



**Privacy International's response  
to the call for contributions on artificial intelligence in education  
and its human rights-based use at the service of the advancement of  
the right to education**

**May 2024**

1. Privacy International (PI)<sup>1</sup> welcomes the opportunity to provide input to the forthcoming report of the UN Special Rapporteur on the right to education to the General Assembly in October 2024, which will consider artificial intelligence in education and its human rights-based use at the service of the advancement of the right to education.<sup>2</sup>
2. Artificial Intelligence (AI) is becoming increasingly common in educational institutions, as part of teaching, in classrooms, and administration. What has not kept pace has been appropriate safeguards, despite an increasing body of evidence around the potential harms of AI systems and the particularly vulnerable situation of children. PI recognises that AI systems can help in promoting human rights, however its use in education technologies (EdTech) raises specific concerns in relation to the right to education and the right to privacy that have not always been given due consideration in the process.
3. PI's submission hereafter highlights in the first part the potential harms for human rights associated with the use of AI within educational institutions and the additional safeguards and precautions that should be taken when implementing AI in EdTech and provides in the second part examples of AI tools and systems used in education process.<sup>3</sup>

---

<sup>1</sup> PI is an international non-governmental organisation, which campaigns against companies and governments who exploit individuals' data and technologies. PI employs specialists in their fields, including technologists and lawyers, to understand the impact of existing and emerging technology upon data exploitation and our right to privacy. <https://privacyinternational.org/>

<sup>2</sup> Special Rapporteur on the right to education, Call for contributions: artificial intelligence in education and its human rights - based use at the service of the advancement of the right to education, <https://www.ohchr.org/en/calls-for-input/2024/call-contributions-artificial-intelligence-education-and-its-human-rights>

<sup>3</sup> The relevant standards include the UN Convention on the Rights of the Child (UN CRC), the UN International Covenant on Civil and Political Rights and the UN Convention (ICCPR) and the UN International Covenant on Economic, Social and Cultural Rights (ICESCR). There are also soft law measures such as the Abidjan Principles on the human rights obligations of States to provide public education and to regulate private involvement in education that lay out human rights guidance for States to provide public education and to regulate private involvement in education.

## Recommendations

4. PI recommends the UN Special Rapporteur for the upcoming report to:
  - Underline the need for a human rights-based approach to all AI systems in the education sector and describe the necessary measures to achieve it including human rights due diligence, including human rights and data protection impact assessments, human rights by design, as well as ensuring the meaningful participation of affected communities in decision-making processes.
  - Reassert that any interference with the right to privacy and the advancement of the right to education due to the use of AI technologies should be subject to the overarching principles of legality, necessity, and proportionality.
  - Encourage states to adopt or review effective data protection legislation and sectoral laws to address the negative human rights implications of AI systems in education – at individual, group and society level.
  - Identify the human rights risks of specific AI applications, due to the technologies employed and/or the context of their use; and describe the circumstances when AI applications should be banned in education because of human rights concerns.
  - Explore the relationship between public-private partnerships in technology in the education sector. Define the scope of responsibility for private actors to ensure a human rights-based approach to their practices and to abide by the UN Guiding Principles on Business and Human Rights.
  - Recommend that companies providing AI systems to education institutions should be required to waive commercial confidentiality and make their technologies fully auditable by any third party. If details of the workings of a particular technology cannot be disclosed for specified and valid grounds of serious commercial harm to the company, an independent oversight body bound by duties of confidentiality should be granted full access to the technology required to carry out an independent audit of the technologies.

## Part 1: The impact of AI in educational institutions on children's rights

5. The use of AI in educational institutions will have direct impact on children's enjoyment of their fundamental rights, who by default, have the right to additional measures of protection as is required by their status as minors.
6. The Special Rapporteur on the Right to Education has previously highlighted that the digitalisation of education brings serious risks to human rights, including the right to education. Some risks are the exact opposite of the potential benefits: heightened exclusion instead of improved access, standardisation instead of personalised teaching, enhanced

stereotypes instead of diversity, reduced autonomy and freedom instead of creativity and participation, and data mining for the benefits of a few at odds with the public interest.<sup>4</sup>

7. The use of EdTech in educational institutions has risen globally. PI has mapped some of these concerns about the lack of adherence to human rights principles in relation among others of India and Brazil being two examples of states with widespread adoption or plans to adopt this technology.<sup>5</sup> We are concerned that the use of AI systems in educational institutions and education technologies can exacerbate existing inequalities further and lead to further violations of the rights to privacy, freedom from discrimination and ultimately limit the access to education for millions of students.

### **1.1 AI in education and the right to privacy**

8. AI systems require the generation, collection, processing, and retention of mass amounts of personal data and therefore directly interfere with the right to privacy.<sup>6</sup> The right to privacy encompasses the physical and psychological integrity of a person, and can, therefore, embrace multiple aspects of the person's physical and social identity. Considering the vast amounts of personal data that AI systems in educational institutions process and use to make inferences about a student, including potentially sensitive data, they can significantly impact on the enjoyment of the right to privacy. Any interference with the right to privacy must be proportionate, necessary and in accordance with the law. Yet we have observed AI being introduced without any impact assessment, appropriate legal frameworks, and safeguards in place.
9. We are particularly concerned by increasing reliance on consent as a legal basis for processing data in educational settings. People shouldn't be asked to exchange their privacy for access to education. It is often hard to fully understand what kinds and how much data devices, networks, and platforms generate, process, or share. Furthermore, how personal data is

---

<sup>4</sup> UN Special Rapporteur on the Right to Education, Report on impact of the digitalization of education on the right to education, UN Doc A/HRC/50/32, 19 April 2022, <https://www.ohchr.org/en/documents/thematic-reports/ahrc5032-impact-digitalization-education-right-education>

<sup>5</sup> See PI's submissions on the use of EdTech for the Universal Periodic Reviews (UPR) of India, <https://privacyinternational.org/advocacy/4981/right-privacy-indian-schools-universal-periodic-review>, and Brazil, <https://privacyinternational.org/advocacy/4982/right-privacy-brazilian-schools-universal-periodic-review>

<sup>6</sup> Article 17 UN ICCPR upholds the right to privacy, providing that no one shall be subjected to arbitrary or unlawful interference with his privacy, family, home or correspondence and that everyone has the right to the protection of the law against such interference or attacks. Article 16 UN CRC provides that the child shall not be subjected to arbitrary or unlawful interference with his or her privacy and that the child has the right to the protection of the law against such interference or attacks. In its General Comment No. 25 (2021), the Committee on the Rights of the Child underlined that children had a right to privacy in the digital space, which was vital for protecting their agency, dignity and safety.

collected and used is often far from transparent, with, in some cases, total opacity and disrespect for the right to privacy.

10. In the UK, for example, a national survey revealed that less than 30% of children were made aware by their educational institutions of why EdTech was being used and how it operated. Moreover, only less than 1 in 10 students reported that they were fine with the companies sharing their data with third parties.<sup>7</sup> Furthermore, 'meaningful' consent is extremely difficult, and potentially impossible, to achieve in an educational setting.
11. Also, in Sweden, the Data Protection Authority (DPA) fined a municipality 20,000 euros for violating the EU's General Data Protection Regulation (GDPR) regarding consent. While the municipality argued they had obtained consent, the DPA noted that the power imbalance between students, their guardians and the school meant the consent could not be deemed freely given.<sup>8</sup>
12. Public and private actors deploying AI systems in educational institutions should take the necessary steps to ensure that the right to privacy and data protection are protected in the process, including by introducing appropriate and effective regulation to ensure the data is processed according to internationally recognised human rights standards, including data protection principles.<sup>9</sup>

## 1.2 AI in education and freedom from discrimination

13. AI systems use identification, profiling, and automated decision-making which can lead to unfair, discriminatory, or biased outcomes. This can occur for several reasons and at many levels in AI systems and is often difficult to detect and mitigate. Often, the quality of the data used to train the system and biases within it are the source of potential discrimination and unfair treatment. People can therefore be misclassified, misidentified, or judged negatively, and such errors or biases may disproportionately affect certain groups of people.<sup>10</sup>

---

<sup>7</sup> Digital Futures Commission, 'What do children think of EdTech or know of its data sharing? Read our survey findings', 24 October, 2022, <https://digitalfuturescommission.org.uk/blog/what-do-children-think-of-edtech-or-know-of-its-data-sharing-read-our-survey-findings/>.

<sup>8</sup> European Data Protection board, 'Facial recognition in school renders Sweden's first GDPR fine', 22 August 2019, [https://www.edpb.europa.eu/news/national-news/2019/facial-recognition-school-renders-swedens-first-gdpr-fine\\_sv](https://www.edpb.europa.eu/news/national-news/2019/facial-recognition-school-renders-swedens-first-gdpr-fine_sv).

<sup>9</sup> See PI's explainer: 101: Data Protection, 12 October 2017, <https://privacyinternational.org/explainer/41/101-data-protection>. See also UN High Commissioner for Human Rights report on the right to privacy in the digital age, UN doc.A/HRC/39/29, paras 29 and 30, <https://www.ohchr.org/en/documents/thematic-reports/ahrc3929-right-privacy-digital-age-report-united-nations-high>.

<sup>10</sup> The right of the child not to be subject to discrimination is provided in Article 2 UNCRC, which provides that States shall take all appropriate measures to ensure that the child is protected against all forms of discrimination and that in all actions concerning children, whether undertaken by public or private social welfare institutions, courts of law, administrative authorities or legislative bodies, the best interests of the child shall be a primary consideration. Moreover, Article 26 of the ICCPR guarantees freedom from discrimination, and Article 10 of the ICESCR highlights the duty to provide special measures

14. When discrimination occurs, this can lead to exclusion, which can ultimately interfere with the child's right to education.<sup>11</sup> This is of significant concern as the quality of education a child receives, their attainment in school, even ultimately their dropping out, can have significant consequences across the course of their life. This exclusion is evident across various uses of EdTech, but particularly so in the use of facial recognition technology (FRT), (explored further below).
15. That educational institutions have implemented AI technologies already, despite the well documented existence of this discriminatory bias suggests that existing procurement and assessment processes are wholly inadequate in evaluating AI systems.

### **1.3 AI technologies and the right to education**

16. EdTech is often perceived as an avenue to guarantee the right to education, however, in many cases it places students in onerous positions to access this right and could be undermining the right itself.
17. General Comment 1 of the Committee on the Rights of the Child made clear that, in practice, education must be "child-centred and empowering" and "must be provided in a way that respects the inherent dignity of the child and enables the child to express his or her views...and to participate in school life."<sup>12</sup> The intrusiveness of the surveillance children are now subjected to is not empowering, does not respect their dignity and actively hinders many children from being able to participate fully.
18. Moreover, the monitoring of students' movements and facial expressions, and their communications both amongst themselves and with their teachers, addressed below, limits their ability to develop to their fullest potential - integral to Article 29 of the UN CRC.<sup>13</sup>

### **1.4 Recommendations**

19. As a result of these concerns, PI recommends the UN Special Rapporteur for the upcoming report to:

---

of protection and assistance for all children and young persons without any discrimination, and calls for children and young persons to be protected from economic and social exploitation.

<sup>11</sup> The right to education is provided in Articles 13 and 14 UN ICESCR and Articles 28 and 29 UN CRC.

<sup>12</sup> UN Committee on the Rights of the Child, "General Comment No. 1 (2001) Article 29 (1): The Aims of Education", CRC/GC/2001/1, 17 April 2001.

<sup>13</sup> UNICEF, 'The right to education: Introducing Articles 28 and 29', <https://www.unicef.org.uk/rights-respecting-schools/the-rsa/the-right-to-education/>

- Underline the need for a human rights-based approach to all AI systems in the education sector and describe the necessary measures to achieve it including human rights due diligence, including human rights and data protection impact assessments, human rights by design, as well as ensuring the meaningful participation of affected communities in decision-making processes.
- Reassert that any interference with the right to privacy and the advancement of the right to education due to the use of AI technologies should be subject to the overarching principles of legality, necessity, and proportionality.
- Encourage states to adopt or review effective data protection legislation and sectoral laws to address the negative human rights implications of AI systems in education – at individual, group and society level.

## **Part 2: AI tools and systems used in education process and related decision making**

20. The impact of AI in education is further exemplified in specific surveillance technologies that have been introduced in education. We provide some examples in the second part.

### **2.1 AI systems equipped with facial recognition technology**

21. In a growing number of countries, facial recognition technology (FRT),<sup>14</sup> which can often be coupled with AI systems, is being increasingly used to mediate children’s access to education. This is despite the persistent evidence of discrimination within facial recognition systems, including systems being deployed by educational institutions.<sup>15</sup>
22. The risks to human rights, in particular the right to privacy, associated with the use of FRT have been well-documented.<sup>16</sup> These concerns are further compounded when additional analytics

---

<sup>14</sup> FRT involves the use of cameras to capture digital images of individuals’ facial features, and the automated processing of these images to identify, authenticate or categorise people. The technology extracts biometric facial data, creates a digital signature of the identified face, stores it and searches records in a database or a watchlist to find a match. FRT may involve the use of cameras, which can capture individuals’ facial images and process them in real time ("live FRT") or at a later point ("Static" or "Retrospective FRT"). The collection of facial images results in the creation of “digital signatures of identified faces”, which are analysed against one or more databases (“Watchlists”), usually containing facial images obtained from other sources to determine if there is a match. See PI and Liberty’s Explainer on Facial Recognition, <https://privacyinternational.org/sites/default/files/2019-02/Explainers-Facial%20Recognition.pdf>; PI’s learn page on Facial Recognition, <https://privacyinternational.org/learn/facial-recognition>

<sup>15</sup> Yoder-Himes DR, Asif A, Kinney K, Brandt TJ, Cecil RE, Himes PR, Cashion C, Hopp RMP and Ross E (2022) Racial, skin tone, and sex disparities in automated proctoring software. *Frontier Education*, 7:881449, <https://www.frontiersin.org/articles/10.3389/educ.2022.881449/full>

<sup>16</sup> Report of the United Nations High Commissioner for Human Rights on the right to privacy in the digital age, 13 September 2021, UN Doc. A/HRC/48/31; UNGA Resolution on the right to privacy in the digital age, 15 December 2022, UN Doc. A/RES/77/211, page 3. See also the Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, 28 May 2019, UN Doc. A/HRC/41/35, para 12, wherein he noted that FRT “seeks to capture and

features increasingly rely on AI systems<sup>17</sup> to carry out facial recognition as noted by the High Commissioner for Human Rights.<sup>18</sup> PI has previously highlighted how the deployment of FRT is happening in a regulatory void and it is not subject to public and democratic scrutiny.<sup>19</sup> Many educational institutions around the world have implemented these technologies without the appropriate oversight, transparency, or review.<sup>20</sup>

23. Among others, there has been persistent evidence of discrimination within FRT systems, including systems being deployed by educational institutions.<sup>21</sup> One researcher's testing of the software 'Proctorio' found that the AI system seemed to be "using a facial detection model that fails to recognize Black faces more than 50 percent of the time."<sup>22</sup> Some data protection authorities have taken steps to prevent the technology from being used in classrooms,<sup>23</sup> and some other authorities - such as New York State - have banned the use of the technology in educational institutions because of the "potentially higher rates of false positives for people of color".<sup>24</sup> This is in line with the growing body of research which suggests that AI systems will mismatch black faces at a higher rate than white faces.<sup>25</sup>
24. Moreover, we consider that the fact that these systems have been continually rolled out, without the well documented risk of discrimination having been addressed, is an indictment of the existing procurement rules and safeguards. Hence, reinforcing the need to call for

---

detect the facial characteristics of a person, potentially profiling individuals based on their ethnicity, race, national origin, gender and other characteristics, which are often the basis for unlawful discrimination".

<sup>17</sup> Note: Even though most FRT is considered to be AI-fuelled, it seems that not all FRT necessarily amounts to AI. The key difference is whether the FRT algorithm was trained using a neural networks approach. Neural networks are a method in artificial intelligence that teaches computers to process data in a way that is inspired by the human brain. Essentially, after being fed thousands of training examples, neural networks help to spot patterns and classify images without human intervention. One key application of neural networks is computer vision, which allows computers to distinguish and recognize images similar to humans. FRT is a form of computer vision. See AWS, 'What is a Neural Network', <https://aws.amazon.com/what-is/neural-network/>

<sup>18</sup> Report of the United Nations High Commissioner for Human Rights on the right to privacy in the digital age, 13 September 2021, UN Doc A/HRC/48/31, para 32.

<sup>19</sup> PI, "UK MPs Asleep at the Wheel as Facial Recognition Technology Spells The End of Privacy in Public", 7 November 2023, <https://privacyinternational.org/long-read/5155/uk-mps-asleep-wheel-facial-recognition-technology-spells-end-privacy-public>

<sup>20</sup> InternetLab, "Surveillance Technologies And Education: mapping facial recognition policies in Brazilian public schools", Diagnosis and Recommendations nº 8, 2023, <https://internetlab.org.br/wp-content/uploads/2023/06/Educacao-na-mira-EN-03.pdf>; Carolina Batista Israel, Rodrigo Firmino, coordenadores; [autores] Carolina Batista Israel ... [et al.]; capa, Manoela M. Jazar - Curitiba (2023) Reconhecimento facial nas escolas públicas do Paran, p 38, [https://jaracalab.org/cms/wp-content/uploads/2023/12/RF\\_PR\\_2023.pdf](https://jaracalab.org/cms/wp-content/uploads/2023/12/RF_PR_2023.pdf)

<sup>21</sup> See further PI's submission to the UN Special Rapporteur on racism for their upcoming report which will examine and analyse the relationship between artificial intelligence (AI) and non-discrimination and racial equality, as well as other international human rights standards, April 2024, <https://privacyinternational.org/advocacy/5295/pi-seeks-inform-report-ai-and-racial-discrimination-un-special-rapporteur-racism>

<sup>22</sup> Todd Feathers, "Proctorio Is Using Racist Algorithms to Detect Faces", Vice, 8 April 2021, <https://www.vice.com/en/article/g5gxg3/proctorio-is-using-racist-algorithms-to-detect-faces>

<sup>23</sup> Sofia Edvardsen, How to interpret Sweden's first GDPR fine on facial recognition in school, IAPP, 27 August 2019, <https://iapp.org/news/a/how-to-interpret-swedens-first-gdpr-fine-on-facial-recognition-in-school>

<sup>24</sup> Commissioner of Education of the State of New York on behalf of the State Education Department, Order, <https://www.nysed.gov/sites/default/files/programs/data-privacy-security/biometric-determination-9-27-23.pdf>

<sup>25</sup> Tom Simonite, 'The Best Algorithms Struggle to Recognize Black Faces Equally' 22 July 2019, <https://www.wired.com/story/best-algorithms-struggle-recognize-black-faces-equally/>

human rights due diligence prior to the introduction of any AI related system in education. Among the specific concerns around racial discrimination resulting from the use of FRT are: non-representative training data with data sets used to train AI models and algorithms do not necessarily represent the communities on which the final system will be used,<sup>26</sup> and there are reported concerns of lower accuracy of facial recognition technologies with certain groups with skin colour being a key factor in the bias and lack of accuracy and profiling on the basis of race, ethnicity, national origin.<sup>27</sup>

## 2.2 Emotion Recognition

25. Further systems, intertwined with the technology found within facial recognition, intended to monitor children's emotions are also being deployed in educational institutions. Educational institutions across Hong Kong are using emotion recognition software to monitor facial expressions of children to determine their mood and level of motivation, gauge their progress, and predict their scores. This technology is also rife with bias, particularly so as different cultures use different facial expressions to emote.<sup>28</sup>
26. These systems are fundamentally unsound, and have been found to interpret the facial expressions of white and black people differently - attributing negative feelings, such as contempt and anger, more frequently to black people.<sup>29</sup>
27. That this data being recorded and used to assess children's engagement in lessons and their emotional state is deeply disturbing and dystopian. It amounts to a significant interference with children's right to privacy, and to their right to develop guaranteed by the UN CRC and is quite simply inappropriate for use on children in a classroom.

## 2.3 Social media and communications surveillance

28. AI is being integrated into students' day to day interactions, both inside and outside of educational institutions. Surveillance software, such as those provided by Gaggle, Navigate360

---

<sup>26</sup> Joy Buolamwini, Unmasking the bias in facial recognition algorithms, 13 December 2023, Excerpted from the book "Unmasking AI: My Mission to Protect What Is Human in a World of Machines," by Joy Buolamwini (2023), Published by Random House, an imprint and division of Penguin Random House LLC, <https://mitsloan.mit.edu/ideas-made-to-matter/unmasking-bias-facial-recognition-algorithms>

<sup>27</sup> Larry Hardesty, Study finds gender and skin-type bias in commercial artificial-intelligence systems, MIT News, 11 February 2018, <https://www.nist.gov/speech-testimony/facial-recognition-technology-frt-0>

<sup>28</sup> PI, Hong Kong schools adopt facial recognition and security cameras, 23 May 2023, <https://privacyinternational.org/examples/5216/hong-kong-schools-adopt-facial-recognition-and-security-cameras>

<sup>29</sup> Lauren Rhue, Racial Influence on Automated Perceptions of Emotions, SSRN, 9 November 2018, <https://ssrn.com/abstract=3281765>



(previously Social Sentinel), and Bark, are increasingly using AI to flag what they claim to be harmful, inappropriate, or concerning messages in student's messages, social media, or browsing history. From algorithms blocking or inappropriately flagging LGBTQ+ content<sup>30</sup> and allegedly outing students to their parents,<sup>31</sup> to "forestalling" protests<sup>32</sup> - the software being provided by these companies is being used in ways that undermine student's human rights, including their freedom of expression, privacy and non-discrimination. Natural language processing - of the kind broadly used by these systems - has been shown to have issues categorising AAVE (African American Vernacular English)<sup>33</sup> and that AAVE sentences may be more likely to be flagged as 'rude'.<sup>34</sup>

29. These AI systems also threaten students access to quality education. Research has found that this kind of extensive surveillance "discourages cooperation among students and rewards conformity over creative or critical thinking and is also particularly bad for learning"<sup>35</sup> A report into these companies by US Senators Elizabeth Warren and Ed Markey found that "these products may be exacerbating the school-to-prison pipeline by increasing law enforcement interactions with students", that the companies have not "taken any steps to determine whether student activity monitoring software disproportionately targets students from marginalized groups" despite evidence that they do, that students and their guardians aren't being appropriately informed of the "use - and potential misuse" of their data.<sup>36</sup>
30. Even more concerning, a Centre for Democracy and Technology report found that "Forty-four percent of teachers report that one or more students have been contacted by law enforcement because of behaviours flagged by the student activity monitoring system".<sup>37</sup> Moreover, significantly more LGBTQ+ students had been, or knew another LGBTQ+ student who had been, contacted "by a police officer or other adult due to concerns about them committing a

---

<sup>30</sup> Todd Feathers, 'Schools Use Software That Blocks LGBTQ+ Content, But Not White Supremacists', 28 April 2021, <https://www.vice.com/en/article/v7em39/schools-use-software-that-blocks-lgbtq-content-but-not-white-supremacists>

<sup>31</sup> James Factora, 'Surveillance Programs Are Reportedly Targeting, Outing LGBTQ+ Students', 19 October 2021, <https://www.them.us/story/surveillance-programs-reportedly-targeting-outing-lgbtq-students>

<sup>32</sup> Ari Sen & Derêka K. Bennett, 'Tracked: How colleges use AI to monitor student protests' 20 September 2022 <https://interactives.dallasnews.com/2022/social-sentinel/>.

<sup>33</sup> Su Lin Blodgett, Brendan O'Connor, 'Racial Disparity in Natural Language Processing: A Case Study of Social Media African-American English', 30 June 2017, <https://arxiv.org/abs/1707.00061>

<sup>34</sup> Anna Woorim Chung, 'How Automated Tools Discriminate Against Black Language', 24 January 2019 <https://civic.mit.edu/2019/01/24/how-automated-tools-discriminate-against-black-language/>

<sup>35</sup> Valerie Steeves, Priscilla Regan and Leslie Regan Shade, 'Digital Surveillance in the Networked Classroom', 13 January 2023, <http://www.equalityproject.ca/wp-content/uploads/2017/05/7-Digital-Surveillance-in-the-Networked-Classroom.pdf>

<sup>36</sup> Senators Elizabeth Warren and Ed Markey, 'Constant Surveillance: Implications of Around-the-Clock Online Student Activity Monitoring', March 2022, <https://www.warren.senate.gov/download/356670-student-surveillance>

<sup>37</sup> Elizabeth Laird, Hugh Grant-Chapman, Cody Venzke, Hannah Quay-de la Vallee, 'Report – Hidden Harms: The Misleading Promise of Monitoring Students Online', 3 August 2022, p 20, <https://cdt.org/wpcontent/uploads/2022/08/Hidden-Harms-The-Misleading-Promise-of-Monitoring-Students-Online-Research-Report-Final-Accessible.pdf>

crime" (31% to 19%).<sup>38</sup> Since both reports Gaggle have decided to drop LGBTQ keywords,<sup>39</sup> but one small action by one company does not alleviate the significant concerns about the potential systemic discrimination created by this kind of software.

31. These abuses are not being mitigated by human review either. Flags surfaced by Gaggle's algorithm, for example, are reviewed by human content moderators who claim they have little training, little time to make decisions, and little support despite being exposed to child pornography and suicide notes, amongst other more mundane conversations.<sup>40</sup>

## 2.4 Scoring systems

32. Educational institutions in Wisconsin, USA, use a dropout early warning system built by the state to identify students at risk of not graduating. The system's machine learning algorithms make their assessments based on test scores, disciplinary records, lunch price status, and race. In a study of millions of predictions over a decade, it has been found that the system may be wrongly and negatively influencing teachers' impressions of students, especially those of colour, that the system has not improved graduation risks for students dubbed "high risk", and that false alarms are 42 percentage points higher for black students and 18 points higher for Hispanic students, compared to white students. At least eight other US states are building similar systems for future use.<sup>41</sup>
33. There is no properly applicable legal basis - nor should there ever be - for the use of automated systems, and more specifically of AI-based systems that generate scores or predictions, for the taking of any decisions about children that have legal or otherwise "significant" effects on those children, e.g., to decide on admission to a specific school, or to place a child in a particular stream.

## 2.5 Recommendations

34. As a result of these concerns, PI recommends the UN Special Rapporteur for the upcoming report to:

---

<sup>38</sup> *ibid.*, p 21.

<sup>39</sup> Mark Keierleber, 'Gaggle Drops LGBTQ Keywords from Student Surveillance Tool Following Bias Concerns', 27 January 2023, <https://www.the74million.org/article/gaggle-drops-lgbtq-keywords-from-student-surveillance-tool-following-bias-concerns/>

<sup>40</sup> Mark Keierleber, 'Meet the Gatekeepers of Students' Private Lives', 2 May 2022, <https://www.the74million.org/article/meet-the-gatekeepers-of-students-private-lives/>

<sup>41</sup> Todd Feathers, 'False Alarm: How Wisconsin Uses Race and Income to Label Students "High Risk"', *The Markup*, 27 April 2023, <https://themarkup.org/machine-learning/2023/04/27/false-alarm-how-wisconsin-uses-race-and-income-to-label-students-high-risk>

- Identify the human rights risks of specific AI applications, due to the technologies employed and/or the context of their use; and describe the circumstances when AI applications should be banned in education because of human rights concerns.
- Explore the relationship between public-private partnerships in technology in the education sector. Define the scope of responsibility for private actors to ensure a human rights-based approach to their practices and to abide by the UN Guiding Principles on Business and Human Rights.
- Recommend that companies providing AI systems to education institutions should be required to waive commercial confidentiality and make their technologies fully auditable by any third party. If details of the workings of a particular technology cannot be disclosed for specified and valid grounds of serious commercial harm to the company, an independent oversight body bound by duties of confidentiality should be granted full access to the technology required to carry out an independent audit of the technologies.