Honda Research Institute Japan & European Commission, Joint Research Centre

Case study

Name of initiative
Haru Robot

Implementing organization
Honda Research Institute Japan (HRI-JP)

European Commission, Joint Research Centre (JRC)

Initiative type
Artificial intelligence (AI) system

About the initiative
Haru is a prototype robot that aims to stimulate children's cognitive development, creativity, problem-solving and collaborative skills. Once fully developed, it is intended to be used in the home environment, as well as in educational settings, by children from different cultural backgrounds. Haru's design process involved children in Japan and Uganda to assess how they viewed concepts of fairness and explainability.

Alignment with UNICEF Policy Guidance on AI for Children
The initiative has aimed to:

- Prioritize fairness and non-discrimination for children
- Provide transparency, explainability and accountability for children

Location
Japan, Europe and Global

Launched
Prototype; pilot initiated in 2020

Overview

New developments at the confluence of artificial intelligence (AI), robotics and cognitive psychology show promise in supporting children's capacity for learning, creativity, collective problem-solving and socio-emotional development. For instance, personalized learning with the aid of autonomous and adaptive AI systems could serve children's unique needs, helping them co-create their personal learning trajectories. In addition, the embodiment of such systems may play a powerful role in impacting children's behaviour and development. Yet such innovations also raise potential ethical implications – such as issues of trust and non-discrimination – that are critical to children's well-being and their fundamental rights.

In this context, the Honda Research Institute (HRI-JP) collaborated with the European Commission's Joint Research Centre (JRC) to integrate UNICEF's Policy Guidance on AI for Children in the participatory design of a prototype robot called Haru. Haru aims to be a socially intelligent robot, which can sense and process human social and emotional cues, and then respond to these behaviours in a developmentally appropriate way, creating sustained conversation, play and learning. Honda's ultimate vision and goal is to equip Haru with the ability to be fully integrated into the classroom and the home as a helpful, dependable social companion that guides and nurtures children in a safe, responsible and ethical manner. JRC's goal is to understand the impact of robots on child development and to develop methodologies for trustworthy robots for children.

This project is an example of applying interdisciplinary collaboration and iterative, inclusive design thinking to AI and robotics to prioritize the perspectives and needs of children from diverse cultural backgrounds. In collaboration with schools in Japan and Uganda, HRI-JP and JRC worked together on a series of classroom sessions where children engaged in scenario and story-telling processes to elicit how they viewed fairness and explainability. This form of participatory design helped raise awareness of emerging ethical considerations and build the technical requirements and conceptual framework that will guide the integration of children's rights in social robotics and embodied AI.[1]

1. Embodied AI refers to algorithms that enable the functioning of real-world physical systems, such as the responses and behaviours of a robot.

Context and project origins

Context and project origins

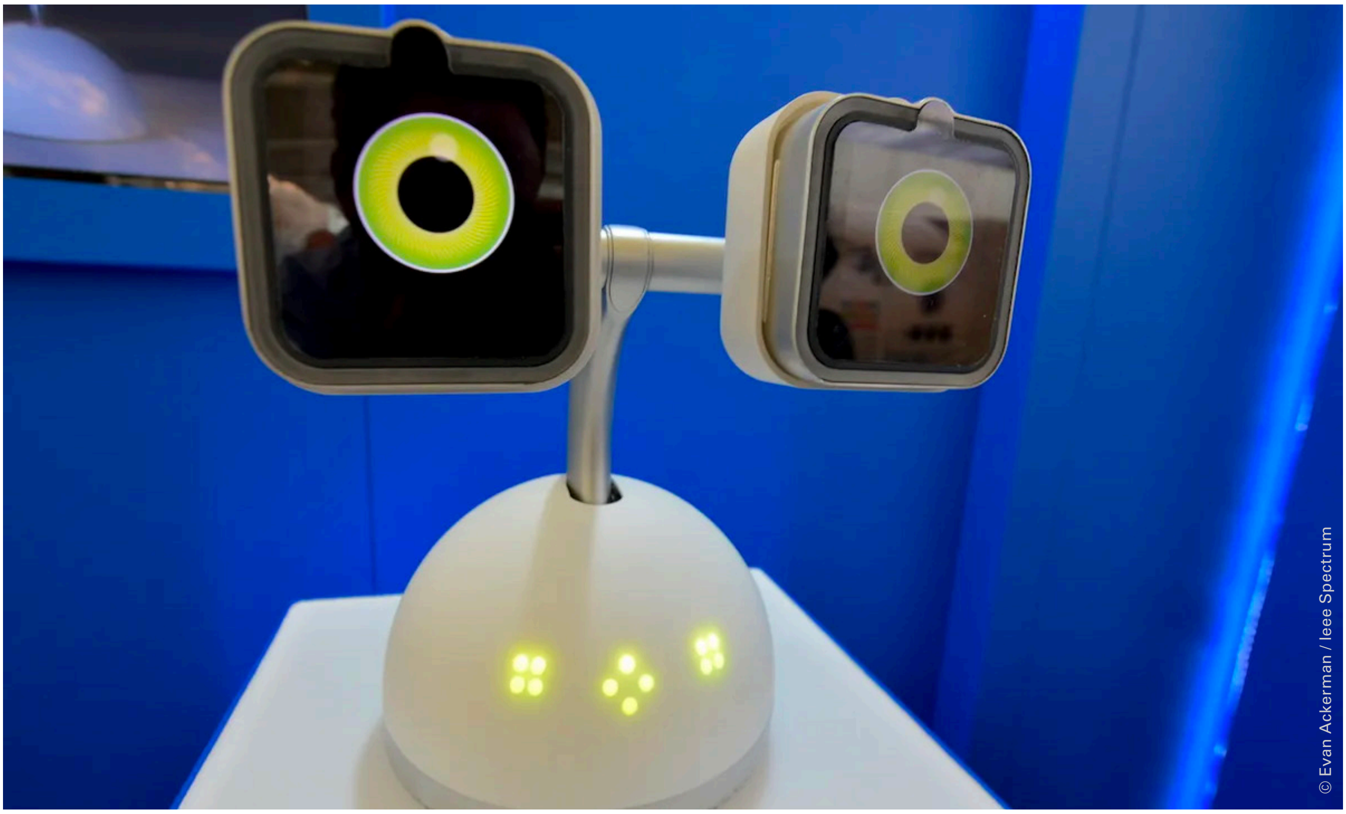
AI and robotics have already influenced human life in a multitude of ways. From chatbots that provide adolescents with legal advice or mental health support, to AI-based games and robots that help raise awareness about autism, AI is increasingly simulating social functions that have become fundamental parts of our lives. This integration of robotics and AI within society has not been without ethical consideration, however. Questions remain as to what constitutes an ethical AI system, particularly when it is designed as a companion for collaborative activities. Children interact with AI and robotics differently than adults. For instance, children may sometimes overtrust or attribute human emotions to robots, which could lead to deception or manipulation. Special care must therefore be taken when designing social and educational robots to safeguard the rights of children.

HRI-JP and JRC are conducting research on how the confluence of AI and robotics can support children by potentially benefiting their emotional and cognitive well-being. Such cross-field efforts between experts in ethics, child rights, developmental psychology, robotics and AI are an integral part of how HRI-JP envisions a society in which the relationships between autonomous cyber-physical systems, humans and nature might be mutually beneficial. In this context, the Haru robot provides a rich and multifaceted platform to investigate cognitive, emotional and empathetic engagement of humans with robots. As a learning companion capable of mimicking social and emotional behaviours, a robot like Haru should ideally exhibit the communication skill sets of a mediator, as well as the ability to stimulate creativity and collaborative problem-solving in advanced, diverse educational settings beyond the classroom.

HRI-JP and JRC are conducting research on how the confluence of AI and robotics can support children by potentially benefiting their emotional and cognitive well-being.

For this particular pilot, Haru's design focused on integrating UNICEF's policy guidance to support child-centred AI, with a particular emphasis on prioritizing fairness and explainability for children. In other words, the Haru robot design process and pilot period were devised specifically to show how the policy guidance could work in the design of a socially intelligent robot.

This was carried out by including children from different regions and socio-cultural backgrounds throughout the design process and by examining their perceptions of and responses to certain robot behaviours. HRI-JP and JRC conducted a cross-cultural pilot study on children's perceptions of fairness in robot-related scenarios with 20 school students in Japan and 24 school students in Uganda. On this basis, HRI-JP and JRC built a framework for future designers, engineers and policymakers to use when creating autonomous agents that operate in social situations with children.



Above: Haru robot designed by HRI-JP and JRC

About the AI system

About the AI system

HRI-JP and JRC involved a global consortium of experts with expertise across diverse fields (AI, robotics, ethics, social sciences and psychology) through a three-stage design process to better tailor the Haru robot to the needs and rights of the main user demographic: children aged 6 to 18. Haru was therefore conceptualized from scratch to try to maximize physically enriched social interactions with children, while aiming to uphold children's rights.

In the initial design phase, the group of experts identified and aimed to translate the fundamental behavioural elements of how children communicate into the robot's character, responses and actions. The consortium of experts then collaborated on defining the core technical requirements that would best enable positive child-robot interactions. The initial set of user requirements focused on enabling the robot to act as a storyteller. In this context, the first element that emerged through the participatory studies with children was their desire to share personal stories with the robot, under the condition that it would not share this information with others. Based on this finding, the research team and Haru's designers built a module into the system to safeguard children's privacy. This is a multi-level module intended to give control over private information primarily to children, their parents and caretakers. The children involved in user testing also discussed the importance of transparency, indicating their need to understand Haru's behaviour and intentions. As a result, a second module was introduced, which focused on the robot's explainability. The research team and designers developed a set of child-friendly phrases to explain privacy safeguards with the young users. For instance, Haru informs the child about the different sensors and cameras it uses to perceive its environment. The robot gives the young user the opportunity to turn off the camera, explaining that the collaborative activity can continue but that it will not track the child's movements. The translation of user requirements into technical requirements resulted from a collaboration between the developmental psychologist, robot designers, and developers through a learning process that is still ongoing.

During the design process, the consortium identified two target environments for which the AI robotic platform has been developed, namely a co-habitation robot in a home setting and a mediation-robot in a school setting. The co-habitation robot focused on tablet-based interplay, in which children would use a tablet to communicate with Haru. The robot would then respond through verbal and non-verbal interactions. In the home setting, engagement with the robotic companion was focused upon simple storytelling and conversations of lengths and complexities appropriate for children. With the mediation-robot, the design focused on group interactions through storytelling and collaborative problem-solving, such as the Tower of Hanoi – a game that involves stacking concentric rings according to size.

FIGURE 1: CO-HABITATION SCENARIOS

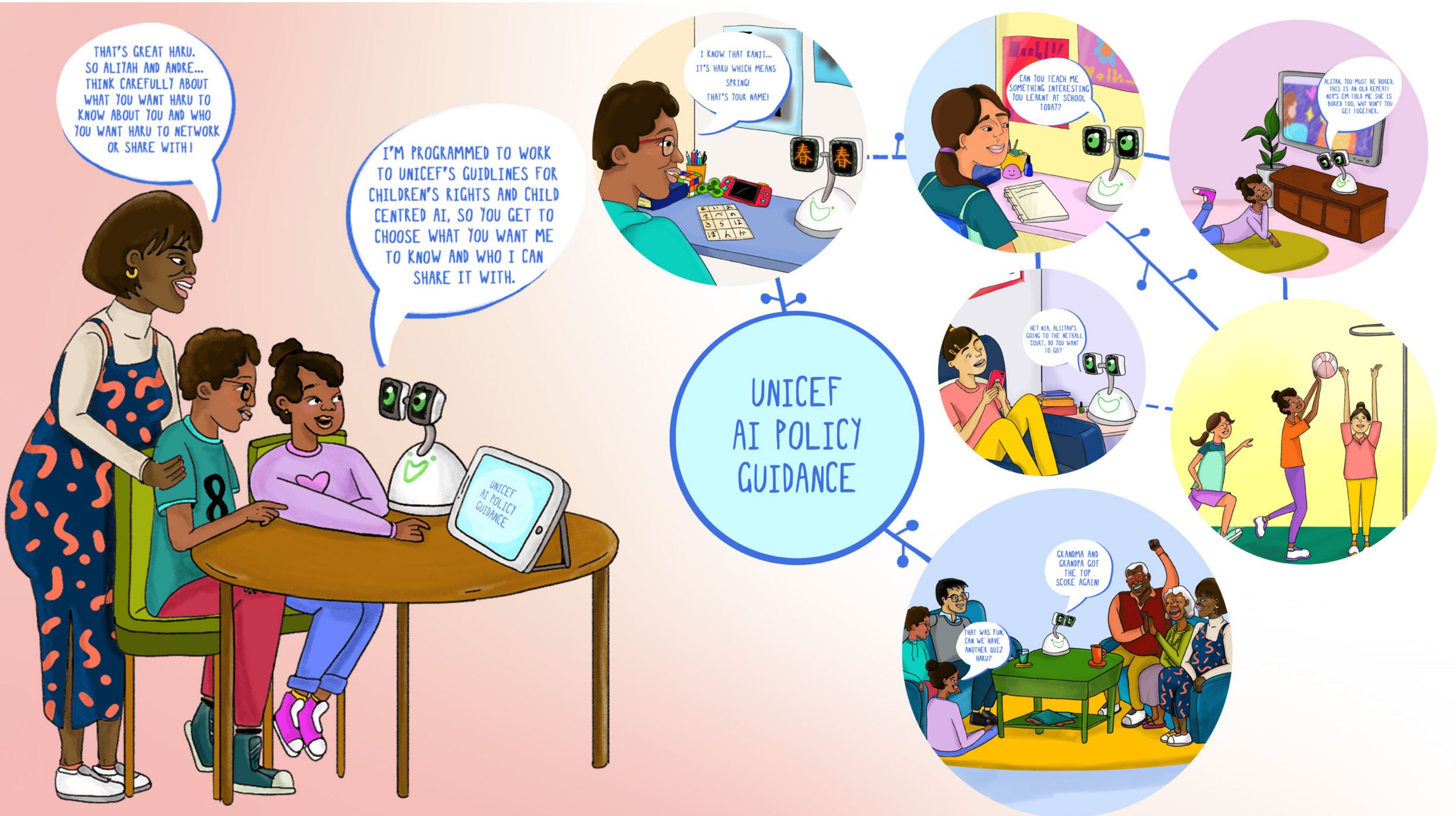


Image courtesy of HRI-JP / European Commission, JRC; artwork by Deborah Szapiro

FIGURE 2: CULTURAL DIVERSITY SCENARIOS

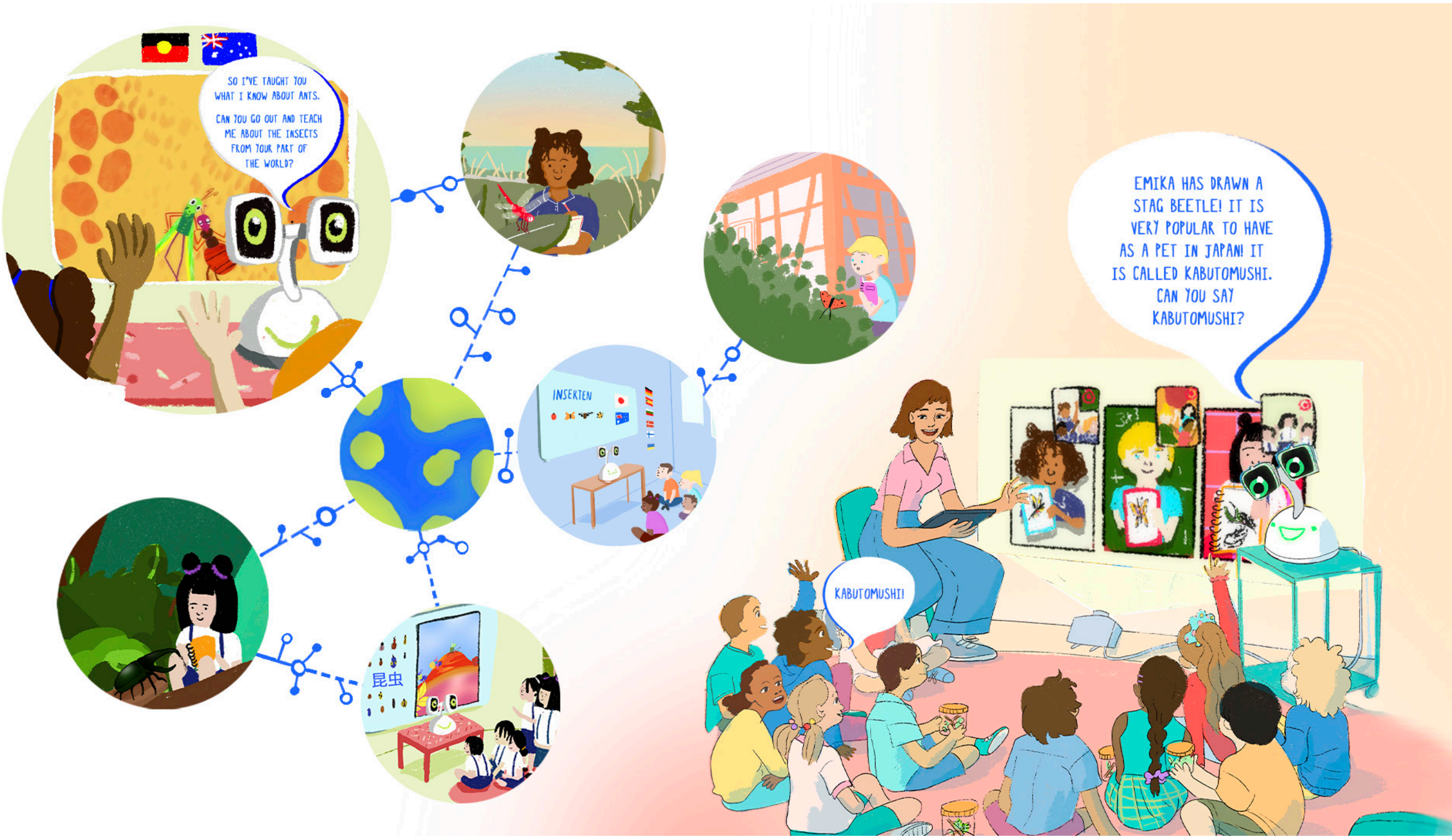


Image courtesy of HRI-JP / European Commission, JRC; artwork by Deborah Szapiro

In more technical terms, Haru relies on a child-centred infrastructure which is made up of three components: characterful robot behaviour, perception software, and interaction composer software. The first component (characterful robot behaviour) was inspired by animation styles that are well loved by children: bright and easy-to-understand expressions. In this pilot, the behaviour was specifically created to try to maximize children's acceptance and understanding of Haru. The routine itself is expressed through the companion robot using a combination of sensors, electronic displays, text-to-speech software and speakers, and actuations.

The second component (perception software) is an AI suite through which the Haru robot attempts to perceive its environment and humans, and understand its interlocutor, without identifying the specific person it is interacting with. In this way, children's privacy is preserved. This pilot feature functions as a trial run for future versions of Haru, where children and parents can opt out of data capture and collection, while preserving the integrity and performance of Haru. Importantly, the perception software does not mine and exploit data for AI system improvement.

The third component (interaction composer software) enables current and future AI engineers to develop robot behaviours and other interactions through simple, graphical programming. This allows for the creation of decision-making processes that guide the behaviour of the robot itself. Haru's core functions are scalable and used by partner universities and research institutes for research in child cognitive and socio-emotional development, as well as intergenerational and cross-cultural studies. Interestingly, this system also enables children to create their own applications through programming, testing and sharing their contents.

Haru's core functions are scalable and used by partner universities and research institutes for research in child cognitive and socio-emotional development, as well as intergenerational and cross-cultural studies.

In a nutshell, HRI-JP understood the importance of revisiting Haru's original design and modifying its core functions and capabilities to best fit its interactions with children. Through this iterative design process, HRI-JP has relied on the systematic inclusion of children (and, when opportune, parents and teachers), not just in the participatory user testing, but also at the level of software conceptualization.

Application of child-centred requirements →

Application of child-centred requirements

The HRI-JP research team and its JRC child-robot interaction specialists began focusing on two child-centric requirements that align with the UNICEF Policy Guidance on AI for Children, namely: prioritizing fairness and non-discrimination for children and providing transparency, explainability and accountability.

Prioritize fairness and non-discrimination for children

After conducting a literature review, the design consortium realized that there was a dearth of research and publications on the understanding and implementation of fairness and non-discrimination in child-robot interactions. HRI-JP and JRC resolved to contribute to filling this gap by conducting a series of participatory studies involving 20 school children from the Jiyugaokagakuen High School in Tokyo, Japan and 24 school children from the Bunabumali Good Samaritan Primary School in Eastern Uganda in order to gauge how children understood notions of fairness and non-discrimination. Four strategic research questions served as a guiding canvas:

- What does fairness and non-discrimination mean in the context of a robotic system?
- How do children perceive fairness and non-discrimination in robotic systems?
- What are the low-level, measurable elements that could be used to measure fairness and non-discrimination in the context of specific user scenarios with the Haru robot?
- How do engineers design a robotic system that is perceived to be fair for all children?

The participatory studies in Japan and Uganda worked through a four-session intervention, based on storytelling techniques, in collaboration with teachers. The studies were based on pedagogical approaches that emphasize the importance of eliciting children's active participation in the learning process through their interaction with specific artefacts or 'objects to think with' (for instance, the design elements of a robot companion). According to this approach, children were first introduced to the project and were subsequently asked to discuss what 'fairness' looked like in their own lives and within familiar stories. In addition to local traditional narratives, widely known stories, such as Pinocchio, were used as a basis for comparison to understand possible differences between the two geographical groups when considering the same story. In the last session, the children were asked to build on scenarios and discuss what fairness and unfairness in a robot would look like. Using this model, the researchers were able to glean insights into children's understanding of fairness.

To further prioritize fairness and non-discrimination, the participatory testing involved close and constant collaboration between Haru's interdisciplinary team and the schools in Japan and Uganda to elicit a variety of children's viewpoints from diverse backgrounds. Through this process, significant differences in discussions of fairness and unfairness emerged. In their description of unfairness, children in Uganda emphasized physical safety and resource scarcity, such as having to sleep outside when a house is overcrowded. By contrast, the children in Japan stressed concerns for psychological harm and emotional manipulation, such as adults reprimanding children's behaviours without understanding the underlying reasons.

The participatory studies helped HRI-JP and JRC understand the importance of utilizing storytelling and narrative techniques to elicit children's feedback in robotic design. Researchers at HRI-JP are also in the process of developing the infrastructure to support creative contents for robot applications that are designed and shared by children. A cloud-based platform that collects content and stories is being created to make it easier for children across the world to share their own stories that reflect their cultural background and demographics.

FIGURE 3: THE PROCESS OF A PARTICIPATORY DESIGN STUDY WITH THE STORYTELLING ACTIVITIES ELABORATING THE NOTION OF FAIRNESS

SESSION 1:
Introduction to the project
- Children's rights
- Assistive robots

SESSION 2
Fairness in everyday situations
- Child-driven

SESSION 3:
Fairness in a story
- Imaginary story
- Widely known story (Pinocchio)
- Traditional local story

SESSION 4:
Fairness in robot scenarios
- Gate-guard
- Learning companion
- Teacher
- Friend

Note: During the study, the activities were designed to be undertaken by individuals, in pairs, in small groups, and a whole class.

Provide transparency, explainability and accountability for children

In an attempt to build more trustworthy AI systems, the research team wished to provide explanations for Haru's AI processes that children could understand and appreciate. In this context, the research team at HRI-JP and the JRC developed specific user testing sessions that elaborated on the following two questions:

> What should explanations look like in the context of a robotic system, what should be included, and to whom should they be presented?
> Considering that the system is designed for children, what are the specific considerations that should be taken into account when designing for explainability (e.g. the cognitive developmental stage of the child or child-friendly communication)?

In relation to the first guiding question, the team devised a way for Haru to provide a level of AI explainability through the Tower of Hanoi game that it would facilitate. In the game, users organize rings by size upon three tower-like structures. Based on the user's movements, Haru would offer a response to help the child learn the game, along with a reply that explained the algorithm at play. The algorithm, based on the child's previous performance, applies certain pedagogical goals which would determine the type of explanation the robot would use. The team conducted a proof-of-concept study with the physical Haru robot with the children in Tokyo. As a result, the young participants not only improved their playing skills, but also demonstrated an improved understanding of Haru's behaviour and functions.

Ensure safety for children

Through Haru's design process, HRI-JP and JRC made sure that experts from the schools and a system engineer were present during the prototype testing with children. Haru itself – a physical robot that exists within a 3D world – was designed to ensure that it interacted with children in a safe manner. In the case of the robot that was designed for direct and long-term contact with children, since the physical Haru was still a prototype, in the co-habitation and home setting, HRI-JP decided to replace the hardware robot with a tabletop tablet-based robot. The robot had the same AI functionality, but the screen-based version was considered more suitable for piloting Haru in a long-term unsupervised manner to ensure further safety for children in a home setting.

Results, findings and impact →

Results, findings and impact

The Haru robot design process has already delivered valuable results and findings. Firstly, there was a positive social impact from the participatory design studies, as teachers and students became collaborators during the creation of Haru. The teachers' ability to facilitate their students' interactions and collect their responses helped them engage in meaningful, participatory research on AI applications for child development. At the test site in Uganda, the design studies led to an uptick in children's school attendance (particularly among girls), as they were excited to go to school to participate in the Haru project. In the high school in Japan, the Haru project has helped inform the design of current curricula for school-aged children.

Based on the Haru case study, HRI-JP and JRC researchers are developing a framework, founded on four technical components, for the conceptualization of social robotics and children's rights. Research indicates that children tend to attribute anthropomorphic characteristics and agency even to non-humanoid autonomous systems. As a consequence, embodiment brings unique gains for a child's development since it facilitates physical interaction with the environment, including social interactions. However, it might also result in a child overtrusting the robot, deception and possible affective and cognitive manipulation.

While HRI-JP and JRC researchers plan to further report on their design process, specifically around children's perceptions of fairness with robots, the results of the participatory user testing with children in Japan and Uganda were published at the 2021 Association for Computing Machinery Conference on Interaction Design and Children, winning the Best Work in Progress Paper Award.

Lessons learned and future challenges

Lessons learned and future challenges

The HRI-JP and JRC experts came away with three specific lessons that they wished to highlight. The first lesson is that children's cultural and economic status significantly influences their perceptions and needs regarding how robots interact within their daily lives. When the researchers examined the concept of fairness, children in Uganda focused primarily on material needs, such as physical protection, food and clothing. In Japan, children focused more on social and psychological issues. It is important for the AI robot to perceive, as well as possible, the different nuances and contexts in which it will operate, since it is meant to interact with children and become a social and learning companion. Designers must conduct significant research on specific user backgrounds before deploying Haru, or similar robots, to help build validity and trust.

The second lesson learned was that storytelling techniques – such as the ones used in the cross-cultural participatory studies – were an effective tool to understand children's perceptions of robots and their rights when it comes to AI. Because this storytelling tool was so effective, researchers have begun to include a large library of stories for Haru to access, and are creating a system where children can upload their own stories to Haru via a connected cloud, so that other children around the world can engage with diverse narratives.

It is important for the AI robot to perceive, as well as possible, the different nuances and contexts in which it will operate, since it is meant to interact with children and become a social and learning companion.

The HRI-JP and JRC experts considered a third and crucial lesson, which is the importance of interdisciplinary collaboration with a wide range of stakeholders. Priority was given to children's voices, which contributed to the understanding of the complexity involved in developing an inclusive robotic companion for children that promotes and aligns with their unique rights. The constant interaction among roboticists, developmental psychologists, child-robot interaction experts, designers, educators and specialists in child rights resulted not only in the first robotic prototype that attempts to align with UNICEF's policy guidance, but also in a valuable learning experience for all the stakeholders.[2]

2. Collaborators include: Tomoko Imai, Tiija Riinta, Joy Nakhayenze, Christos Zotos, Nikos Giannakopoulos, Luis Merino, Fernando Caballero, Natalia Diaz Rodriguez, Serge Thill, Séverin Lemaignan, Joe Daly, Nigel Newbutt, Louis Rice, Deborah Szapiro, Angelo Cangelosi, Guillermo Perez, Gloria Alvarez, Paulo Alvito, Georgios Andrikopoulos, Selma Šabanovic and Urs Gasser.

While inclusive and participatory testing is already underway with autistic children from the Mendip School in Shepton Mallet, England and students from the Arsakeio Lyceum, in Patras, Greece, further design trials will be run in other countries to continually refine and expand Haru's capabilities. Trials will be divided into long-term studies with existing partnerships and newer, large-scale studies; this pilot study is only the beginning of a larger scheme of future activities to embed child rights into robots for children and to include diverse populations to determine the evolution of the Haru system.

About this case study

About this case study

This case study is part of UNICEF's Artificial Intelligence for Children project, which aims to better understand how AI-powered technologies can protect, provide for and empower children. Through a partnership with the Government of Finland, the project developed a draft Policy Guidance on AI for Children, which contains practical recommendations and principles for governments and industry to create child-centred AI policies and systems. In order to inform and improve future versions of the guidance and inspire others, organizations were selected to pilot the guidance and share their diverse experiences to illustrate how AI-based initiatives could be designed to be more child-centred. The approaches taken, lessons learned and insights gathered will be one contribution to the global effort towards AI policies and systems that support children's development. View all of the case studies on our website.

Contact
Randy Gomez
r.gomez@jp.honda-ri.com

Vicky Charisi
vasiliki.charisi@ec.europa.eu

Acknowledgements
This case study was produced by the Office of Global Insight and Policy (OGIP) under the guidance of Melanie Penagos and Steven Vosloo, and was authored by Eleonore Pauwels. UNICEF is grateful to Randy Gomez from the Honda Research Institute Japan and Vicky Charisi from the European Commission, Joint Research Centre for their support in developing this case study. It was copy-edited by Eve Leckey, art directed by Kathleen Edison, and designed by Grace Leong. Cover: Photography by Brent Ninaber / Unsplash; artwork by MacWell.

UNICEF works in the world's toughest places to reach the most disadvantaged children and adolescents – and to protect the rights of every child, everywhere. Across 190 countries and territories, we do whatever it takes to help children survive, thrive and fulfill their potential of early childhood through adolescence. And we never give up.

The Office of Global Insight and Policy serves as UNICEF's internal think-tank, investigating issues with implications for children, equipping the organization to more effectively shape the global discourse, and preparing it for the future by scanning the horizon for frontier issues and ways of working. With dedicated expertise in seven policy areas – digital technology, human capital, governance, the environment, society, markets, and finance – the Global Insight team assists the organization in interpreting, and engaging in, a rapidly changing world. Visit us online to learn more: unicef.org/globalinsight

Office of Global Insight and Policy
United Nations Children's Fund
3 United Nations Plaza, New York, NY, 10017, USA

This document is interactive and designed for digital viewing.

Please consider the environment and refrain from printing.