

**SUMMARY REPORT OF EXPERT ROUNDTABLE ON
COLOMBIA'S DRAFT AI ETHICAL FRAMEWORK:**

LENS ON YOUTH ISSUES

JANUARY 2021

RYAN BUDISH • URS GASSER



A. Introduction

The Berkman Klein Center’s **Policy Practice: AI**, in close collaboration with the Colombian government, hosted an expert roundtable offering feedback on Colombia’s draft “**Ethical Framework for Artificial Intelligence in Colombia**.” The roundtable discussion, which took place in late November 2020, focused on the youth-related elements of the draft framework. It specifically explored how Colombia’s ethical framework can best ensure that young people are protected from the possible negative impacts of AI, can best take advantage of AI’s opportunities, and have the skills necessary to shape both the future development of AI technologies and the policies around the technologies.

The invitation-only, expert roundtable was conducted under Chatham House Rules in order to facilitate an open dialogue collecting feedback on Colombia’s draft framework. Over 25 experts participated in the roundtable. These experts — participating in their individual capacities — came from academic centers around the world, intergovernmental organizations, as well as global, regional, and national civil society organizations. Participants had varied expertise; some have extensive experience with youth and education issues, others on policy relating to AI or other emerging technologies, and others with a deep understanding of national and international policymaking processes. From these diverse perspectives, the roundtable was able to collect feedback for Colombia in two broad areas: (1) feedback on the AI framework, with a focus on Principle 9 on youth issues; and (2) best practices for implementing the framework, based on experiences from around the world.

In advance of the event, roundtable participants were provided material to help lay a common foundation for the discussion. Participants received a copy of the full draft ethical framework, an excerpt from the framework focused on Principle 9 entitled “Prevalence of the Rights of Children and Adolescents,” **UNICEF’s Policy Guidance on AI for Children**, and background on **Singapore’s Skills Framework**. Additionally, the roundtable discussion began with a level-setting presentation from Armando Guio Español, an advisor to the Colombian government and consultant for the Development Bank of Latin America (CAF) who authored the framework. Armando’s presentation shared additional information about the goals of the framework, the process used in creating the draft, and the importance (and uniqueness among AI ethics frameworks) of the youth focus within the framework. Following that presentation, three respondents offered brief presentations sharing important perspectives on key issues like youth-centered co-design processes, the importance of teaching youth both the fundamentals of AI systems (e.g., what is an algorithm?), as well as ethical questions around AI and protecting youth privacy. For the remainder of the session, Sandra Cortesi, BKC Fellow and Director of BKC’s Youth and Media Lab, moderated an open discussion among participants.

The purpose of the roundtable — one of several different thematically oriented roundtables initiated by the Colombian government — was to provide expert guidance and feedback to the Colombian government as it iterates on the ethical framework and begins to consider how best to implement it. The comments and feedback collected from the roundtable are synthesized here, and this summary is being shared directly with Guio Español and the Colombian government as a key input to their future work.

These contributions are an embodiment of the work BKC engages in as part of **BKC Policy Practice: AI**, a public interest-oriented program that helps governmental, nonprofit, and private sector organizations implement AI best practices and turn AI principles into operational realities. Our goal with the Policy Practice is to create spaces for learning, knowledge-sharing, and capacity-building by bringing

together public and private decision makers with experts from around the world to tackle concrete implementation challenges. This roundtable and ongoing consultations with the Colombian government represent an opportunity for the Policy Practice to help shape Colombia's next steps as it develops and implements its own AI strategy and framework.

B. Summary of Colombia's Youth-Focused Principle

The framework establishes a series of principles for the design, development, implementation, and deployment of AI systems in Colombia. The framework aims to provide a foundation that will enable fair and responsible innovation while advancing the public good. Although many countries and international organizations have developed their own ethical frameworks for AI in recent years, one thing that is particularly unique about Colombia's framework is that it specifically recognizes the important role that young people play, both as a community impacted by AI technologies, but also as the future creators, users, and regulators of these technologies.

Principle 9 of the Colombia framework, entitled "Prevalence of the rights of children and adolescents," lays out a set of youth-focused elements to guide the development of AI technologies around young people. In particular, it makes it necessary to "recognize, respect and privilege the rights of children and adolescents" and importantly to ensure that no AI system works "to the detriment of their best interest." The principle goes on to focus on three broad areas: (1) data ethics; (2) ethics of algorithms; and (3) ethics of practices, and for each one describes the importance of ensuring that young people are part of the design and development of these systems, that young people understand these systems, and that they operate in their best interests.

The roundtable discussion focused primarily on participant reflections, reactions, and insights to this principle. The participants provided input on both the principle itself, as well as thoughts on how it could be most effectively implemented in Colombia.

C. Feedback on Principle

Participants provided key feedback on several aspects of the principle. In particular, participants' comments focused on two broad areas: (1) ensuring the principle was sufficiently dynamic and flexible; and (2) finding the right balance of protection, provision, and empowerment through law, education, and other mechanisms.

Dynamism and Contextuality

Overall, participants reflected a general belief that it might help to expand the principle's conception of AI technologies. Several participants suggested that Colombia find a way to incorporate the concept of **contextuality** into the principle. What participants meant by this was that AI is not a single technology applied in a single way; instead, AI encompasses different technologies, in different applications, used in different settings and for different purposes, across a range of diverse geographies and people (even across the already narrower subset of young people).

- This contextuality should reflect *local and community differences*. For example, one participant from an international organization that has conducted youth workshops on attitudes toward AI

noted that youth perceptions about AI differed across local communities, even within a single country.

- Throughout the conversation, participants emphasized the importance of recognizing the different needs and perspectives of *different stakeholders*. More than just teachers and students, but caregivers, parents, and the broader community, each of which should be consulted and engaged in order to understand their unique concerns and objectives for AI systems.
- This contextuality should also reflect the *differences in domain areas* in which AI is applied (e.g., health, education, transportation). One participant, with expertise in educational technology, gave the example that the use of AI as tutors in an educational context involved pedagogical, ethical, political, and practical questions that may not exist when similar technology is applied in a commercial context.
- Participants also hoped the contextuality could reflect that AI systems must do more than just “recognize the needs of users” because the *ultimate users may be unexpected and unintended*. In particular, there are AI systems designed for youth and educational contexts, and there are AI systems that are designed for general use or other contexts but may nonetheless be applied in a youth-centered context, and the principle should reflect those possibilities.

Similarly, participants reflected critically on whether the principle was sufficiently flexible and adaptable to the **dynamic** nature of AI technologies and the very human systems those technologies will operate within.

- One participant noted that the principle should be dynamic enough to respond to *generational differences*. By that, they meant that the principle should not be so narrowly scoped in its conception to only apply to the challenges of today, but should also be flexible enough to be adaptable to the next generation of young people and the next generation of AI technologies.
- Several participants discussed the *unknown unknowns of AI*. In particular, AI technology is evolving and improving so rapidly that we must ensure that all AI principles — and in particular, those applied to young people — must not make assumptions about future AI applications, uses, and impacts, and instead be open to the lessons that we will inevitably learn over the coming years.
- Given this highly dynamic nature, several participants stressed the need for *educating young people to enable meaningful participation*. Many participants described how, in their own work with young people, students wanted to be engaged in the process of developing AI systems and the rules that govern them. But to facilitate that meaningful participation, the principle should recognize the need for education and training of teachers, children, and parents and caregivers, so that they can understand the purpose of AI technologies, the potential benefits of the technologies, what data is being collected and why, the value of youth data, the steps taken to protect that data, the process of developing AI, and more.
- Although AI is highly dynamic, and its uses are expanding, several participants emphasized the importance of *delineating the limits of automation*. In particular, one participant discussed how their research with children identified areas that children felt were “human” and should remain untouched by AI automation. Several participants felt it was important to ensure that the principle

look beyond just protecting children but also empowering them to develop, define, and evolve their own dynamic sense of the limits of automation.

Balancing Protection, Provision, and Empowerment

Another important theme that permeated the discussion was the need to carefully balance protecting young people, providing them with the resources necessary to fully utilize and take advantage of AI technologies, and empowering them to have agency and control over their own role in the development, deployment, and use of AI technologies.

Participants made important observations about the tensions between **data inclusion and data privacy**. If governments and corporations are going to rely on AI technologies to make predictions about young people, or make decisions that impact young people, then it is important that those systems be as accurate as possible. Collecting more data about youth can potentially improve AI, but collecting more data also presents clear privacy risks.

- Participants noted that children have the potential to be *sophisticated controllers* of their own privacy. One participant described how their research showed that children are aware that they must often provide data to access services and that there is an opportunity to empower them through greater awareness about the exchange between data and access. Other participants agreed that guiding principles and education could help children improve their decision making about their own data, while others noted that placing the burden of consent on children may be unfair.
- Multiple participants discussed the importance of maintaining trust in institutions. When that trust is violated and data about children is abused, children often respond with *privacy protective behavior* such as avoiding situations and systems that require the submission of their personal information. As a consequence, children may actually avoid necessary services.
- Several participants were concerned about the potential *inaccuracies caused by data minimization* approaches. Avoiding collecting data about young people could lead to less accurate AI systems and worse outcomes when those systems are applied to young people. Instead, some participants suggested other approaches like youth participation, sandboxes, and continuous evaluation of AI systems would achieve a better balance of effectiveness and control.
- However, other participants argued that data minimization might be necessary to *prevent revictimization of vulnerable populations*. In particular, one participant who works with governments and NGOs working with vulnerable populations noted that collecting information about children who are already vulnerable revictimizes those populations both through the collection and, potentially, every use of that data. Instead, they argued that there must be clear needs for certain data before collecting it to avoid reenacting harms.

Another common theme that emerged was the importance of empowering young people through **education, active engagement, and co-design** of AI systems. By giving youth the skills, education, and resources they need, they can become advocates for their own rights, views, and ideas, and they can help shape AI policy and development to be in their best interests. However, participants noted that this responsibility cannot, and should not, be entirely on young people and will require a multis-takeholder approach, including industry, policymakers, and the community at large.

- Several participants wanted the principle to more clearly emphasize the importance of providing *agency for young people*. The discussion made clear that this agency had many facets. For example, some people emphasized the need for education about AI so that they can understand it, use it, and develop it. Other participants noted the possibilities of using AI to improve curriculums and digital skills. And others emphasized a set of cognitive, meta-cognitive, and social and emotional skills so that young people can understand how AI might impact other vulnerable communities.
- Many participants discussed how one important part of agency is giving young people the opportunity to actively *participate in the design of AI systems and AI policy*. The challenge, as many participants noted, is to engage young people without overburdening them. Additionally, one participant pointed out that young people are not all the same, and engaging young people in a multistakeholder process requires recognizing the differences among groups of young people.
- In addition to recognizing differences among young people, participants also discussed the importance of engaging diverse stakeholders, to ensure these systems reflect youth *by design* throughout their development and deployment process. Just as “privacy by design” requires sustained engagement by diverse stakeholders throughout the design, development, and implementation process, so too should addressing youth needs by design.

Finally, throughout the conversation, participants emphasized the **role of formal legal and regulatory systems** in ensuring that young people are both protected and empowered.

- Several participants recommended that Colombia’s principle should more *explicitly connect to existing human rights frameworks*. They noted that many of the same issues that were discussed within the roundtable were issues that have come up in human rights contexts. Therefore, Colombia could benefit from grounding the principle in existing human rights standards and legal frameworks.
- Many of the participants felt that the principle should more clearly articulate who bears the responsibility for ensuring its execution. Principles are not self-executing, and several participants hoped the principle could provide a more *concrete assignment of responsibility*. The participants debated whether this responsibility should be at the national or international level and wondered how differences between overlapping legal frameworks could be addressed.
- Finally, several participants wondered whether any meaningful change could be achieved without first addressing the *harms of surveillance capitalism* and data commodification.

D. Feedback on Implementation

In addition to providing feedback on the principle itself, roundtable participants also provided ideas, best practices, and guidance on how Colombia could best implement the principle. This feedback, grounded in examples and experiences from around the world, clustered into two broad categories: (1) multistakeholder approaches; and (2) legal and policy approaches.

Multistakeholder Approaches to Implementation

Participants discussed several models for engaging a wide variety of stakeholders, including children and youth, in the process of implementing Colombia's AI ethics framework. In particular, participants focused on ensuring appropriate **incentives and support for a variety of stakeholders, including children, parents and caregivers, teachers, and other community members.**

- Participants shared successful past examples of *hosting co-design workshops* and meetings, both in-person and online, in order to consult students and children directly in the creation of AI systems and policies. Some participants noted that a benefit of co-design and youth involvement is that it ensures that content and policies resonate with young people's interests, backgrounds, and needs, making them active participants and owners in the process and not just passive recipients.
- In addition to co-design workshops, participants suggested that Colombia explore other collaborative models of engaging young people, including *youth labs and youth councils*.
- Participants stressed the need to engage additional stakeholders, such as teachers, parents and caregivers, and other community members. In particular, multiple participants noted that *children rely on parents and caregivers, teachers, and others to provide for their safety*, and that extends to the privacy and AI contexts. Thus, fully implementing these principles will require the full engagement of these stakeholders as well.
- Similarly, one participant observed that *children should expect support from their entire community* in navigating these complex and challenging questions. They urged Colombia to consider how it can create opportunities for this kind of communal support.
- Finally, participants noted the importance of engaging the private sector as the primary driver of new AI technologies. One participant suggested that Colombia consider how to create incentives and resources to facilitate effective engagement of the private sector, given that their buy-in and support will be necessary to ensure the principles are fully adopted through the design and development phases.

Legal and Policy Approaches to Implementation

Participants described several legal and policy frameworks that could be helpful as Colombia considers the best approaches in implementing the AI ethics principles. One set of tools that participants described were those that provided **legal protections and incentives for experimentation and adherence to best practices.**

- One participant described several examples of governments creating *regulatory sandboxes for both policy and technical experimentation*. Such sandboxes could provide opportunities to develop, test, and evaluate ideas on a smaller scale without risking widespread impact on children.
- Several participants described how *voluntary guidance, industry best practices, and safe harbors* could encourage conformance to principles and make it easier for those developing, deploying, and using AI systems to ensure that they are meeting expectations concerning how they use data from children.
- Similarly, several participants mentioned the value of creating *case studies* that highlight the im-

portance of protecting children in the use of AI as a way of building support among policymakers and the private sector.

Throughout the conversation, participants mentioned several approaches for bringing **additional transparency** to the use of AI technologies. The discussion emphasized how transparency can encourage conformance with best practices and also aid policymakers, children, parents and caregivers, and other key stakeholders in making informed, evidence-based decisions.

- One participant described how *a data provenance mapping tool* could guide people through the data lifecycle to make transparent who are the responsible individuals (and accountable individuals) for data, particularly as it impacts children.
- Other participants described legal frameworks and tools that can *ensure responsible data use and accountability*. In particular, participants emphasized how children, parents and caregivers, and other stakeholders want to be engaged and active participants in determining how their data is used within AI systems, but facilitating that accountability requires sufficient transparency and insight into data collection, data use, and AI systems as a whole.
- Participants described how transparency can also be a carrot and not only a stick. For example, one participant described how their organization gave an award (in part selected by young people) for innovative and privacy-preserving uses of AI. This award *incentivized positive uses of AI*.

Finally, participants stressed the importance of **collaboration across legal and policy frameworks**. Participants noted that the Colombian AI ethics framework must be implemented within a complex network of overlapping stakeholders, laws, and organizations. Only through collaboration can Colombia implement this framework and ensure it is leveraging and supporting their existing efforts, rather than duplicating or competing with them.

- One participant noted that the Code of Childhood and Adolescence already includes many of the principles and rights established in the AI framework and wondered how to *ensure that these frameworks are complementary* and not contradictory or overlapping.
- Several participants suggested that Colombia collaborate with (or continue collaborating with) organizations such as UNICEF, UNESCO, ITU, WEF, Council of Europe, OECD, GovLab, IEEE, which are also working on AI principles — many involving youth.
- Importantly, however, participants also emphasized that collaboration may need to include *continued efforts in AI literacy and digital citizenship education*. Several participants worried that without those educational efforts, decisions would be left to specialists rather than including youth, parents and caregivers, teachers, and other key stakeholders.

E. Conclusion

Overall, participants in the BKC Policy Practice: AI roundtable discussion were enthusiastic and supportive of the work Colombia has done in developing its AI ethics framework. Participants were particularly supportive of including a specific principle about children — something relatively unique among AI frameworks.

The suggestions that the participants described, outlined in the above discussion, ultimately come down to a single question: How can Colombia ensure that children’s best interests are understood, considered, and protected throughout the process of designing, developing, implementing, and using AI technologies? In answering that question, participants had many suggestions, including AI education and skill-based training, co-design and youth engagement mechanisms, additional transparency, and more. All of these are important considerations for Colombia to take under advisement. The remaining challenge, however, is for Colombia to figure out how to capture those suggestions within the scope of the principle and within Colombia’s future efforts at implementing the principle.

F. Additional Materials Recommended by Participants

- P2089 - Standard for Age Appropriate Digital Services Framework - Based on the 5Rights Principles for Children <https://standards.ieee.org/project/2089.html>
- P7004 - Standard for Child and Student Data Governance <https://standards.ieee.org/project/7004.html>
- The Ethics Certification Program for Autonomous and Intelligent Systems (ECPAIS) <https://standards.ieee.org/industry-connections/ecpais.html>
- IEEE 3527.1-2020 - IEEE Approved Draft Standard for Digital Intelligence (DQ) -- Framework for Digital Literacy, Skills and Readiness https://standards.ieee.org/standard/3527_1-2020.html
- **Youth and Artificial Intelligence: Where We Stand** - Youth and Media team, Berkman Klein Center
- **AI Innovators Should be Listening to Kids** - Urs Gasser
- **Digital Citizenship+ (Plus) Resource Platform** - Youth and Media team, Berkman Klein Center
 - The following group-based educational resources on AI may be of interest:
 - **What Is an Algorithm?** - Learners will understand what an algorithm is, why algorithms matter, and how algorithms are used in both everyday life and computer science.
 - **Social Media and Algorithms** - Learners will explore the various ways algorithms help shape the content they see on social media and apply these understandings to their own social media feed(s).
- UNICEF's The children's rights-by-design standard for data use by tech companies <https://www.unicef.org/globalinsight/media/1286/file/%20UNICEF-Global-Insight-DataGov-data-use-brief-2020.pdf>
- AI4k12.org's Github Repository <https://github.com/touretzkyds/ai4k12/wiki>
- Dr. Samantha-Kaye Johnston's Presentation on Building an AI Ready Generation <https://docs.google.com/presentation/d/114PKijxTlzfAyBcFeQwZpry7XaJX-42H6dpKQxZh5qk/edit#slide=id.p>

- Resources to Complement the UNICEF Policy Guidance on AI for Children <https://docs.google.com/spreadsheets/d/1zKmFPZgnaOeuQafmcWRRp6l8BeyxSts2pC7wPndcYaM/edit?usp=sharing>
- Responsible Data for Children Initiative at GovLab <https://rd4c.org/>
- UNICEF's AI for Children Project <https://www.unicef.org/globalinsight/featured-projects/ai-children>
- Wayne Holmes, et al., Artificial Intelligence in Education Report <http://bit.ly/AIEDpdf>
- RD4C's Decision Provenance Mapping Tool https://files.rd4c.org/RD4C_Decision_Provenance_Mapping.pdf?utm_source=RD4C&utm_campaign=37145dfbb0-EMAIL_CAMPAIGN_2020_03_26_06_18_COPY_01&utm_medium=email&utm_term=0_8990a71369-37145dfbb0-
- YaM's Youth and Digital Citizenship+ <https://cyber.harvard.edu/publication/2020/youth-and-digital-citizenship-plus>
 - Part V of the report explores various ways recent formal and informal educational initiatives for youth around digital skills are addressing artificial intelligence, computational thinking, and data.
 - Additionally, Part V, section 3 — “Artificial Intelligence and Ethical Thinking”—offers a brief overview of ways youth may be able to approach emerging technologies, such as artificial intelligence, in ways that take into account the impact of their actions online not only on an individual level, but the broader online community (e.g., taking into account cultural, social, and regional nuances).
 - Part VI, section 2 of the spotlight is a case study about the gender gap in computer science, including how formal and informal learning settings can help address this divide.
- GovLab's ReinventED <https://youreducationyourvoice.org/>
- Lionel Brossi, et al., Inteligencia artificial y bienestar de las juventudes en América Latina <https://drive.google.com/file/d/1OenZSNPqHUUd39ltaUIZw83B0Lt0GF6K/view>