TAGe

Talking Across Generations on the Ethics of Artificial Intelligence

unesco

Mahatma Gandhi Institute of Education for Peace and Sustainable Development

URGENT

# POLICY BRIEF

## Talking Across Generations on the Ethics of Artificial Intelligence

Ref: UNESCO/MGIEP/2021/PB 05

| S. NO. | MEMBER STATE | DATE | ADOPTED |
|---|---|---|---|
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |

# 1. Executive Summary

Around the world, 69 percent of people between 15 and 24 years of age are online,[1] while only half of people across all age groups use the Internet. Young people today, perhaps the first wholly digital generation, are a driving stake holder in the potential of artificial intelligence (AI) in unlocking educational, entrepreneurial and broader development opportunities and improving basic services for all.

Young people are also the constituency facing the challenges and risks of AI development in a regulatory and ethical vacuum: neuro-adaptation; privacy and data ownership issues in the context of AI-driven applications and devices that gather personal data; algorithmic bias in critical services such as law enforcement and credit lending; legal rights and redress against AI-driven decision making; and algorithmic manipulation in cyberspace that can lead to filter bubbles, polarization, and the shaping of preferences and biases at an impressionable age.

Young people should therefore be key participants, with a prominent seat at the table, when it comes to policy making and planning on the ethical design, adoption and deployment of AI, in public services, development programmes and the commercial applications that young people use every day.

In 2021, two key efforts to shape

**Around the world**

**69%**

**of people between 15 and 24 years of age are online while only half of people across all age groups use the Internet.**

the guiding principles for ethical and responsible AI will make substantial headway: the multilateral exercise to adopt the UNESCO Recommendations for Ethical AI (to be submitted to the UNESCO General Conference in November 2021), and India's publication of an approach document on Responsible AI for All, drafted by NITI Aayog. Inspired by these efforts, UNESCO Mahatma Gandhi Institute of Education for Peace and Sustainable Development (MGIEP), in partnership with the United Nations in India, UNICEF and United Nations Volunteers (UNV), convened a two-week-long digital consultation with nearly 3000 young people, undertaken through UNICEF's U-Report, social media platforms, and email, culminating in a Talking Across Generations Live Dialogue with senior experts.

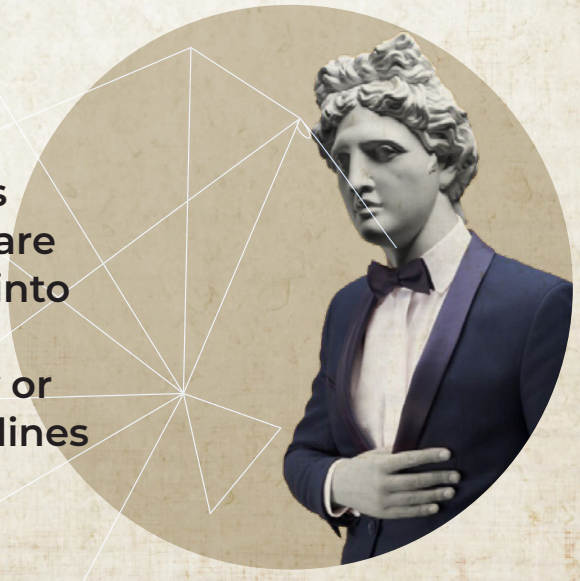
The aim was to convene a youth-led agenda for ethical AI, articulated through this youth-led policy brief on ethical and responsible AI. It outlines the key risks identified by young people, concerning algorithmic bias, privacy and data governance, and explores the policy solutions proposed by young people.

The virtual consultation raised both critical and urgent concerns for young people when it comes to their AI future, as well as surprising results: more than a quarter (26%) of the youth respondents indicated that AI eases their day-to-day tasks such as ordering food, using ride-share services, and sending emails and messages. A sizeable group of participants reported that they see cyber-crime (31%) and the loopholes in technological systems (30.8%) as key risks to their privacy and security. More than one third of respondents (35%) also indicated that youth opinions and feedback are seldom taken into account when framing policy or drafting guidelines for ethical AI.

This brief proposes that youth, as key constituents in shaping the contours of the digital future, must have a seat at the table when it comes to dialogue on AI governance. Successful regulation and a truly multidimensional approach to building a healthy relationship with frontier tech, will require the buy-in of a critical mass of stakeholders – the 600 million young people in India who constitute the largest generation of young people in history.

**35%**

**respondents indicated that youth opinions and feedback are seldom taken into account when framing policy or drafting guidelines for ethical AI.**



# 2. Context: The AI Future is Now

**Today, on average, adults spend**

**6 hr**

**and**

**43 mins**

**online**

In less than one generational shift, we are in the age of the 'digital native'[2] – young people who were born to the ubiquity of computers and the internet. Even before the unique circumstances of the COVID-19 pandemic, young people were drawing on AI-driven technology to access education (tailored content, intelligent tutoring,[3] recreation – smart toys, virtual reality gaming) and socialization (social media that uses algorithm-driven targeting). Today, on average, adults spend 6 hours and 43 minutes online, of which mobile devices account for half,[4] while children spend 2 hours a day online.[5] These applications have access to user data that is used to optimize services.

Young people are not just recipients of AI and digital technologies, but they are also their shapers and co-creators as innovators, producers of digital content, and network builders.

In 1949, there were no 'intelligent' machines; in 1989, a neural network took three days to successfully recognize handwritten ZIP codes;[6] in 2004, no autonomous vehicle finished the test course set by the Defense Advanced Research Projects Agency (DARPA); and in 2021, machine learning is part of our daily routine, industry and recreation.

By 2030, AI[7] could add $13 trillion to the world's economic output and augment global GDP by 1.2 percent every year. In healthcare alone, the AI-driven global healthcare market is worth $6.6 billion in 2021, a 40 percent growth over 2014. AI and machine learning have enabled predictions of crop yield value and enhanced yield through sowing apps,[8] upgraded business productivity,[9] giving a leg up to small businesses in competitive environments, and facilitated product innovation.

AI, and big data and digital tools using AI systems, can improve life outcomes, but their misuse and unintended side-effects raise significant concerns for securing the fundamental principles of equity, rights, and individual and social well-being through regulatory, policy and multi-stakeholder approaches.

The deployment of AI technology without clear regulation and rules on privacy, legal rights and technological guardrails, can throw up a host of challenges. Algorithmic and statistical bias occurs where machine learning based on

an already biased data set, may exacerbate inequality, discrimination or violence. In the more analogue age, we saw these biases in, for example, cameras that were better at capturing lighter skin tones in faces, and crash-test dummies whose design was based on a tall male and endangered the lives of female drivers.[10]

When machine learning is founded on historic practice or data selected in a biased way by the human designer, it may incorporate the same discriminatory biases that were used before it, or reinforce such biases. A deep learning model[11] may also make biased decisions through a lack of intersectional factors or even empathy, for instance a banking algorithm that only dispenses loans or creditworthiness based on financial factors, favouring those who are already secure and wealthy.

There are also ethical questions regarding the future of work and job displacement. In less than ten years, 70 per cent of the world's businesses will have AI in one or more of their functions, and 85 per cent will be using big data analytics.[12]

While AI advances in healthcare can help us screen diseases faster and more efficiently, especially at the last mile, the issue of data privacy in the healthcare sector has been raised as a key concern. AI automation has improved access to credit,[13] but algorithms can lack empathy and intersectional insight and exclude people by assigning creditworthiness to those who are already privileged. And as AI is rolled out for facial recognition, data processing, and criminal justice in the law enforcement sector, the question of legal rights in the face of decisions made by machines and specific-use cases like biased facial recognition, raise important concerns.

Youth constitutes an especially vulnerable category of users and recipients of AI-powered services – in education, healthcare, development programmes, and in their recreational and social use of online and digital tools.

**By 2030, AI could add $13 trillion to the world's economic output and augment global GDP by 1.2 percent every year**

# SENIOR POLICY MAKERS

**P. Anandan**
Former CEO of Wadhwani Institute of Artificial Intelligence

**Renata Lok-Dessallien**
(Former) United Nations Resident Coordinator, India

**Sasha Rubel**
Programme Specialist, Digital Innovation and Transformation Section of the Communication and Information Sector at UNESCO

**Anantha Duraiappah**
Director, UNESCO MGIEP

# YOUTH SPEAKERS

**Dr Shitanshu Mishra**

**Kashish Agarwal**

**Krupalini N. Swamy**

**Shivam Singh**

**Busthan Pokkassery**

**Sruthikeerthi Nandita**

# 3. Approach: Inspirations and Key Questions to Gather Youth Voices

The joint effort by UNESCO MGIEP, UN India, UNICEF and UNV to convene young people's voices, draws on UNESCO's global effort to coordinate a multilateral initiative on the adoption of a Recommendation on Ethical AI, and India's initiative to establish a framework for responsible AI for all. This effort has aimed to ground these principles in the lived experiences of India's youth and their everyday interaction with technology – both as users and creators of digital goods.

## Relevant UN and Policy Milestones on Ethical AI

- UNESCO is leading the multilateral process to develop a comprehensive global standard-setting instrument for ethical AI to protect and promote human rights and human dignity. Comprising a multidisciplinary team of 24 AI specialists, the Ad Hoc Expert Group has undertaken a broad multi-stakeholder consultation to produce a draft Recommendation on the Ethics of Artificial Intelligence.
- The Recommendation outlines four values:
    1. Respect for, promotion and protection of human dignity, human rights and fundamental freedoms;
    2. Environment and ecosystem flourishing;
    3. Diversity and inclusiveness; and
    4. Living in harmony and peace,

  as well as 10 principles for ethical AI: **Outcome Document: First Draft of the Recommendation on the Ethics of Artificial Intelligence (https://unesdoc.unesco.org/ark:/48223/pf0000373434)**
- The **United Nations Secretary-General's Roadmap for Digital Cooperation (https://www.un.org/en/content/digital-cooperation-roadmap/)** builds on the report of the High-level Panel on Digital Cooperation and identifies a set of eight Key Actions, including ensuring the protection of human rights in the digital era.
- The 2020 **Policy Guidance on AI for Children (https://www.unicef.org/globalinsight/media/1171/file/UNICEF-Global-Insight-policy-guidance-AI-children-draft-1.0-2020.pdf)** draws on the Convention on the Rights of the Child, to present the foundations for child-centred AI: AI policies and systems should aim to protect children, provide equitably for their needs and rights, and empower them to participate in an AI world by contributing to the development and use of AI.

India is a founding member of the Global Partnership on Artificial Intelligence, hosted by the OECD, a coalition of 15 countries, which **'aims to bridge the gap between theory and practice on AI by supporting cutting-edge research and applied activities on AI-related priorities.'**

- NITI Aayog has developed **an approach document for responsible AI (http://www.niti.gov.in/sites/default/files/2021-02/Responsible-AI-22022021.pdf)** for all, including legal and regulatory approaches for managing AI systems, technology-based approaches, and principles for the responsible management of AI systems.
- In addition, the European Union, the OECD and other multilateral organizations and partnerships have adopted ethical AI principles and frameworks. India is a founding member of the **Global Partnership on Artificial Intelligence (https://gpai.ai/)**, hosted by the OECD, a coalition of 15 countries, which 'aims to bridge the gap between theory and practice on AI by supporting cutting-edge research and applied activities on AI-related priorities.'

Given the breadth of risks and opportunities of AI for youth, two main questions (and seven sub-questions beneath them) were posed to provoke a substantive and policy-facing discussion:

## Q1. What do young people see as the key opportunities for and risks of a data- and algorithm-driven cyberspace for their well-being and privacy?

Q1a. How do you think artificial intelligence (AI) makes your everyday life better?

Q1b. What do you see as the key risks of a data- and algorithm-driven cyberspace to your privacy and security?

Q1c. In what ways do you think that the way data is used, owned and shared currently could impact individuals and society?

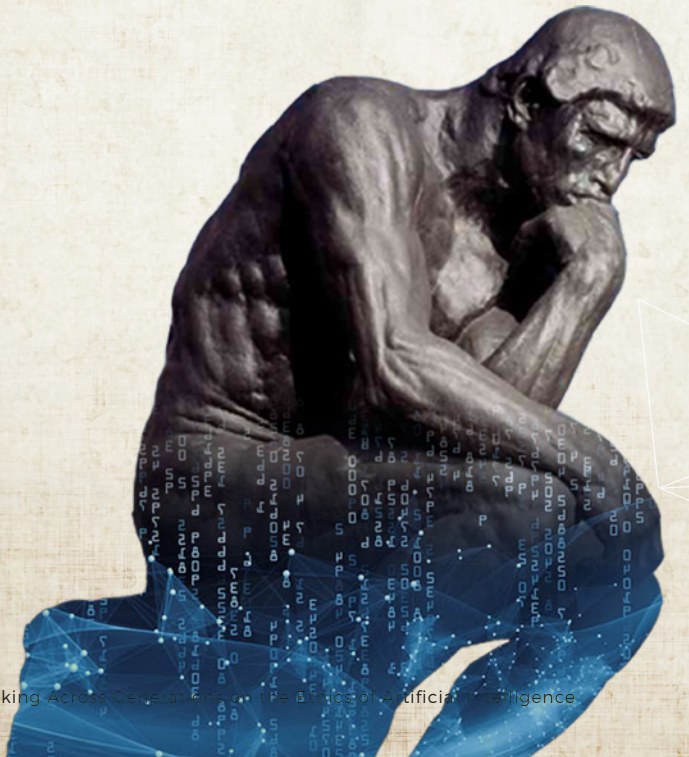
Q1d. What suggestions to address your concerns?

## Q2. What are youth priorities for the accelerated, safe, and transparent deployment of AI-powered tools for sustainable development across sectors?

Q2a. How do you think AI-powered tools can be deployed safely?

Q2b. Which industry sectors do you think have the biggest potential for AI?

Q2c. How can youth voices be included in the development of India's ethical guidelines for AI for social good?



## Process

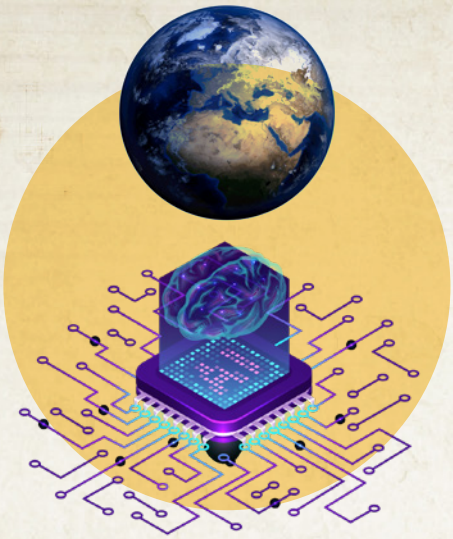
In March 2021, a pan-India online discussion over a period of two weeks, was organized as a TAG$^e$ (Talking Across Generations on Education) Dialogue on the Ethics of AI. TAG$^e$ is a youth-driven intergenerational dialogue series launched by UNESCO MGIEP. Following the online discussion facilitated by three moderators on UNICEF's U-Report, MGIEP's social media platforms and a virtual survey, 3,023 responses were gathered in total. These included 77 responses from Facebook and Instagram discussions, 242 from the virtual survey and 2,704 from U-Report. Six youth champions were selected after a careful evaluation of their contribution to the online discussion. These youth champions engaged in a live, online dialogue on 1 April 2021 with seniors – Renata Dessallien (UN resident coordinator in India), P. Anandan (CEO, Wadhwani Institute for Artificial Intelligence), Sasha Rubel (programme specialist in the Digital Innovation and Transformation Section of the Communication and Information Sector, UNESCO) and Anantha Duraiappah (Director, UNESCO MGIEP) — wherein the youth took centre stage to voice their opinions on a non-hierarchical platform. The live dialogue amassed over 5,000 views on Facebook.

Responses from all sources were collated for analysis. After a thorough cleaning (removal of duplicates, blank entries and gibberish responses), a total of 827 responses were left to be analysed. These responses were coded by a team of four people – three youth speakers and an in-house data analyst. A structured qualitative approach was used to capture themes that appeared in the collected data. The team of coders was first oriented by an expert: here, the process of thematic coding was explained to the coders, who also went through pilot coding of a few responses to ensure that inter-coder reliability was established. After the team felt that satisfactory inter-coder reliability had been achieved, the process of manual coding was performed on a subset of 200 responses from the data pool. For these, answers to the seven key questions posed (see above) were placed into 600 open codes, and then further synthesized into higher-order focus codes. Following this, a stacked machine-learning (ML) model was used to perform a prediction of focus codes on the data. After a verification of model accuracy on the already coded data, the ML model was used to perform a prediction of codes on the rest of the data. These codes were then checked manually and corrected for discrepancies. Finally, the frequencies of codes were set out in a graph.

**3,023 Responses**

**were gathered in total. These included 77 responses from Facebook and Instagram discussions, 242 from the virtual survey and 2,704 from U-Report.**

## Limitations



- TAC$^{e}$ usually culminates in a live, physical dialogue, which stimulates dynamic interactions among youth panellists and senior policy makers. However, due to the ongoing COVID-19 pandemic, it was not possible to organize a face-to-face dialogue with the seniors.

- Although this was a pan-Indian online dialogue, the discussion was held in English, thereby excluding youth who are not as comfortable in the language. The challenge is to reach the most marginalized and silenced youth in this type of exercise, which is designed to reflect youth voices on matters that affect them significantly.

- Due to time constraints, the duration of the online dialogue was shorter than that for a usual TAC$^{e}$, which might have compromised opportunities to deepen discussion on certain themes. Additionally, the drafting team had to work within the limitations of the virtual space and tight deadlines.

- And finally, the qualitative research process does not provide any statistical representation. However, some quantification of responses is necessary and was done by visualising the focus code frequencies.

## Key Findings

### Q1a. How do you think artificial intelligence (AI) makes your everyday life better?

The coding team identified 100+ open codes for Question 1a, which were later synthesized into nine major focus codes. The frequency distribution of the focus codes is illustrated in Figure 1. A third of respondents (31.3%) said that AI helps them solve complex problems that might be dangerous or require smart decision making/automation. One of the respondents said:

> **There are various dangerous jobs that humans have to do like working in mines or in construction, or working with dangerous chemicals and explosives, or under difficult weather conditions. So it'll be better for us to leave such risky jobs for AI-based robots."**

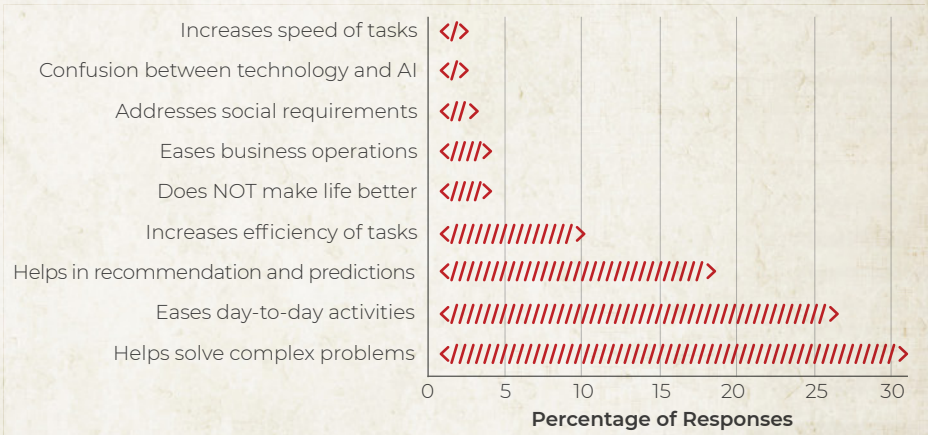



**Figure 1. Frequency distribution of focus codes for Question 1a**

While about 5.6% of respondents indicated that they do not consider AI to be beneficial for everyday life, 25.8% indicated that AI eases their day-to-day tasks such as ordering food, using ride-share services, sending emails and messages, and so on.

> **AI helps man [sic] to do various tasks on a daily basis. It reduces the workload, makes one more comfortable and [helps them] use their energy for more creative tasks."**

A considerable percentage of respondents indicated that AI helps them with recommendations and predictions, such as film and music suggestions, or travel directions, as exemplified by the quote below:

> **Personalized search results and Google Maps algos are a god-send for a frequently travelling person like me."**

Other focus codes that emerged from the analysis included the point that AI helps increase the speed and efficiency of tasks, and eases business operations. Some participants also said that AI helps address their social needs. For example, one of the respondents said that they 'talk to chatbots to relieve stress' and another said that AI helps them 'connect with friends' through social media.

A small fraction of respondents (1.9%) confused AI with technology and responded to the posed question by noting the benefits associated with technology in general. Similarly, only 1.9% of respondents pointed out that AI helps increase the speed of everyday tasks – a theme that the coding team anticipated would have a higher share in the responses. These unexpected findings warrant the general need to build AI literacy in youth in India.

### Q1b. What do you see as the key risks of a data- and algorithm-driven cyberspace to your privacy and security?

Over 100 open codes were synthesized into six focus codes for Question 1b. The majority of participants reported that they see cyber-crimes (50.5%) and the loopholes in technological systems (42.0%) as the key risks to their privacy and security (Figure 2). Cyber-crimes such as 'malware and virus attacks, sabotage of data or networks, and phishing/social engineering', which may be accelerated due to the use of AI, can pose a serious threat to the privacy and security of individuals. Additionally, participants indicated that the inherent loopholes in AI algorithms, a lack of encryption and hashing, limited configuration security and patch management may also pose threats to their security.

**50%**

**of participants reported that they see cyber-crimes as the key risks to their privacy and security**

**The key risk associated is the way our data and information can be misused. Today, everything is under the public purview and can easily be hacked by hackers. Even top-notch IT firms which boast of their security solutions have fallen prey to cyber hacking. Thus, this is a huge concern for the privacy and security of people."**

**I see limited configuration security – companies implement systems that lack encryption, and little to no attention is given to patch management and weakness in code security – I see these as the key risks.**

Respondents indicated that the theft of data and the lack of substantial policy to ensure the ethical use of AI, can also accelerate cyber-crime.

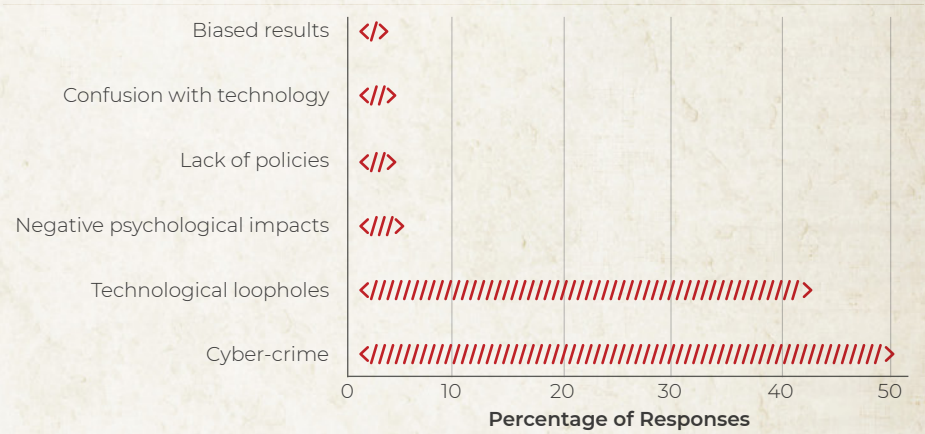



**Figure 2. Frequency distribution of focus codes for Question 1b**

Respondents indicated that the theft of data and the lack of substantial policy (1.9%) to ensure the ethical use of AI, can also accelerate cyber-crime. Like the previous question, a small section of respondents (1.9%) confused AI with technology and answered this question with their views on technology in general. Another small section of respondents indicated that a key risk of algorithm-driven cyberspace might be the negative psychological impact on users.

**Collecting data in the name of providing more efficient and user-friendly results has made me highly sceptical of data and algorithm intensive cyberspaces. It robs me of mental peace as I am always occupied or fearful of the risks posed by such developments in cyberspace."**

Contrary to expectations, the risk of biased results was pointed out by very few respondents (0.5%). However, it is a well-known fact that bias can infiltrate AI algorithms

in various forms and create serious risks for individuals and for society overall. Having diverse teams creating algorithms is therefore crucial to keeping bias at bay. Ensuring diversity, not just in terms of gender but also in socioeconomic status and views, can help tackle anything from xenophobia through homophobia to casteism. It is also important that historical biases are not perpetuated while training AI models. The use of such external mechanisms, as well as hard-wiring bias constraints into the algorithms themselves, can help ensure that AI systems are deployed safely.

### Q1c. In what ways do you think that the way data is used, owned and shared currently could impact individuals and society?

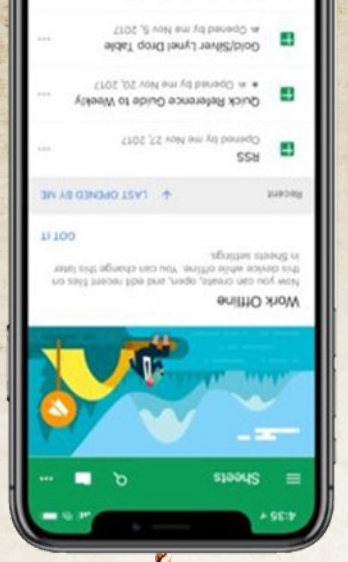
More than 200 open codes were synthesized into seven focus codes for Question 1c (see Figure 3). Half of respondents (50.1%) felt that the way data is used and owned currently puts their privacy under threat. One of the respondents said:

> **You need to be careful with how much personal information you reveal online. Sharing your address, phone number, birthday and other personal information can mean you are at a greater risk of identity theft, stalking and harassment."**

Another important point that emerged from the responses was about how an unethical usage and sharing of data can polarize public sentiment and opinion, and spur the spread of fake news (28.2%):


> **AI used in social media is transforming a person's timeline into a false social bubble. Misinformed people are further pushed into the depths of misinformation and disinformation. This inflicts schism in a society and we arrive at a situation where there are only extremists on either side."**

> **The way data is used and shared can have social implications as it can lead to cutting down marginalized communities such as women, and the LGBTQ. It can also lead to more conflicts.'**



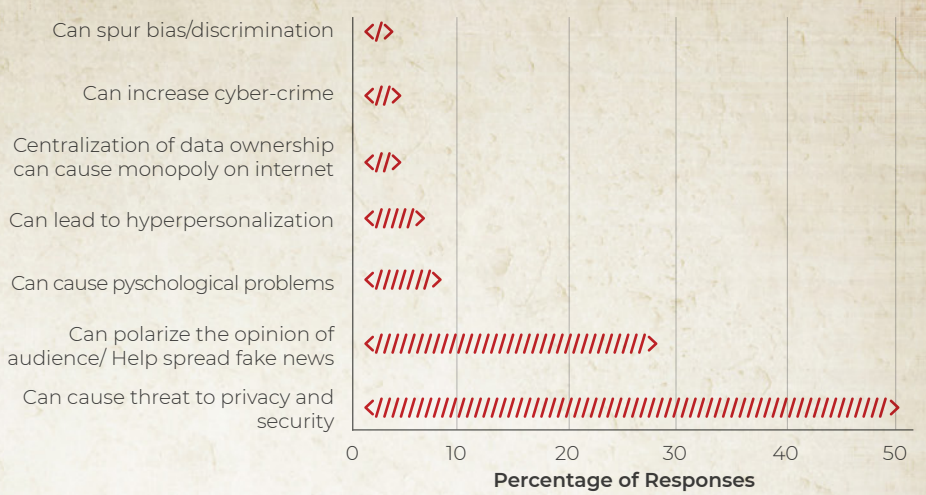



**Figure 3. Frequency distribution of focus codes for Question 1c**

Other focus codes revealed that the way data is handled currently can lead to psychological problems (9.6%), hyper-personalization due to recommendations and predictions (6.3%) (which can be dangerous rather than helpful), a monopoly on the internet (2.1%), and spur an increase in cyber-crime (2.1%) and discriminatory practice (1.5%).

### Q1d. What are your suggestions to address your concerns?

For Question 1d, the coding team identified 150+ open codes, which were merged to create a set of eight focus codes that are illustrated in Figure 4. The majority of respondents (49.4%) seemed to agree that building awareness around issues related to unethical AI (and AI in general) is crucial to addressing current concerns about data sharing, collection and usage. Particular codes on policy making highlighted in responses included the development of policies on the ethical handling of data (14.5%), data privacy (5.9%) and guidelines for developers (3.1%). Respondents said:

**49.4%**

**of respondents seemed to agree that building awareness around issues related to unethical AI (and AI in general) is crucial to addressing current concerns about data sharing, collection and usage**

**‘The data collected for specific purposes should be kept private and not shared. There should be a sense of social responsibility in everything that a person does.”**

**Protect the data itself, not just the perimeter. Pay attention to insider threats. Encrypt all devices. Establish strong passwords and update your programs regularly.”**

Create an independent, transparent body that would evaluate and assess the ethicality of AI systems, and ensure better representation in policy teams.

The respondents also indicated that improving user practices (11.2%), such as reading data privacy statements carefully, can help to address some of the concerns around data ownership and sharing. Although at a lesser frequency, two important suggestions emerged from the analysis – the creation of an independent, transparent body that would evaluate and assess the ethicality of AI systems (1.9%), and ensuring better representation in policy teams (2.4%). Representation matters because we all bring different perspectives to the table, and those perspectives are really informed by our experience and our identity – for example, our education, gender identity, sexual orientation or race.

**Have a good representation of people from different social strata in the policy-making teams.”**

**Formulate an independent body to look into data policy-related matters. Such a body should not just be at the level of nations but should be an international collaboration.”**


Respondents also thought that fostering innovation and research in AI might help address some of the concerns around the unethical use of AI. Of the 22,000 PhD-educated researchers worldwide in artificial intelligence (AI), only 386 are located in India. Moreover, research work in India in the field of AI is limited mostly to premier institutes such as the Indian Institutes of Technology (IITs), the Indian Institutes of Information Technology (IIITs) and the Indian Institutes of Science (IISc).

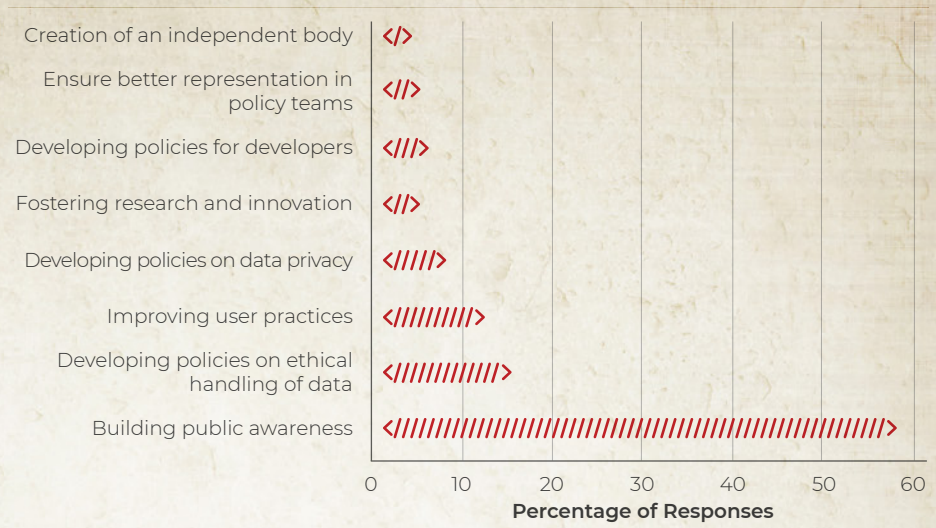



**Figure 4.** **Frequency distribution of focus codes for Question 1d**

## Q2a. How do you think AI powered tools can be deployed safely?

The coding team developed 50+ open codes that were synthesized into five focus codes for this question. As shown in Figure 5, most respondents felt that the best way to deploy AI-powered tools was to collect and use data ethically (27.4%). One of the respondents said:

> **The data collected should be in a safe and secure network. Any risky elements in the application should be mitigated immediately.'**

A well-echoed point was also the importance of building awareness among the public, with the development of AI literacy (21.9%), especially in rural parts of India. The prioritization of security and privacy by putting in place stringent policies, can help deploy AI tools safely. Lastly, AI systems must be reconfigured smartly, and continuously improved, in order to eliminate the risks of the inherent loopholes or biases that may be present in them (14.5%).

> **AI powered tools should be privacy-preserving. Aggregating data based on demography or interest can avoid the privacy concerns while still enabling the companies to generate revenue via ads.'**

One of the respondents had a particularly interesting take on the process of reconfiguration and testing of AI systems, as they compared them to the development and

testing of vaccines. The respondent suggested that AI algorithms should be tested with a group of potential users in processes similar to the phases of a vaccine trial. This process is usually performed by the development and quality assurance teams in software companies, before the algorithm is examined for manufacturing and safety defects for deployment.

> **Only way is keeping their way of execution, their vulnerability being updated from time to time in parallel to growing security problems against AI. Testing the AI rigorously not only by the developer but also by the people is necessary before deploying them (like how a vaccine is developed and tested)."**

**Respondents also felt that youth opinions and feedback are very seldom taken into account when framing policies, or creating guidelines for the ethical usage of AI.**

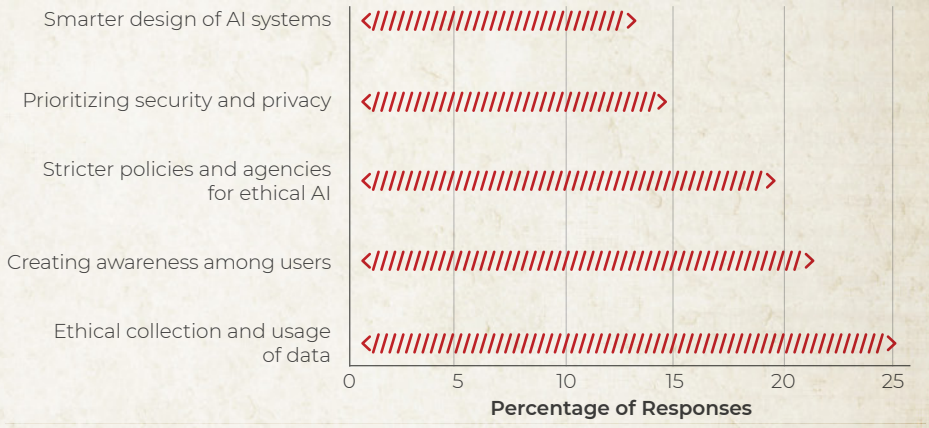



**Figure 5. Frequency distribution of focus codes for Question 2a**

## Q2b. Which industry sectors do you think have the biggest potential for AI?

Responses to this question were analysed using unigram analysis instead of focus coding. After a clean-up of the data to filter out irrelevant details (punctuation marks, stop words, etc.), the most frequently occurring unigrams in the text corpus were visualised (Figure 6). Top responses included the healthcare industry (37.3%), education (15.5%) and the banking and commerce (15.2%) sectors. These were followed by the marketing (8%), automobile (8%), manufacturing (7%) and retail sectors (7%). Other, less frequent responses included climate change analysis, agriculture, waste management and the defence sector. One of the respondents said:

> **I think the space and medical industry are going to have the biggest potential for AI, things like multiplanetary travel, and surgery of delicate organs can become more feasible in the future."**

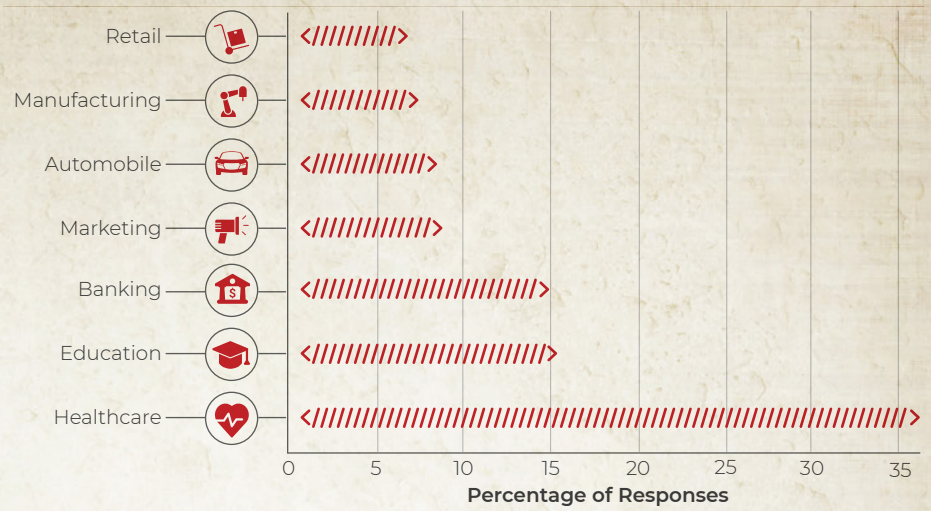



**Figure 6.** **Most frequently occurring unigrams for Question 2b**

## Q2c. How can youth voices be included in the development of India's ethical guidelines for AI for social good?

For Question 2c, 30+ open codes were created and later synthesized into four focus codes, illustrated in Figure 7. In general, almost all youth voices echoed the fact that there is a lack of awareness about AI among the public, including the youth (34.7%). Therefore, the first step in including young voices in India's AI guidelines would be to build their AI literacy. The respondents also felt that youth opinions and feedback are very seldom taken into account framing policies, or creating guidelines for the ethical usage of AI (35.1%), despite the fact that the young are the generation of users that is and will be affected directly by using AI unethically:

> **"Almost 34 per cent of the Indian population consists of youth. They have the power to change the nation. But adequate opportunities should be given to the youth to represent their ideas and policies for upliftment of the nation through social media, and discussion platforms organised by the government, etc."**

> **"Youth and experience are the two sides of the same coin. The opinions of the youth and experienced people may vary on some aspects but if both the viewpoints are channeled in one direction it may become fruitful for society at large. The young minds we have in India can**

**give a fresh and innovative perspective to India's ethical guidelines for AI."**

**"By spreading the word. Nothing is more powerful than words. Youth can do that naturally. They can reach out to those in need, explain [to] them the features and help them to get access. The process might look difficult. But it can be done in no time by involving many people."**

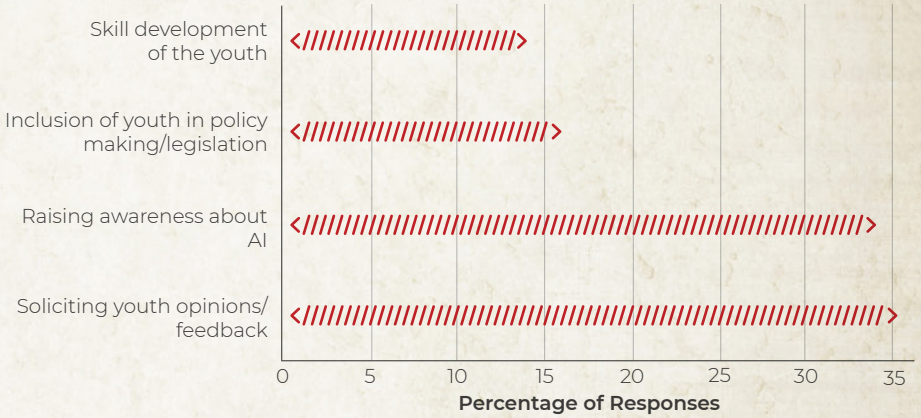



**Figure 7. Frequency distribution of focus codes for Question 2c**

# 14.4%

**of respondents felt that the continuous skill development of young people is essential to the development of sound ethical guidelines in India**

Respondents also brought to light the fact that soliciting the opinions of youth is insufficient if there is no representation of younger people within policy-making teams (15.6%). Tackling ageism in policy making continues to be a profound challenge in India. It is important to understand that with 'creative, rational and contemporary ideas', the youth can hasten India's developmental guidelines for the greater good. One participant clearly summarized why young voices are essential in creating policy:

> **"Youth voices are not experienced, but aware and thoughtful. If the policy makers listen to youth voices, they can be better decision makers."**

And finally, respondents felt that the continuous skill development of young people is essential to the development of sound ethical guidelines in India (14.4%). Participants felt that AI education should start at school level, as is now enshrined in the National Education Policy (NEP) released in 2020. Other suggestions included the creation of AI societies in schools and colleges, the holding of nationwide AI competitions and the creation of new internship opportunities for youth.

# 4. Key Messages

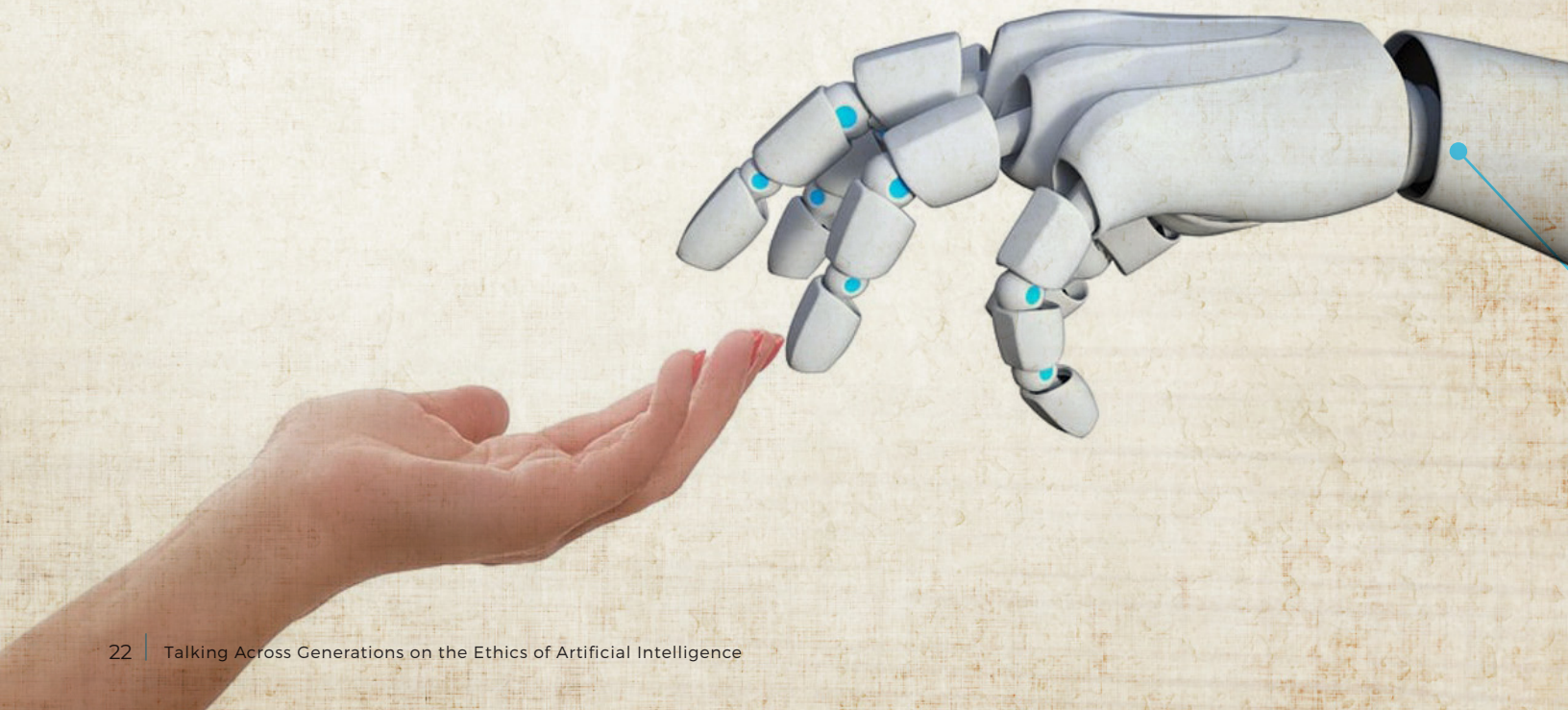
**Five key messages emerged from this diverse and dynamic set of youth submissions, encompassing actions that can be taken through policy and by young people themselves:**

## 1

## Responsible AI can be a force for good.

Concerns about ethical AI do not negate the potential of AI to improve life in many significant ways, by making education, communication, work and other aspects of our life more convenient or effective. Youth voiced that even the AI that collects data and profiles users can make life easier when the profiling is intended to target better services and help machine-learning tools better understand users' habits and preferences – such as map applications, email, and mobile personal assistants. Digital tools can process more data than human beings can, and may improve safety through the evolution of such devices as driverless cars.These dividends are especially useful in improving life outcomes, for instance for persons with disabilities, the elderly and other vulnerable groups.

As the UN Secretary-General's Roadmap for Digital Cooperation acknowledges, digital public goods are key to achieving the Sustainable Development Goals. Like the use and deployment of electricity, we can't make decisions about whether to deploy AI, only on how to protect individuals and societies from discrimination, discord and the erosion of human rights, where advances outpace policy guardrails.



2

**AI systems must adhere to a set of agreed-upon ethical principles, both at the policy level and hard coded into algorithms.**

At global, national and firm levels, we need to have an understanding of the core principle that AI systems must avoid violating human rights. Significant progress has been made in this regard globally, but, in some respects, the deployment of AI tools has been faster than the rollout of these frameworks.

Young people are calling for minimum standards rooted in the Universal Declaration of Human Rights, to safeguard the data rights of those who are most vulnerable: for transparency, legal accountability, human oversight, data security, consent and privacy, and more control over individual data. By identifying and establishing youth as a uniquely vulnerable group that requires protection from harm from AI, through an explicit elaboration of ethics and enforcement mechanisms, countries around the world can fully unleash AI as a force for social good.

Youth expressed several concerns about the deployment of AI in the development sector, and called for policy and ethical solutions. In some sectors, the very act of assigning decision making to AI must be closely examined, such as in military contexts through autonomous weapons. We need greater AI literacy and awareness not only among young people but across all population groups, particularly where AI systems are impacting people.

Ethical guidelines must include responsible data governance, such as ensuring the privacy of users and a transparency in data collection and use. The importance of human and regulatory oversight was emphasized by many young voices. The robustness and safety of AI tools were also aspects raised, to be ensured through frequent testing before deployment, and by giving due consideration to any foreseeable impact. Young people also raised the issue of monopoly and commercial models that depend on the monetization of data.

Young people are concerned about data breaches and cybersecurity, the mining of data by AI systems, and the lack of transparency in data collection, they are also occupied with the larger legal, criminal and socio-political ramifications of data deployment

3

## Data protection and privacy concerns go beyond individual security.

While young people are concerned about data breaches and cybersecurity, the mining of data by AI systems, and the lack of transparency in data collection, they are also occupied with the larger legal, criminal and socio-political ramifications of data deployment. This includes unease around hyper-personalization due to recommendations and predictions, the security of judicial system data on digital and/or AI platforms, black markets of data once data becomes the 'new oil', the weaponization of data, financial fraud, the complete erosion of privacy when digital companies can analyse and profile users based on every click, and human beings and AI not sharing the same ethical values when it comes to privacy and transparency.

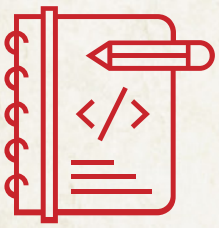
Strong data sets that are fully inclusive and representative of young people, can strengthen development programmes and service delivery systems. But any data collection must be sensitive to protecting the interests of its users. There must be transparency in how much data can be accessed and used, who owns it, and the consent that some of the younger youth – young teenagers – can give for the use of this data.

Youth respondents pointed to both technological guardrail solutions (better data and cloud security, protection against cyberattacks, and encryption) and policy and ethical tools, including greater transparency from apps on data collection, taking the bias out of algorithms, legal redressal mechanisms against AI, and drawing the line between service provision and algorithms being able to 'interfere' in people's lives by impacting their decisions.

## 4

## Ethical AI must become a social movement, through better AI literacy and multi-stakeholder approaches.

Some youth focused on creating more awareness – through curriculum, school and campus events, government-led public campaigns, youth networks, and AI literacy workshops – on the risks and opportunities of AI. A convergence of stakeholders – engineers, AI designers, youth and other impacted and vulnerable populations, government and policy makers, and private companies – is necessary to both establish ethical AI guidelines and ensure their enforceability and implementation. There were also calls to ensure better representation in policy teams.



**A convergence of stakeholders – engineers, AI designers, youth and other impacted and vulnerable populations, government and policy makers, and private companies – is necessary to both establish ethical AI guidelines and ensure their enforceability and implementation**

## 5

## Youth voices must be part of the development of ethical AI guidelines.

Young voices must form part of core consultations on the establishment of ethical AI principles at all levels: through dialogue with diverse groups of youth, embedding AI ethics in curriculum and education systems, inviting youth solutions for ethical AI design, ensuring young innovators and engineers are included in the design of AI systems, and recognizing the scale and impact youth can have in creating the social momentum for ethical AI. There must also be a focus within AI ethics policy conversations on the well-being aspect of AI deployment, especially as it pertains to young people, in terms of mental health impacts, addiction, and distraction costs.

# 5. Recommended Actions

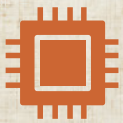
**A set of concrete recommendations for action has emerged for each of the key stakeholders – policy makers, private companies, AI designers and administrators, schools and universities, and youth networks and leaders – and these are provided below.**

## Policy Makers

Ensure representation from all stakeholders in establishing guidelines for ethical AI, and prioritize the protection of individuals and vulnerable groups from the adverse impacts of unregulated data collection and use by:

- Ensuring diverse representation on policy teams tasked with overseeing responsible AI, and including the voices of those who will be impacted by an AI-powered public service tool in the establishment of safeguards pertaining to that technology;
- Collaborating with key AI stakeholders, including developers and private companies, to create data privacy and data ethics standards;
- promoting AI literacy and awareness among individuals and communities, as well as promoting transparency in how digital tools are used and in data collection;
- Laying out clear data ethics and data privacy guidelines, ethics, laws or standards to safeguard the privacy of young digital natives;
- Ensuring the cyber-security of young people, both in terms of data theft and protection from harms like abuse;
- Creating independent and transparent mechanisms to evaluate and assess the ethicality of AI systems.

## Private sector and tech companies



Ensure that robust, well-tested and secure AI is deployed and that user data is collected with clear and transparent processes of consent-taking, and that data is not abused by:

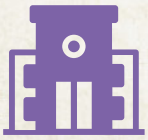
- Enhancing transparency through features that ensure clear user consent for data collection;
- Securing user data through robust cybersecurity and data-security measures, so that there are no inadvertent leakages;
- Establishing internal principles and standards for ethical AI;
- Eliminating, to the extent possible, the potential for statistical and algorithmic bias, especially where young people and their long-term life outcomes are likely to be impacted;
- Reconfiguring AI systems smartly and continuously to keep pace with emerging social needs and to eliminate bias.

## AI Designers


Ensure that AI systems are unbiased and do not end up exacerbating historic inequalities, and design technological guardrails to secure youth from harm from digital tools by:

- designing secure and safe data collection and analysis systems that can keep user data safe from leaks and cyber-attacks;
- designing ways to counter algorithmic bias, and consciously using fair, unbiased and accurate data sets;
- participating in policy dialogues on ethical AI so that they are informed by technical knowledge of AI's opportunities and potential future risks;
- testing AI-based systems with potential user populations, and not just on internal development teams, similar to how a vaccine trial is conducted. This process of testing and reconfiguration should be a continuous one, conducted throughout the life-cycle of an AI-based system.

## Schools and Universities



Build youth understanding around ethical AI, ensure that young people are protected from any harm that may be caused by AI systems, and raise awareness and literacy of AI ethics by:

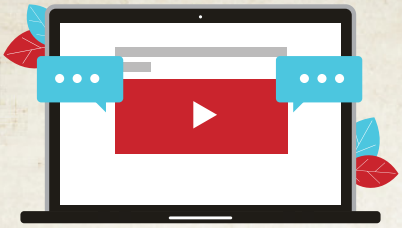
- including AI ethics and risks in curricula and teaching, and making the interdisciplinary link between technology, ethics and policy in the classroom;
- encouraging conversations on mental well-being, data privacy and the harms of internet addiction and cyber-bullying among youth, and soliciting their ideas and solutions on countering these harms;
- encouraging students to assist and participate in developing and using AI for good, and thinking about an AI future where technology can be used to solve real-world problems without bias.

## Youth

Create literacy and awareness around the risks and ethics of artificial intelligence and advocate for ethical AI policies that are representative of, and safeguard the rights and well-being of, youth by:

- spreading awareness among their peers, colleagues, classmates, and communities and networks, on the risks of AI and what young people can do to keep themselves safe;
- creating campaigns around the safe use of AI to both raise awareness and inform policy dialogues on ethical AI, especially around emerging issues of concern;
- co-creating a safe cyberspace for everyone and responsibly using apps and tools that collect data and use algorithmic decision making;
- developing technological tools that can combat AI risks, and building new cultures of AI for good;
- speaking up and voicing their concerns and ideas about the use of big data and digital tools.

# Notes



1. ITU (2017): www.itu.int/en/ITU-D/Statistics/Documents/publications/misr2017/MISR2017_Volume1.pdf
2. A concept coined by Marc Prensky in 2001: www.marcprensky.com/writing/Prensky%20-%20Digital%20Natives,%20Digital%20Immigrants%20-%20Part1.pdf
3. Hasse, Cortesi, Bermudez and Gasser (2019): Harvard Berkman Klein Centre 'Youth and Artificial Intelligence' Spotlight Series: https://dash.harvard.edu/handle/1/40268058
4. Digital 2020 Global Overview Report: https://datareportal.com/reports/digital-2020-global-digital-overview
5. UNICEF (2019): www.unicef-irc.org/publications/pdf/GKO%20Summary%20Report.pdf
6. Forbes (2016): www.forbes.com/sites/gilpress/2016/12/30/a-very-short-history-of-artificial-intelligence-ai/?sh=45dfe4ca6fba
7. Artificial intelligence has been defined by the UNESCO Recommendations for Ethical AI as: 'information-processing technologies that embody models and algorithms that produce a capacity to learn and to perform cognitive tasks leading to outcomes such as prediction and decision-making. AI systems are designed to operate with some aspects of autonomy and may include: (i) machine learning, (ii) machine reasoning, including knowledge representation, search, and optimization, and (iii) cyber-physical systems, including the Internet-of-Things, robotic systems, and human-computer interfaces.' (https://en.unesco.org/artificial-intelligence/ethics)
8. UNESCAP (2019): www.unescap.org/sites/default/d8files/knowledge-products/AI%20Report.pdf
9. International Finance Corporation (2019): www.ifc.org/wps/wcm/connect/corp_ext_content/ifc_external_corporate_site/annual+report-2019/download/ifc-ar-2019
10. Forbes (2019): www.forbes.com/sites/tanyamohn/2019/10/28/dummies-used-in-motor-vehicle-crash-tests-favor-men-and-put-women-at-risk-new-report-says/?sh=696b417b47bc
11. MIT (2019): www.technologyreview.com/2019/02/04/137602/this-is-how-ai-bias-really-happensand-why-its-so-hard-to-fix
12. World Economic Forum (2018) – The Future of Work: www3.weforum.org/docs/WEF_Future_of_Jobs_2018.pdf
13. IFC (2019): www.ifc.org/wps/wcm/connect/32e54505-3bfb-4198-b939-e1e88477715f1/EMCompass-Note-69-Role-of-AI-in-EMs.pdf?MOD=AJPERES&CVID=mNdPiNf

The publication can be cited as: Jamshed, N., Sharma, M. & Mochizuki, Y. (2021). Policy Brief: *#TAGeAI – Talking Across Generations on the Ethics of Artificial Intelligence: Youth-led guide on the ethics of AI and nurturing a responsible and transparent digital future.* New Delhi: UNESCO MGIEP, United Nations India, and UNICEF India.

TAG[e] AI was co-organized by UNECO MGIEP, UN in India and UNICEF India, in partnership with UN Volunteers India, YuWaah! and Generation Unlimited.

Organizers

unesco Mahatma Gandhi Institute of Education for Peace and Sustainable Development

UNITED NATIONS INDIA

unicef for every child

Partners

UN VOLUNTEERS India

YuWaah! The Potential of Young People

GENERATION UNLIMITED

unesco
Mahatma Gandhi Institute of Education for Peace and Sustainable Development

UNESCO MGIEP
United Nations Educational, Scientific and Cultural Organization |
Mahatma Gandhi Institute of Education for Peace and Sustainable Development
35 Ferozshah Road, ICSSR Building, 1st Floor, New Delhi 110001

**mgiep.unesco.org**
**To download:** mgiep.unesco.org/policy-reviews-and-interventions