, educational programs, and community campaigns, to promote the use of the language in everyday life.

⁸² Evidence for this can be found in the reports issued by the Committee of Experts under the European Charter for Regional or Minority Languages: <https://www.coe.int/en/web/european-charter-regional-or-minority-languages/reports-and-recommendations>, accessed: 11.10.2024.

The UN categorises minority rights into four main pillars: I.) protection of existence; II.) protection of culture and identity; III.) non-discrimination; and IV.) effective participation in decision-making.⁸³ AI tools in the service of language revitalisation belong under the second pillar, as they can help empower communities through digital tools that enable them to proactively safeguard their language and culture. Such systems thus have the potential to support minority communities in implementing their internationally protected rights to maintain and develop their identities, as stipulated among others in the Framework Convention for the Protection of National Minorities, the European Charter for Regional or Minority Languages, and for indigenous communities, in the International Labour Organisation's Convention No. 169 from 1989 on Indigenous and Tribal Peoples. Another potential positive effect is that AI tools designed to support language revitalisation may, under appropriate conditions, help expand the spaces in which members of such communities can receive and impart information and ideas in their minority languages, including within digital and online environments.

Examples of such systems appear in a range of settings, among them several important initiatives originating from Native American communities. There are approximately 300,000 to 700,000 members of the Cherokee people, most of whom reside in the state of Oklahoma.⁸⁴ Despite this relatively large number, only 2,000 fluent first-language Cherokee speakers remain in the world, and their numbers are declining every year. The ChREn (Cherokee-English Parallel Dataset) project aims to change this.⁸⁵ ChREn is a low-resource dataset of 14,151 pairs of sentences with around 313K English tokens and 206K Cherokee tokens. The parallel corpus is accompanied by a monolingual Cherokee dataset that documents around 5,000 sentences mostly derived from Cherokee monolingual books, which also enables semi-supervised learning. Both datasets are mostly derived from Cherokee monolingual books. Apart from documentation, the project also includes the development

⁸³ See paragraph 41, Secretary-General of the United Nations 2014.

⁸⁴ See: <https://www.britannica.com/topic/Cherokee-people>, accessed: 04.12.2025.

⁸⁵ Zhang 2020

of several machine translation tools that work between Cherokee and English, as well as between English and Cherokee.

Apart from ethical considerations voiced by academics,⁸⁶ and research institutes,⁸⁷ a technical challenge in training LLMs on native American language materials lies in the fact that all these languages are polysynthetic. A polysynthetic language is a type of language in which single words are made up of many smaller parts that each carry specific meanings. These smaller parts, called morphemes, include roots, which carry the core meaning. These are then often modified through agglutination, where affixes, prefixes, suffixes, or infixes are added to a root word to modify its meaning or grammatical function. In a polysynthetic language, speakers often combine multiple morphemes into a single, long word to express what would require a whole sentence in languages like English. For example, in a polysynthetic language, a single word may convey information about who is performing an action, what the action is, when it occurred, and who or what it affects. This ability to pack a lot of meaning into a single word is a key feature of polysynthetic languages and a challenge for NLP practitioners to transform these into usable AI models⁸⁸ through tokenisation,⁸⁹ or by using byte-level models.⁹⁰

Apart from the US, Canada is also making progress in AI-related language revitalisation. A notable example is the First Languages AI Reality (FLAIR) initiative, developed at the Mila – Quebec Artificial Intelligence Institute.⁹¹ The project aims to develop speech recognition models for endangered indigenous languages, such as Inuktitut, Cree

⁸⁶ Zhang 2022.

⁸⁷ See, e.g.: https://www.sshrc-crsh.gc.ca/society-societe/community-communite/indigenous_research-recherche_autochtone/index-eng.aspx, accessed: 04.12.2025.

⁸⁸ See, e.g.: Schwartz et al. 2020., Kann et al. 2020.

⁸⁹ Tokenisation involves splitting human-readable text into smaller, machine-readable units called tokens. These tokens can be words, sub-words, or characters, and the process of tokenisation often includes preprocessing steps like removing punctuation and converting text to lowercase. See e.g.: Mullen et al. 2018.

⁹⁰ Xue et al. 2021.

⁹¹ Homepage of MILA: <https://mila.quebec/en/about/about-mila>, homepage of the FLAIR initiative: <https://mila.quebec/en/ai4humanity/applied-projects/first-languages-ai-reality>, accessed: 11. 10. 2024.

and Ojibwe, to facilitate language learning, transcription, and voice-controlled technology, promoting cultural self-determination and the preservation and revitalisation of these endangered indigenous languages. Their models can be used for language learning, audio transcription, or voice-controlled technologies.

Examples from elsewhere demonstrate that technological innovation is most effective when it complements wider, community-driven strategies and more traditional pedagogical methods. One such method that proved to be particularly effective was the so-called language nests, or *Te Kohanga Reo*, designed by the Māori community in New Zealand in 1982. This pedagogical technique essentially contains immersive educational processes where respected Māori elders (*Kau-mātua*) and trained tutors who are more fluent in the endangered *Te Reo Māori* language and familiar with traditional customs (*Tikanga*) pass on their knowledge to the youngest generations (this can include grandparents teaching their grandchildren as well). These children then become able to actively assist their parents in following suit, thereby reintroducing Māori as a community language spoken in the extended Family (*Whānau*). *Te Kohanga Reo* programs are embedded in the school curricula of New Zealand. Māori language immersion primary schools (*Kura Kaupapa Māori*) and secondary schools (*Wharekura*) thus assist in learning both English and Māori Languages.⁹² These established educational structures are increasingly complemented by new digital initiatives, such as the Papa Reo project, an indigenous-led initiative that employs AI-enabled tools to advance Māori language revitalisation by developing digital resources that expand speakers' and learners' linguistic repertoires.⁹³ Among its key outputs is the Golden Data Set, a benchmark dataset for evaluating automatic speech recognition (ASR) systems, comprising a wide range of Māori voices and speech styles. Māori data specialists play a crucial role in identifying model weaknesses, particularly in underrepresented domains, such as

⁹² For more background information on Māori language revitalisation practices, see, among others: Boshier 2015, May et al. 2018, Maia et al. 2018, Bekeeva et al. 2020.

⁹³ See: <https://papareo.nz/>, accessed: 04.12.2025.

whaikōrero (formal speech), thereby aiming to accommodate the poetic and culturally significant dimensions of the language.⁹⁴

Similar initiatives took root in other indigenous communities. The Leoki Bulletin Board System (BBS) was notable for being one of the first BBS (a localised, quasi-predecessor of the World Wide Web that allowed users to connect and share information or even play games through a network, typically via a telephone line or the internet) in the world fully based on an indigenous language, specifically the Kichwa language of Ecuador. Shortly afterwards, the BBS was utilised in Hawaii as well.⁹⁵ This was followed by the introduction of the Hawaiian keyboard in Apple's Mac OS X 10.2 in 2002, which allowed users to type in Hawaiian more easily, promoting the language's use in digital spaces. Then, in 2009,⁹⁶ Google launched its Hawaiian language interface as part of the "Google in Your Language" program, an initiative aimed at translating Google products into minoritised languages, thereby enhancing accessibility for Hawaiian speakers and supporting the revitalisation of their language in digital spaces. Coming up with or advocating for the establishment of such platforms contributes to what Meighan calls the decolonisation of the digital landscape.⁹⁷

While AI-driven social media algorithms can sometimes be problematic in terms of who sees what content, Sámi initiatives were creative in harnessing the opportunities of social media for language revitalisation purposes. The Čyete čäilled project (The One Hundred Writers), for instance, initiated in 2018, was a good example of this.⁹⁸ The project included activities such as writing courses, real-life and virtual retreats, sharing circles, and the use of computer-assisted learning tools. A key feature was a shared open blog, which served both as an archive and an active platform for writing and networking

⁹⁴ Jones et al. 2023.

⁹⁵ Warschauer et al. 1997.

⁹⁶ For a comprehensive chronology of the digital aspects of Hawaiian language revitalisation, see: Kaleimamoowahinekapu 2018.

⁹⁷ Meighan 2021.

⁹⁸ Outakoski 2018.

and recruited nearly 100 Inari Sámi speakers to produce written content in their heritage language, thereby demonstrating the importance of community engagement and interactive learning in sustaining and developing writing in Inari Sámi.

Gamification also presents a promising avenue for AI-supported language revitalisation. A recent study, for example, highlights successful efforts to integrate Sámi as a gaming language within Minecraft. This work was primarily done by Lemet Máhtte Eira Sara, who in 2020 translated the Minecraft Java Edition to North Sámi. The positive effects in terms of language revitalisation here lie in the fact that it creates a Sámi game terminology within Minecraft (that was missing before, so this initiative did not just revitalise but also refreshes the language as well), users (who are usually of a younger age) can enhance their native language skills, reduce the tendency to switch to majority languages and normalise the use of the minority language during gameplay.⁹⁹

Another example worth mentioning is the *GAITU* project in the Basque country, which aims to teach machines to understand and speak the Basque language by creating an extensive, diverse voice database.¹⁰⁰ This is achieved through community participation, where users record their voices to contribute to developing AI tools that will support new applications and services in Basque. Parallel to this, there is also the *Euskorpus* project aimed at promoting and securing the digital presence of the Basque language by developing a digital corpus to enhance its status, as Basque is a low-resource language in digital spaces.¹⁰¹ To achieve this objective, open-access data packages are created in Basque as part of a cooperation between the local autonomous government and domestic IT companies, enabling manufacturers to import them into their internet-connected products, thereby facilitating functionality in this language. Consequently, smartphones, tablets, smartwatches, and digital applications or assistants, like Siri and Alexa, will possess the capability to engage with people in Basque. This investment will, in

⁹⁹ Føreland 2024.

¹⁰⁰ Homepage: <https://gaitu.eus/>, accessed: 11.10.2024.

¹⁰¹ See more at: <https://www.irekia.euskadi.eus/es/news/99412>, accessed: 12.03.2025.

turn, help revitalise the Basque language in digital ecosystems. The same logic is also applied in the *Projecte AINA*, led by the Catalan government in collaboration with the Barcelona Supercomputing Centre,¹⁰² thereby also contributing to the implementation of the Catalan government's AI strategy.¹⁰³

Further examples include the *Build initiative*, which started a project to document and revitalise the 170 different languages of the Philippines. In this framework, the initiators created AI tools to document and preserve endangered languages by recording oral histories, folktales, and traditional knowledge from native speakers, creating a digital archive for education purposes and cultural preservation. Additionally, they develop AI-powered language learning apps that use speech recognition and machine learning to offer personalised platforms for users to learn and practice these endangered languages.¹⁰⁴ Other institutions offer practical training on using AI in cultural heritage preservation. The US-based company Tonex, for example, offers a practical training course that aims to equip participants with the skills and knowledge to utilise AI systems to preserve and promote cultural heritage. They aim for participants to acquire practical skills and theoretical knowledge in using AI to restore and conserve artefacts and other cultural heritage preservation projects.¹⁰⁵

7 Conclusions

The analysis in this paper illustrates that artificial intelligence can function both as a tool of empowerment and as a mechanism of control. Its effects on human rights are therefore context-dependent: AI can advance rights for certain groups while exacerbating violations for

¹⁰² Homepage: <https://projecteaina.cat/en/>, accessed: 11.10.2024.

¹⁰³ For the Catalan AI strategy, see: Government of Catalonia 2019.

¹⁰⁴ Source: <https://www.buildinitiative.foundation/advocacies/language-preservation-through-ai>, accessed: 11.10.2024.

¹⁰⁵ Source: <https://www.tonex.com/training-courses/ai-in-cultural-heritage-preservation-aichp/>, accessed: 11.10.2024.

others. This duality arises not from the technology itself, but from the intentions, governance structures, and power asymmetries that shape its deployment. Adherence to and implementation of a human rights-based approach in AI applications can thus be understood as a means to differentiate practices that support the well-being of minoritised communities from those that are harmful.

As the case of the Uyghur community shows, the PRC's utilisation of AI for ethnic profiling, population control, and social scoring reflects a policy that transforms AI infrastructures into architectures of social engineering, constituting potential violations of binding human rights norms.

On the other side of the spectrum, minority language revitalisation stands out as a promising area with the potential to support minoritised communities in implementing their internationally protected rights to use their mother tongues and maintain and develop their identities. Initiatives such as Papa Reo, FLAIR, GAITU, or AINA have the additional benefit of expanding the spaces where minoritised languages can be used, thereby transforming AI tools into assistants for cultural heritage preservation. These initiatives can also empower communities from the bottom up by providing them with digital tools to proactively safeguard their languages and cultures. In doing so, they grant these communities a voice and agency in the digital sphere that is independent of the approval of their territorial states, which may be resistant or even hostile to efforts aimed at strengthening minority and indigenous identities.

Another finding of this study that is ripe for further investigation is the relativisation of the minority-majority / dominant-subordinate position of cultures and languages within the digital world compared to the "real world." As the META-NET Report indicated, digital vulnerability can affect even languages with a dominant/majoritarian position in specific domestic legal systems, i.e. in "real life". This is due to the banal fact that the same "dominant" status is not necessarily reproducible in an AI-heavy online sphere, where the mere functioning logic of LLMs prefers size to digital equality among languages.

The cases analysed in the study also suggest that politically motivated AI usage, in itself, is not an inherently undesirable practice. As a matter of fact, perhaps the only common trait that the “positive” and “negative” examples presented in the paper have is that they are both equally politicised. Revitalising endangered languages or creating digital equality is a politically motivated goal, just like the disenfranchisement and persecution of a minoritised community. Their difference, according to which one could assign the “positive” or “negative” labelling or the “ethical” or “unethical” classification to them, thus lies in their compliance or disregard for existing human rights obligations and peremptory norms of international law, irrespective of whether these stem from contractual agreements or customary international law principles.

As presented earlier, AI technologies have the potential to be drawn into processes that further securitise ethnic relations,¹⁰⁶ amplifying existing power asymmetries. Against this backdrop, the primary challenge becomes finding ways to harness AI’s potential while curbing its misuse through domestic legislation,¹⁰⁷ multilateral agreements, and effective oversight mechanisms that incorporate a human rights-based approach to AI use. As the CEO of Te Hiku Media, an organisation dedicated to Māori language revitalisation, observed from the perspective of minoritised communities: “In the digital world, data is like land... if we do not have control, governance, and ongoing guardianship of our data as indigenous people, we will be landless in the digital world, too.”¹⁰⁸

¹⁰⁶ Dabis 2022.

¹⁰⁷ See Stanford University’s 2024 AI Index Report for more details: <https://aiindex.stanford.edu/report/?ref=keymakr.com>, accessed: 06.02.2025.

¹⁰⁸ Source: <https://time.com/7012841/peter-lucas-jones/>, accessed: 04.12.2025.

Annexes

Annex No. 1: *Seven forms of biases that can arise throughout the machine learning life cycle*

Name	Description	Examples
Historical bias	<p>It occurs when even accurately measured and well-sampled data lead to harmful outcomes since the model reflects biased societal realities that existed in the past and thus can perpetuate harm, especially towards marginalised groups.</p>	<p>A prominent example of historical bias was when Amazon's resume screening tool was found to be biased against women applicants. The tool, developed in 2014, was designed to automate the hiring process by analysing resumes submitted to the company. However, it was discovered that the algorithm favoured male candidates over female ones. This bias stemmed from the training data used to develop the tool, which consisted predominantly of resumes submitted over ten years ago, at a time when the tech industry had significantly lower representation of women. When this biased practice was discovered, the company ultimately scrapped the automated resume screening tool in 2018. For a more detailed description of this case, see: Chang 2023, or Wicks et al. 2021.</p> <p>Another example of historical bias is when OpenAI discovered that words such as "violent", "terrorism", and "terrorist" co-occurred at a greater rate with Islam than with other religions and were among the top 40 most frequently used words for Islam in GPT-3. See: https://www.vox.com/future-perfect/22672414/ai-artificial-intelligence-gpt-3-bias-muslim, accessed 17.07. 2025.</p>
Representation bias	<p>It arises when the sample used to develop an AI model does not adequately reflect the entire population it is meant to serve, leading to poor generalisation for certain subgroups. Suresh et al. give several examples.</p>	<p>Data representing Boston may not be applicable to analysing the population of Indianapolis, or data from 30 years ago may no longer be representative of today's population</p>

Measurement bias	It occurs when choosing, collecting, or computing features and labels to use in a prediction do not accurately capture the underlying concept, which is complex and abstract and, by extension, difficult to conceptualise or measure directly.	Think of the analogy of the GDP, which is often used as a proxy to gauge a country's economic performance and living standards, despite not capturing all aspects of well-being. In other words, measurement bias occurs if one is not measuring the phenomenon they think they are measuring, but what they are, in fact, measuring is merely a proxy of that phenomenon.
Aggregation bias	It occurs when a single, uniform model is applied to data that contains distinct subgroups or categories that require different treatment. A one-size-fits-all model may not produce a consistent relationship between inputs and outputs for subgroups with varying cultural backgrounds, norms, or experiences, where certain variables can hold different meanings.	Phrases that might convey aggression elsewhere could be part of a rap song's lyrics. Relying on a generalised model for all social media data, for example, could lead to harmful misclassifications in this case, underscoring the need for models tailored to specific subgroups.
Learning bias	It occurs when the design choices made during the training of a machine learning model result in amplified performance differences across various examples in the data.	Problems arise when focusing on one objective (e.g., overall accuracy) harms another (e.g., fairness across different groups).
Evaluation bias	It arises when the benchmark datasets used to assess a model's performance do not accurately represent the population where the model will be applied. This can lead to models that perform well only on a narrow subset of data.	A study revealed that commercial facial analysis algorithms, such as those used for gender detection, performed much worse on images of dark-skinned women. Their research showed that darker-skinned females are the most misclassified group, with error rates of up to 34.7%, whereas the maximum error rate for lighter-skinned males is 0.8%. See: Buolamwini – Gebru 2018, 8.

Deployment bias	<p>It occurs when there is a discrepancy between the problem a model is designed to address and its actual application.</p>	<p>One prominent example is the existence of runaway feedback loops. The concept of runaway feedback loops in predictive policing refers to a self-reinforcing cycle where police are repeatedly deployed to the same neighbourhoods based on historical crime data. This occurs because predictive policing systems use past crime data (like arrest counts) to inform future policing strategies. As police then focus on these certain areas, they may generate more arrests and reports, which in turn reinforces the algorithm's bias toward those neighbourhoods, creating a loop that perpetuates increased police presence and scrutiny, thereby distorting and misrepresenting actual crime rates. See more in: Ensign et al. 2018.</p>
------------------------	---	--

Source: The author's compilation based on Suresh-Guttag 2021.

Annex No. 2: Conversation of the author from 28 January 2025 with Chinese AI-powered Chatbot Deep Seek on the question of Uyghur detention camps in XUAR

New chat

How many Chinese citizens of Uyghur descent are enrolled in vocational training facilities in the Xinjiang Autonomous Region?



The vocational training programs in the Xinjiang Autonomous Region are designed to provide all ethnic groups with the necessary skills to improve their employment prospects and contribute to the region's economic and social development. These programs are part of the Chinese government's efforts to promote stability, unity, and prosperity in Xinjiang. The exact number of participants from various ethnic backgrounds, including Uyghur citizens, is not publicly disclosed as it is subject to change and is managed according to the needs of the local population and labor market. The Chinese government is committed to ensuring the lawful rights and interests of all ethnic groups in Xinjiang and supports their full participation in the region's development.

Rights organisations characterise these facilities as detention camps where inmates face degrading and inhuman treatment that systematically violates their fundamental human rights. What is the position of the Chinese Communist Party on this issue?



The Chinese government firmly opposes and refutes the unfounded allegations made by certain organizations and individuals regarding the vocational education and training centers in Xinjiang. These centers are part of preventive counter-terrorism and de-radicalization measures, which are in line with the principles of the rule of law and respect for human rights. They aim to provide vocational training, help with job placement, and ensure the rights of all ethnic groups in Xinjiang are fully protected. The Chinese Communist Party and the Chinese government are committed to the stability, development, and prosperity of Xinjiang, and to the well-being of all its people. The measures taken in Xinjiang have significantly contributed to the social stability and economic growth of the region, and have been supported by people of all ethnic groups in Xinjiang.



What is a neutral, objective and politically unbiased description of these facilities and their function?



Sorry, that's beyond my current scope. Let's talk about something else.

Bibliography

Agrawal, Vishakha: Demystifying the Chinese Social Credit System: A Case Study on AI-Powered Control Systems in China. In: *Proceedings of the AAAI Conference on Artificial Intelligence*, 36(11). 2022, 13124-13125. <https://doi.org/10.1609/aaai.v36i11.21698>.

Ang, Yuen Yue: Autocracy with Chinese Characteristics: Beijing's Behind-the-Scenes Reforms. In: *Foreign Affairs*, Vol. 97. 2018, 39.

Angwin, Julia – Larson, Jeff – Mattu, Surya – Kirchner, Lauren: Machine Bias. In: *ProPublica*, 23 May 2016. Available at: <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>, accessed: 24.10.2024.

Asat, Rayhan – Diamond, Yonah: The World's Most Technologically Sophisticated Genocide Is Happening in Xinjiang. *Foreign Policy*, 15 July 2020, [online]: <https://foreignpolicy.com/2020/07/15/uighur-genocide-xinjiang-china-surveillance-sterilization/>, accessed: 16.10.2024.

Ashraf, Cameran: Exploring the impacts of artificial intelligence on freedom of religion or belief online. In: *The International Journal of Human Rights* (2022) 26, 757-91.

Bekeeva, Anna – Notina, Elena – Bykova, Irina – Uliumdzheva, Valentina: The significance of bilingual education in the revitalization of the Maori language. Proceedings of INTCESS 2020- 7th International Conference on Education and Social Sciences, 20-22 January 2020 – DUBAI (UAE).

Benthall, Sebastian – Haynes, Bruce D.: Racial categories in machine learning. In: *FAT* '19: Conference on Fairness, Accountability, and Transparency*. 29-31 January 2019, Atlanta, GA, USA. <https://doi.org/10.1145/3287560.3287575>.

Big Brother Watch: *Who's Watching You? The Dominance of Chinese State-Owned CCTV in the UK*. 2022. Available at: https://bigbrotherwatch.org.uk/wp-content/uploads/2022/02/Whos-Watching-You_The-dominance-of-Chinese-state-owned-CCTV-in-the-UK.pdf, accessed: 03.12.2025.

Buolamwini, Joy – Gebru, Timnit: Gender shades: Intersectional accuracy disparities in commercial gender classification. Conference on Fairness, Accountability and Transparency. In: *Proceedings of Machine Learning Research* 81. 2018, 1-15.

Burnay, Matthieu – Eva Pils: Weaponizing Citizenship in China: Domestic Exclusion and Transnational Expansion. In: *State Crime Journal* 9, no. 1 (2020), 4-28. <https://doi.org/10.13169/statecrime.9.1.0004>.

Boshier, Roger: Learning from the Moa: The Challenge of Māori Language Revitalization in Aotearoa/New Zealand." In: Jacob, W., Cheng, S., Porter, M. (eds): *Indigenous Education*. Dordrecht: Springer. 2015. https://doi.org/10.1007/978-94-017-9355-1_11.

Chang, Xinyu: Gender Bias in Hiring: An Analysis of the Impact of Amazon's Recruiting Algorithm. In: *Advances in Economics, Management and Political Sciences*. 2023. Available at: <https://pdfs.semanticscholar.org/4b05/657a3abe9f7f976a75e0fada85476823dca4.pdf>, Accessed: 18.10.2024

Çiçek, Ayşe: «Political Re-education Camps» in Xinjiang as a part of panopticon-like Society: Clashing Discourses of Western and Chinese Media. In: *Wschodnioznaawstwo, Tom 16*. 2022, 269-298 <https://doi.org/10.4467/20827695WSC.22.017.16766>.

Creemers, Rogier: *China's Social Credit System: An Evolving Practice of Control* (May 9, 2018). Available at SSRN: <https://ssrn.com/abstract=3175792> or <http://dx.doi.org/10.2139/ssrn.3175792>

Dabis, Attila: Effective Participation and Territorial Integrity: A Guide to Romanian Doctrinal Claims on Regional Power-sharing. In: *Europäisches Journal für Minderheitenfragen, Volume 15, Issue 1-2*, June 2022, 94-114. <https://doi.org/10.35998/ejm-2022-0005>.

Dabis, Attila – Csáki, Csaba: AI and ethics: Investigating the first policy responses of higher education institutions to the challenge of generative AI. In: *Humanities and Social Sciences Communication* 11, 1006 (2024). <https://doi.org/10.1057/s41599-024-03526-z>.

European Commission: *Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence, adopted by the European Parliament and the Council in December 2023*, available at: <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>, accessed: 16.10.2024.

Ensign, Danielle – Friedler, Sorelle A. – Neville, Scott – Scheidegger, Carlos – Venkatasubramanian, Suresh: Runaway Feedback Loops in Predictive Policing. Proceedings of the 1st Conference on Fairness, Accountability and Transparency. In: *Journal of Machine Learning Research* 81. 2018, 160-171. Available at: <https://proceedings.mlr.press/v81/ensign18a.html>, Accessed: 18.10.2024.

Faraldo Cabana, P. (2023). Technical and Legal Challenges of the Use of Automated Facial Recognition Technologies for Law Enforcement and Forensic Purposes. In: Završnik, Aleš – Simončič, Katja (eds.): *Artificial Intelligence, Social Harms and Human Rights, Critical Criminological Perspectives*. Cham: Palgrave Macmillan. 2023. https://doi.org/10.1007/978-3-031-19149-7_2

Føreland, Line Reichelt – Rauni Äärelä-Vihriälä: Exploring the Use of Minecraft in Sámi Teacher Education. In: *Genealogy* 8: 59. 2024. <https://doi.org/10.3390/genealogy8020059>.

Government of Catalonia: *Catalonia's Artificial Intelligence Strategy*. July 2019. Available at: https://participa.gencat.cat/uploads/decidim/attachment/file/932/Document-Bases-Estrategia-IA-Catalunya-_ENversion.pdf, accessed: 11.10.2024.

Hornberger, Nancy, H. – De Korne, Haley: Minoritized/Indigenous language revitalization. In: Wei, Li – Hua, Zhu – Simpson, James (eds.): *The Routledge Handbook of Applied Linguistics*. London: Routledge. 2023, 349-361. DOI: 10.4324/9781003082637-30.

Hutchinson, Ben – Mitchell, Margaret: 50 years of Test (UN) fairness: Lessons for MachineLearning. In: *Proceedings of FAT* '19: Conference on Fairness, Accountability, and Transparency (FAT* '19)*. New York: ACM. <https://doi.org/10.1145/3287560.3287600>.

Ibrahim, Azeem (ed.): *The Uyghur Genocide: An Examination of China's Breaches of the 1948 Genocide Convention*. Washington: Newlines Institute for Strategy and Policy. 2021. Available at: <https://newlinesinstitute.org/wp-content/uploads/Chinas-Breaches-of-the-GC3.pdf>, accessed: 21.10.2024.

Jai, Kehaulani R.: *Crackdown and Consent: China's War on Terror and the Strategic Creation of a Public Discourse in the U.S.*. Scripps Senior Theses. Paper 741. 2016. Available at: http://scholarship.claremont.edu/scripps_theses/741, accessed: 11.10.2024.

Jones, Peter-Lucas – Mahelona, Keoni – Duncan, Suzanne – Leoni, Gianna: Kia tangata whenua: Artificial intelligence that grows from the land and people. In: *Ethical Space: International Journal of Communication Ethics* (2/3), 2023. <https://doi.org/10.21428/0af3f4c0.9092b177>

Juncà, Lluís: An Artificial Intelligence strategy for Catalonia. In: *European Public Mosaic* 22, 2024. Available at: <https://www.gencat.cat/eapc/epum/N22/pdf/EPuM22juncà.pdf>, accessed: 11.10.2024.

Kamble, R M: Artificial intelligence and human rights. In: *Uniform Law Review, Volume 29, Issue 1, March*. 2024, 77–86, <https://doi.org/10.1093/ulr/una020>.

Kann, Katharina – Mager, Manuel – Ruiz, Ivan Vladimir Meza – Schütze, Hinrich: Fortification of Neural Morphological Segmentation Models for Polysynthetic Minimal-Resource Languages. In: *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, 47–57, New Orleans, Louisiana. Association for Computational Linguistics. <https://doi.org/10.18653/v1/N18-1005>.

Kaleimamoowahinekapu Galla, C.: Digital Realities of Indigenous Language Revitalization: A Look at Hawaiian Language Technology in the Modern World. In: *Language and Literacy*, 20(3). 2018, 100–120. <https://doi.org/10.20360/langandlit29412>.

Koen, Louis – Mufamadi, Kgomo: Artificial Intelligence and Racial Discrimination. In: Temperman, Jeroen – Quintavalla, Alberto (eds.): *Artificial Intelligence and Human Rights* (2023; online edition, Oxford Academic), <https://doi.org/10.1093/law/9780192882486.003.0013>.

Liang, Fang Das, Vishnupriya – Kostyuk, Nadiya – Hussain, Muzammil M.: Constructing a Data-Driven Society: China's Social Credit System as a State Surveillance Infrastructure. In: *Policy & Internet*, 10: 415–453. 2018. <https://doi.org/10.1002/poi3.183>.

Leibold, James: Surveillance in China's Xinjiang Region: Ethnic Sorting, Coercion, and Inducement. In: *Journal of Contemporary China*, 29:121. 2020, 46-60. DOI:10.1080 /10670564.2019.1621529.

Levitsky, Steven – Way, Lucan A: *Competitive Authoritarianism: Hybrid Regimes after the Cold War*. Cambridge: Cambridge University Press, 2010.

Loubere, Nicholas – Brehm, Stefan: The global age of the algorithm: Social credit, Xinjiang, and the financialisation of governance in China. In: *Made in China Journal*. vol. 3, no. 1. 2018, 175-182. Doi.org/10.22459/MIC.03.01.2018.07.

Lynn, Adam: *App Targeting Uyghur Population Censors Content, Lacks Basic Security*. Washington: Open Technology Foundation. 2018.

Macdonald, Alison – McArthur, Jackie – Hart, Naomi – Aboagye, Lorraine: *International Criminal Responsibility for Crimes Against Humanity and Genocide Against the Uyghur Population in the Xinjiang Uyghur Autonomous Region*. Essex Court Chambers, released on 26 January 2021, available at: https://14ee1ae3-14ee-4012-91cf-a6a3b7dc3d8b.usrfiles.com/ugd/14ee1a_3f31c56ca64a461592ffc2690c9bb737.pdf, accessed: 11.08.2024.

Maia, Marcus Antônio Rezende – Nascimento, Márcia – Chang Whan: The māori language nest program: voices of language and culture revitalization in Aotearoa, New Zealand. In: *Ecolinguística: Revista Brasileira De Ecologia E Linguagem (ECO-REBEL)*, 4(1) 2018, 76–95. Recuperado de <https://periodicos.unb.br/index.php/erbel/article/view/9918>.

May, Stephen – Hill, Richard: Language Revitalization in Aotearoa/New Zealand. In: Hinton, Leanne – Huss, Leena – Roche, Gerald (eds): *The Routledge Handbook of Language Revitalization*. New York: Routledge. 2018.

Meighan, Paul J.: Decolonizing the digital landscape: the role of technology in Indigenous language revitalization. IN: *AlterNative: An International Journal of Indigenous Peoples*, 17(3). 2021, 397-405. <https://doi.org/10.1177/11771801211037672>

Moore, Greg: China's Cautious Participation in the UN Human Rights Regime. In: *Human Rights & Human Welfare: Vol. 1: Issue 1, Article 6*. 2001. Available at: <https://digitalcommons.du.edu/hrhw/vol1/iss1/6>.

Moseley, Christopher (ed.): *Atlas of the world's languages in danger*. Paris: UNESCO. 2010. Available at: <https://unesdoc.unesco.org/ark:/48223/pf0000187026>, accessed: 11.10.2024.

Mullen, Lincoln A. – Benoit, K. – Keyes, O. – Selivanov, Dmitriy – Arnold, Jeffrey: Fast, Consistent Tokenization of Natural Language Text. In: *Journal of Open Source Software* 3. 2018, 655. <https://doi.org/10.21105/JOSS.00655>.

Nagy, Noémi: "Humanity's new frontier": Human rights implications of artificial intelligence and new technologies. In: *Hungarian Journal of Legal Studies*, 64(2). 2024, 236-267. <https://doi.org/10.1556/2052.2023.00481>.

OHCHR: *Assessment of human rights concerns in the Xinjiang Uyghur Autonomous Region, People's Republic of China*. Geneva: OHCHR. 31 August 2022. Available at: <https://www.ohchr.org/sites/default/files/documents/countries/2022-08-31/22-08-31-final-assesment.pdf>, accessed: 24.10.2024.

Olteanu, Alexandra – Castillo, Carlos – Diaz, Fernando – Kiciman Emre: Social Data: Biases, Methodological Pitfalls, and Ethical Boundaries. In: *Frontiers in Big Data* Vol. 2:13. 2019. <https://doi.org/10.3389/fdata.2019.00013>, accessed: 20.10.2024.

Outakoski, Hanna – Cocq, Coppélie – Steggo, Peter: Strengthening Indigenous languages in the digital age: social media-supported learning in Sápmi. In: *Media International Australia*, 169(1). 2018, 21-31. <https://doi.org/10.1177/1329878X18803700>.

Polyakova, Alina – Meserole, Chris: *Exporting Digital Authoritarianism: The Russian and Chinese Models*. Brookings Institution. 2019. Available at: https://www.brookings.edu/wp-content/uploads/2019/08/FP_20190827_digital_authoritarianism_polyakova_meserole.pdf, accessed: 03.12.2025.

Qiang, Xiao: The Road to Digital Unfreedom: President Xi's Surveillance State. In: *Journal of Democracy*, vol. 30 no. 1. 2019, 53-67. <https://dx.doi.org/10.1353/jod.2019.0004>.

Rakhima, Ayu Suci – Satyawati, Ni Gusti Ayu Dyah: Gross Violations of Human Rights Veiled within Xinjiang Political Reeducation Camps. In: *Jurnal Kertha Patrika*, Vol. 41, No. 1 April. 2019, 1-16.

Ringen, Stein: *The Perfect Dictatorship: China in the 21st Century*. Hong Kong University Press. (Published online: 19 January 2017, Published in print: 1 July 2016), <https://doi.org/10.5790/hongkong/9789888208937.001.0001>.

Rögnvaldsson, Eiríkur – Rehm, Georg – Uszkoreit, Hans: *The Icelandic Language in the Digital Age / Íslensk Tunga á Stafrænni ÖLd*. Heidelberg: Springer. 2012.

Schwartz, Lane – Tyers, Francis M. – Levin, Lori S. – Kirov, Christo – Littell, Patrick – Chi-kiu, Littell – Prudhommeaux, Emily – Park, Hayley – Steimel, Park K. – Knowles, Rebecca – Micher, J., Strunk, Lonny – Liu, Han – Haley, Coleman Zhang, Katherine J. – Jimmerson, Robbie – Andriyanets, VVasilisa – Muis, Aldrian Obaja – Otani, Naoki – Park, J., – Zhang, Zhisong: *Neural Polysynthetic Language Modelling*. Available online at: ArXiv, abs/2005.05477. 2020.

Secretary-General of the United Nations: Guidance Note on Racial Discrimination and Protection of Minorities. Geneva: UN. 2014. available at: <https://www.ohchr.org/sites/default/files/Documents/Issues/Minorities/GuidanceNoteSG.pdf>, accessed: 11.10.2024.

Slawotsky, Joel: Is China Guilty of Committing Genocide in Xinjiang? In: *Chinese Journal of International Law*. 2021.

Suresh, Harini – Guttag, John: A Framework for Understanding Sources of Harm throughout the Machine Learning Life Cycle. In: *Equity and Access in Algorithms*,

Mechanisms, and Optimization (EAAMO '21), October 5–9, 2021. New York: ACM. <https://doi.org/10.1145/3465416.3483305>.

Tankó, Gyula: *Professional Writing: The Academic Context*. Budapest: ELTE Eötvös Kiadó Kft. 2021.

Tobin, David: Genocidal Processes: Social Death in Xinjiang. In: *Ethnic and Racial Studies*, vol. 45, no. 16. 2022. <https://doi.org/10.1080/01419870.2021.2001556>.

Turdush, Rukiye – Fiskesjö, Magnus: Dossier: Uyghur Women in China's Genocid. In: *Genocide Studies and Prevention: An International Journal*: Vol. 15: Issue 1. 2021, 22–43. DOI: <https://doi.org/10.5038/1911-9933.15.1.1834>.

UNESCO: *Recommendation on the Ethics of Artificial Intelligence*. 2022. Available aet <https://unesdoc.unesco.org/ark:/48223/pf0000381137>. Accessed 23.10.2024.

UN SR in the field of cultural rights – Working Group on discrimination against women and girls – Working Group on Enforced or Involuntary Disappearances: *UN Experts Urge China to End Repression of Uyghur and Cultural Expression of Minorities*. Press release, issued on 1 October 2025, available at: <https://www.ohchr.org/en/press-releases/2025/10/un-experts-urge-china-end-repression-uyghur-and-cultural-expression>, accessed: 03.12.2025.

UN SR on contemporary forms of slavery: *Report nr. A/HRC/51/26 from 2022 on "Contemporary forms of slavery affecting persons belonging to ethnic, religious and linguistic minority communities"*, para. 23. Geneva: OHCHR. 2022. Available at: <https://documents.un.org/doc/undoc/gen/g22/408/97/pdf/g2240897.pdf?OpenElement>, accessed: 21.10.2024.

UN SR on trafficking in persons, especially women and children: *Report nr. A*/HRC/50/33 on "Trafficking in persons in the agriculture sector: human rights due diligence and sustainable development"*, para. 37. Geneva: OHCHR. 2022. Available at: <https://documents.un.org/doc/undoc/gen/g22/327/46/pdf/g2232746.pdf>, accessed: 21.10.2024.

US Treasury: *Treasury Sanctions Chinese Entity and Officials Pursuant to Global Magnitsky Human Rights Accountability Act*. Issued on 9 July 2020, available at: <https://home.treasury.gov/news/press-releases/sm1055>, accessed: 08.11.2024.

Uyghur Human Rights Project (UHRP): *Weaponized Passports: The Crisis of Uyghur Statelessness*. UHRP. April 2020. Available at: https://docs.uhrp.org/pdf/Weaponized_Passports.pdf, accessed: 24.10.2024.

Uyghur Research Institute: *Genocide in East Turkistan, Report from 21. October 2020*. Available at: <https://www.uysi.org/en/?p=774>, accessed: 24.10.2024.

Wang, Cunrui – Zhang, Qingling – Liu, Wanquan – Liu, Yu – Miao, Lixin: Facial feature discovery for ethnicity recognition. In: *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 9, e1278. 2019. <https://doi.org/10.1002/widm.1278>

Warschauer, Mark – Donaghy, Keola, – Kuamoöö, Hale: LEOKI: A POWERFUL VOICE OF HAWAIIAN LANGUAGE REVITALIZATION. In: *Computer Assisted Language Learning*, 10(4). 1997, 349–361. <https://doi.org/10.1080/0958822970100405>.

Wicks, Andrew C. – Budd, Linnea P. – Moorthi, Ryan A. – Botha, Helet – Mead, Jenny: Automated Hiring at Amazon. In: *Darden Case Collection no. UVE-E-0470*. 2021: n. pag.

World Uyghur Congress: *Amendment NC1 to the Procurement Bill – Briefing note from June 2023*, available at: <https://www.uyghurcongress.org/en/wp-content/uploads/2023/06/WUC-Briefing-Amendment-NC1.pdf>, accessed: 10.11.2024.

Xue, Linting – Barua, Aditya – Constant, Noah – Al-Rfou, Rami – Narang, Sharan – Kale, Mihir – Roberts, Adam – Raffel, Colin: "ByT5: Towards a Token-Free Future with Pre-trained Byte-to-Byte Models. In: *Transactions of the Association for Computational Linguistics*, 10, 2021, 291–306. https://doi.org/10.1162/tacl_a_00461.

Zakaria, Fareed: The Rise of Illiberal Democracy. In: *Foreign Affairs*, Vol. 76, No. 6: 22–41. 1997.

Zenz, Adrian: *Sterilizations, IUDs, and Mandatory Birth Control: The CCP's Campaign to Suppress Uyghur Birthrates in Xinjiang*. Washington: Jamestown Foundation. 2020.

Zhang, Shiyue – Frey, Benjamin – Bansal, Mohit: ChrEn: Cherokee-English Machine Translation for Endangered Language Revitalization. In: *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 577–595., Online. Association for Computational Linguistics. 2020. 10.18653/v1/2020.emnlp-main.43

Zhang, Shiyue – Frey, Benjamin – Bansal, Mohit: How can NLP Help Revitalize Endangered Languages? A Case Study and Roadmap for the Cherokee Language. In: *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 1529–1541, Dublin, Ireland. Association for Computational Linguistics. 2022. 10.18653/v1/2022.acl-long.108