BREAKOUT GROUPS SUMMARY

Global Forum on AI for Children

The Global Forum on AI for Children focused on five thematic areas, which were explored through dedicated breakout sessions each with a diverse panel of speakers. The panellists included representatives from the public sector and governments, non-profit organizations, private sector and academia. On day one (30 November), breakout speakers discussed existing knowledge and evidence pertaining to each of the five themes, while day two (1 December) focused on sharing experiences related to practice, policy and governance. These discussions are summarised below.

MORE INFORMATION AND RESOURCES CAN BE ACCESSED AT THE FOLLOWING LINKS:

- AI for Children project page
- Global Forum event page (includes recordings from the plenary sessions)
- Case studies of pilot organizations
- Policy Guidance on AI for Children 2.0
- #ai4children on Twitter
- Visual summaries of the plenary sessions
- Resources shared by participants
Alongside the harms and risks associated with AI, it is also important to **acknowledge the benefits** that AI can have for children and to develop appropriate mechanisms to measure and monitor positive outcomes. Articulating benefits and opportunities related to AI is essential for companies, so they can incorporate them into their business models and the lifecycle of their products and services. It is important to keep in mind, however, that not everything is measurable and by optimising only for measurable outcomes, other important aspects of well-being could be missed or overlooked.

When the impacts of AI on children are considered, the focus tends generally to be on well-being, while there is limited understanding about the impact of AI on **children’s mental health** – one key aspect of well-being. AI offers an opportunity to help address mental health challenges among children and young people, but at the same time, AI platforms that are not designed with children in mind can also expose them to psychological harm.

A holistic approach can help understand the **interrelation between the positives and negatives** of AI and identify trade-offs between the opportunities of AI and the mitigation of its potential risks. Understanding these trade-offs is important for the practical implementation of standards, regulations and policy. For example, when implementing the principle of age-appropriateness in design, it is necessary to understand how trade-offs shift, depending on the age and developmental stages of children.

Existing and evolving tools, such as **AI standards**, can help organizations that design and develop AI products and services to consider its potential positive and negative well-being impacts on children.

Speakers highlighted the following examples of AI standards that take into consideration children’s well-being:

- **IEEE* Standard 2089 on age-appropriate digital services frameworks** – based on the 5Rights Principles for Children, it outlines recommended processes for considering and mitigating against potential risks, as well as identifying goals related to benefits that can be achieved throughout the full life cycle of an AI product or service.
- **IEEE Standard 7010 on well-being metrics** – provides processes and methodologies for identifying the right indicators to measure.

* Institute of Electrical and Electronics Engineers
Initiatives and efforts that leverage AI to support the health and well-being of children are emerging in both the private and public sector, but they are not without challenges. For example, Helsinki University Hospital developed the Milli chatbot, an anonymous and non-stigmatizing online service that uses natural language processing to allow children and young people to ask questions regarding their mental health concerns. A challenge of using a chatbot to provide mental health advice lies in the risks of providing incorrect information or not addressing a sensitive mental health issue with the appropriate response. Helsinki University Hospital analyses conversations anonymously to identify areas where the chatbot is not able to respond appropriately as one possible mitigation strategy to deal with this risk.

Children have a different perspective and perception of safety, security and risk in virtual environments. When adults talk about design, they talk about AI, algorithms, machine learning, etc., but very young children do not usually make this distinction. For them it is an experience (i.e. in play and learning). It is important to understand what risks children perceive, but also what they are not aware of, to help design better and more ethical products. For example, research with refugee children in South Asia found that children who are very aware and concerned about their physical safety, have not considered some of the safety implications of the information they are sharing online.

Cross-cultural and cross-disciplinary approaches can help bring together diverse perspectives from different cultural settings and epistemological traditions (social science, psychology, etc.). Research and policy on AI needs to take into account the complex and multiple childhoods around the world, and recognize the distinct needs and vulnerabilities of different children. But often data used to train AI solutions are skewed towards the Global North and the specific needs and characteristics of children from the Global South are underrepresented in data sets and learning models. To achieve fair representation for all children, offline modes of engagement with children in low-resource settings (such as rural areas of less developed countries) are one key way of ensuring the inclusion and engagement of underrepresented children. Another important consideration is how children in these settings are accessing the internet (for instance, mobiles rather than desktop or laptop computers, using poor internet connections) and what this means for product design.

THEMATIC AREA 2

Ensure inclusion of and for children

“Children should be involved in all parts of the design cycle, not because it’s a good selling or marketing point ... it’s because of their rights, and it creates better products for them and ethical products for them, which should be the standard.”

Shuli Gilutz, UX researcher of technology for children
Communication with children during all stages of the product design cycle is key to integrating children’s unique perspectives into the design of AI products and services. This means continuously and regularly engaging with children to get their feedback and to understand their experience of a product or service. It requires building trust, so that children really open up about their experiences, concerns and feel comfortable sharing their ideas. It is important to let children know that their contribution is taken into account and considered.

The inclusion and incorporation of children’s perspectives and ideas in the design and development of AI systems can be achieved through more deliberative methods (i.e. perspective-taking and public participation workshops) that help move conversations about AI with children beyond first impressions to more nuanced understandings, including about the behind-the-scenes aspects of AI. Improving AI literacy for children, as well as for their parents and caretakers, is important for their meaningful engagement in these debates and decisions.

Despite the value of including children’s perspectives in the design of AI products and services, it is still not a priority for many companies. Regulatory instruments such as the Age-appropriate Design Code in the UK can enhance and incentivise the inclusion of children in design processes.

Child voices should also be included in the formulation of policy and regulation on AI. But, at what point in complex policy development processes should children be consulted? The Child rights impact assessment is a useful tool that helps to ensure children’s voices are heard in policy discussion. The impact assessment is conducted throughout the policymaking process; it offers a framework, guides the organization on what expertise is needed and what questions need to be asked and reported on in a child-friendly manner.
Children’s data has been collected by digital providers for some time. If used responsibly, such data offers an opportunity to serve children better, for example, by providing targeted health services, education or social care. However, there has been a growing awareness about the risks associated with the collection and use of children’s data. While most data governance frameworks tend to place the onus of responsibility on the individual with regard to what data they share, it is increasingly recognised that accountability should not lie with individuals, in particular not with children. Children are more vulnerable than adults due to their evolving capacities and ability to provide informed consent, therefore protecting children’s data and ensuring their privacy should be the collective responsibility of corporations and governments. Entities handling children’s data should be aware of the ethical implications and their responsibility to protect children’s privacy.

UNICEF’s Manifesto for the improved governance of children’s data sets out ten actions to inform the development and implementation of child-centred data governance frameworks and policies and advocates for setting out clear duties, standards and responsibilities to ensure that children are protected and that their data is used for good. Similarly, the Responsible Data for Children initiative provides guidance on how to put the child’s best interests first throughout the data use life cycle, from collection and storage to processing and deletion. Governments and businesses need to start thinking about the potential harms of technology and data usage from the design phase. Impact assessments, horizon scanning, and similar tools can help identify both positive and negative implications for data privacy in a way that is contextually responsive to the needs of children and their families.

Despite the recognition that human rights, and more specifically children’s rights, need to play a central role in data governance frameworks, practical challenges still remain around integrating transparency, accountability and ethics into technology systems. The risks associated with data collection and use by AI systems are not always fully understood by the organizations designing AI systems or by those creating regulations and policy. For example, in many developing countries, government actors who have decision-making and legislative power have low levels of understanding of AI technology and the consequences that AI can have on society, and in particular, children.
In order to ensure children’s privacy and security in a world where AI plays an increasingly prominent role, individual and institutional capacity-building on AI will be needed. Children and their caregivers also need to be empowered to maintain agency over their personal and sensitive data, including accessing it, sharing it securely, and understanding how it is used. This can be facilitated by improved AI and data literacy and through open and transparent communication. Institutional capacity-building involves upskilling relevant actors on the technical aspects of AI, as well as enhancing the capacity of policymakers and developers to grasp the ethical implications that socio-technical systems, such as AI, can have on the rights of children.

Example shared
The Hello Baby initiative in Allegheny County, Pennsylvania, USA is a voluntary programme for parents of new babies, which focuses on supporting child and family well-being and preventing maltreatment. Built on the principles of differentiated service delivery, the programme uses a predictive risk model to identify families who are most in need of support services. The model augments traditional knowledge about families and their needs by integrating data on health, crime and birth records. Transparency is key and explanations are given to parents about the purpose of data collection and the use of data within the predictive risk model. Parents are provided an opportunity to opt out of the service.
AI-powered education technology should be designed to consider the many different context-specific needs of students and their teachers. Taking such a nuanced approach, with the goal of ensuring more equitable enjoyment of the benefits of AI, means considering questions such as:

- Equity: Who has access to use AI?
- Skills: Who has the means to learn about AI?
- Governance: Who is responsible for data protection, students’ data ownership, and overseeing ethical AI?
- Impact: What is the impact of AI on learning outcomes?

Crucially, AI systems should be designed to support teachers, not replace them. Teachers can also play a key role in educating children about AI and helping them develop values and cognitive skills for life in an AI era. This includes helping children become conscious and confident users of AI systems. But, very often, teachers may not be informed themselves nor feel confident to teach AI literacy. Teachers need to be empowered through training and professional development to help them foster children’s AI literacy. Government can play a key role in co-designing and scaling solutions for upskilling teachers, as well as in evaluating the impact of EdTech on learning outcomes.

Example shared

MindSpark’s AI-powered Personalised Learning Software: This digital learning solution aims to bring together pedagogy (how students learn) and learning content with technology. MindSpark’s system analyses interaction data to try and identify why students struggle with a particular question and uses this to help teachers identify targeted and specific actions to address these learning challenges. It starts by assessing children’s current state of learning, continuously tracks progress and then adapts the curriculum to deliver a personalised learning experience.

To truly leverage the potential of AI for learning and education, AI EdTech needs to be inclusive by design. Currently, there are between 180 and 200 million youth with disabilities around the world, most of whom live in developing countries, yet many of the technology tools are not being designed to take this into account. A staggering 97 per cent of the top 1 million web pages are inaccessible to people with disabilities.
The aim should be universal design for digital tools and resources that can be accessed equally by everyone and are not only intended to be used by particular sub-groups, like those living with disabilities. This will require sensitising developers and designers about disabilities and understanding how someone living with a disability makes use of technology. Further, this means not only that the products and services must be accessible to users with different abilities, but the coding and data analysis software used to create AI-enabled EdTech must also be accessible.

Governments also have an important role to play and should be involved in the design and development of AI-powered education tools. The governance of AI in education needs to take a ‘whole-of-government’ inclusive approach where AI policy is aligned with existing policies in relevant sectors (education, economic development, etc.). This will take communication and coordination between different ministries. Without that, AI policies and strategies could fail to be operationalised and could remain ‘nice words’ without action.
Example shared

The European Commission’s Digital Competence Framework for citizens is currently being updated to include AI competencies: *knowledge* to be aware of the benefits, limitations and challenges of AI; *skills* to be conscious and confident users of AI; and critical but open-minded *attitudes* towards AI. About half of the education systems in Europe already use the Digital Competence framework or similar guidance. The European Commission’s goal is for 80 per cent of citizens to have basic digital literacy by 2030.
AI can be leveraged to protect children from online abuse and exploitation, intercept harmful content being shared and provide support to law enforcement officers responsible for investigating and policing crime that occurs on virtual platforms. For instance, UNHCR uses predictive analytics to try to identify children who are displaced, unaccompanied or left behind. In particular, it seeks to protect the rights of displaced children who may publicly share their personal information in an attempt to get help. UNHCR advocates for improved education about privacy online.

At the same time, if AI is not designed carefully with children’s safety in mind, it can amplify existing risk and exacerbate inequality, exclusion and discrimination. This may occur, for example, when classification algorithms used to determine eligibility for welfare and other service provision are built on data that is inherently discriminatory (e.g. it contains biases against children from lower socioeconomic or minority groups). In such cases, the very children who are disproportionately at risk of exclusion may fall through the cracks.

To overcome these risks, participants suggested a number of actions, such as establishing mechanisms to assess products before they are launched (i.e. assessing algorithmic impact), improving data-sharing between companies and researchers to understand the nature of the problem, and improving transparency and accountability for industry. To address risks inherent in the design of AI products or services, a human-centred and inclusive design approach and capacity-building among end users are also crucial. This can be achieved through a multi-stakeholder and multidisciplinary approach which involves children and their parents, as well as experts from different disciplinary backgrounds.

A number of helpful guides already exist on how to protect children in digital environments, including those powered by AI (e.g. ITU Guidelines on Child Online Protection, OECD Recommendation on Children in the Digital Environment, and a guide for protecting children online during COVID-19 by UNICEF and partners), but the challenge is making sure that they are put into practice. The OECD takes a number of steps to support the implementation of their recommendations. These include reviewing legal and policy environments, developing implementation guidance with examples of best practices, and an OECD AI policy observatory.