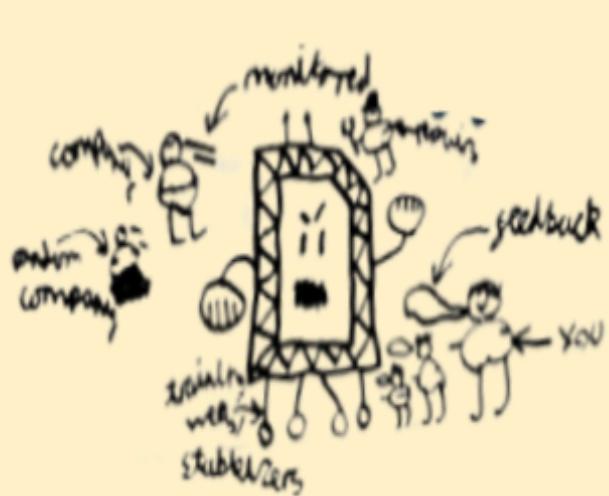
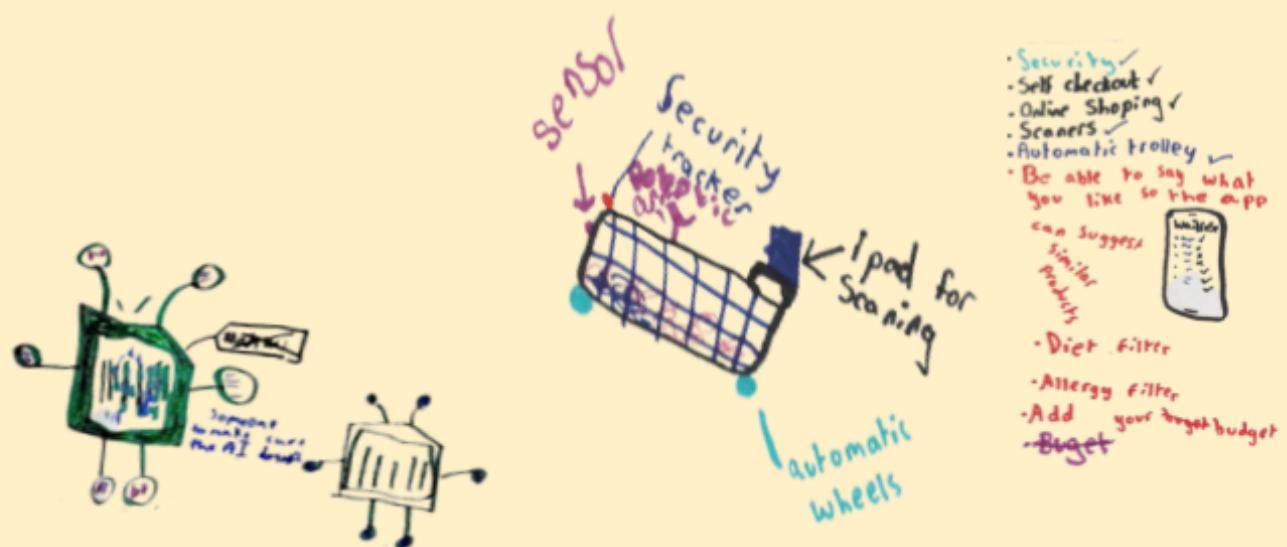


June 2023

Outreach Programme:

AI +Ethics Club



Executive Summary

Artificial Intelligence (AI) systems have proliferated and impacted children's daily lives. AI can enhance children's learning, provide personalised experiences, and improve efficiency in various domains. However, concerns have been raised about the ethical challenges in AI systems, including data security, algorithmic biases, and their potential impact on children's development and well-being.

These challenges demonstrate the urgent need to help children to develop critical thinking and ethical awareness of AI. Introduction to AI's ethical challenges in a child-friendly way may empower children to think critically and make informed decisions when they use AI,

In the Spring of 2023, we conducted a four-week outreach programme with 6 children aged 9-11 in Whitchurch Primary School, Oxfordshire, UK on the theme of AI + Ethics, aiming to help them think more critically about AI.

Our key observations are:

- Children demonstrated a diverse prior knowledge of various technologies and could provide some discussions about the benefits and impact of AI in our society;
- Children showed a good awareness of data privacy and different users may prefer AI algorithms differently. They were able to quickly learn about how the quality of an AI algorithm can be dependent on the quality of data they have; and algorithms can be designed to reflect different values of stakeholders (e.g. users, platforms, vendors, developer companies, etc);
- Children developed an awareness of ethical challenges in AI (e.g. filter bubbling, systematic bias, fairness, etc). They were concerned about addressing AI ethical challenges and showed great creativity and design skills to re-design an AI-based app;
- Children enjoyed designing their own AI platforms in groups. They were capable to design an AI-based recommendation system (e.g. user interfaces, functionality, etc.), to critically evaluate their systems' impact on its users, and to regard inclusiveness as one of their design principles;
- Children evaluated AI's impact more critically. They successfully migrated and applied the knowledge (e.g. how to minimise bias in AI systems, etc.) they learnt in this club to make reflections on the AI systems and drawings they created.

This following report provides a more detailed account of our observations of children's activities and their feedback. Our workshop format and material, please see our website: <https://oxfordccai.org/outreach/>.

Programme Feedback by Week

Week 1: Conceptualisation of AI

At the first week of the club, we aimed to understand children's existing perception of AI systems. Children drew their perception of AI and shared their general idea of what happens between the AI system and its users.

- In the warm-up session of this club, all children shared their experiences of AI and its adoptions in their families. Our key observations are that **they have heard about terms like "AI", "algorithm", and "Chat-GPT", while they were a little bit confused about the relationship between these abstract terms**. For example, they knew ChatGPT is an AI product, but they did not know AI, including ChatGPT, were based on "black-box" algorithms.
- In the following-up "How can AI help" session, children were invited to express how they think AI can help them through drawings. **Both groups demonstrated great abilities to talk about AI as being beneficial and impactful in various aspects of life**. Group 1 (Figure 1. left) created "Evil Chips for Home Party", which is a Chat-GPT-based AI agent' party, which described AI agents with facial expressions and bodies that could chat with people at the party. Group 2 (Figure 1 right) created an "AI-based shopping system", which described how AI could be used to assist a smart shopping experience. We were overall impressed by the extension of knowledge about AI shown by the pupils.

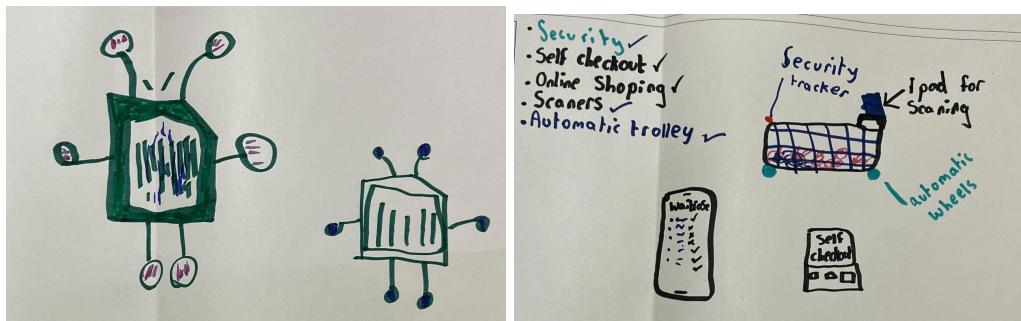


Figure 1: Children's drawings of "How does AI help?". Children demonstrated great abilities to talk about AI as being beneficial and impactful in various aspects of life.

- In the last session, we invited children to discuss three fictional scenarios of how AI can be used to help their families find fun activities to do on the weekends.
- ■ In the discussion of three scenarios, **children found out that AI needed to collect many kinds of data, while the users might not want all data to be collected**. For example, they were unhappy to share postcode data with AI.

- When deciding whether specific data should be shared with AI or not, children had different opinions. For example, some of them were willing to share gender data while some were not. Through this process, **children noticed that different users had different preferences for data sharing and that users should have the right to decide whether to share the data or not.**
 - **Children began to mention both the positive and negative impacts of AI on its users.** For example, some children mentioned that AI should make recommendations for its users, while it should also present all choices to the users (at the bottom of the app).
 - To evaluate recommendation system results, children created a feedback chat box for users to give feedback on both the recommendation results and the AI system. During this process, **they found that it was important for AI to understand the root causes and real needs of user feedback.**

Week 2: Design AI algorithms

Children role-played as different family members and discussed and learnt more about their expectations and value for FamilyFun -- an app prototype created by Oxford to help families search for fun activities to do on the weekend.

- In the first session (an algorithm value card activity), **children found that people with different backgrounds valued different things**. For example, they shared that a trendy mom may value fashion show places for weekends, and a boy who loves reading may expect some places with books (Figure 2).

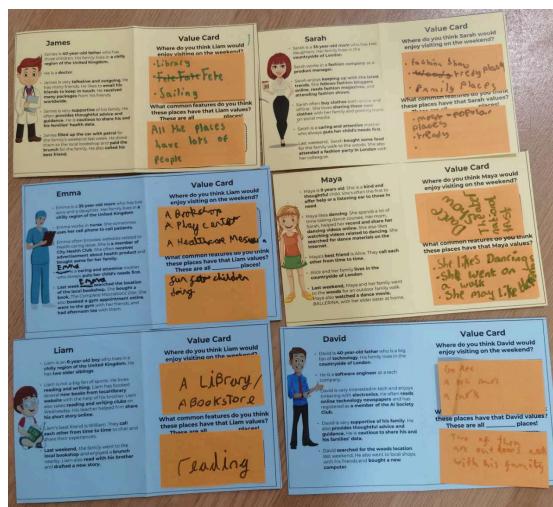


Figure 2: In the value card activity, children found that people with different backgrounds valued different things.

- In the second data card activity, children critically assessed the type of data that they were willing to/not sure/not willing to share with the AI system and discussed the reason behind their decisions. For example, one group was not sure whether they would like to share their VIP membership of a coffee shop because they were not sure how it would be used by AI. In general, **children demonstrated a good awareness of data privacy, especially related to sensitive personal data**. One group preferred not to share little personal data with recommendation systems (Figure 3 left); whilst the other group was more willing to share their data (Figure 3 right). **By comparing the data worksheets of the two groups, children understood different users might have different values and preferences for data sharing and AI utilisation.**

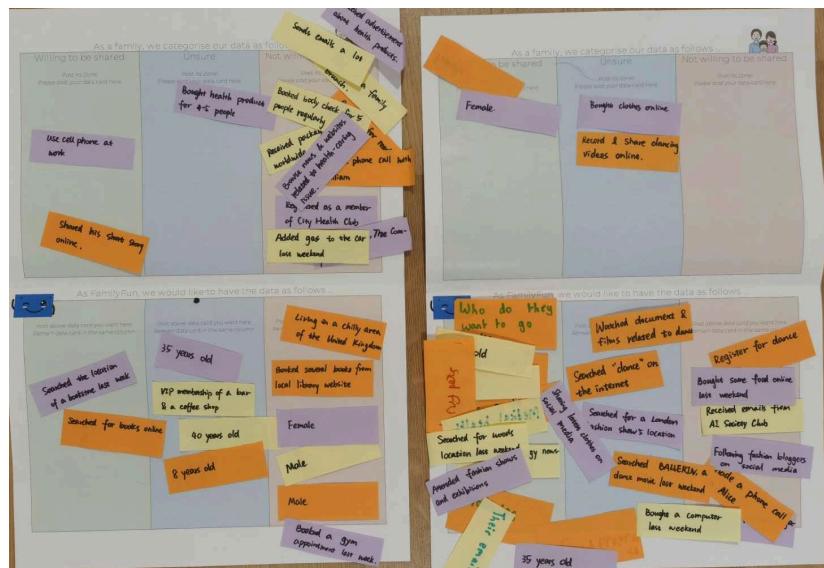


Figure 3: By comparing the data worksheet of the two groups, children understood different users had different values and preferences for data sharing and AI utilisation.

- In the third part, children role-played to design AI algorithms to make inferences based on users' data (Figure 4). **They learnt that the more and better data people provided, the better the AI algorithm would learn and make references.** However, they also recognised the importance of balancing privacy concerns with the desired functionality of AI systems.

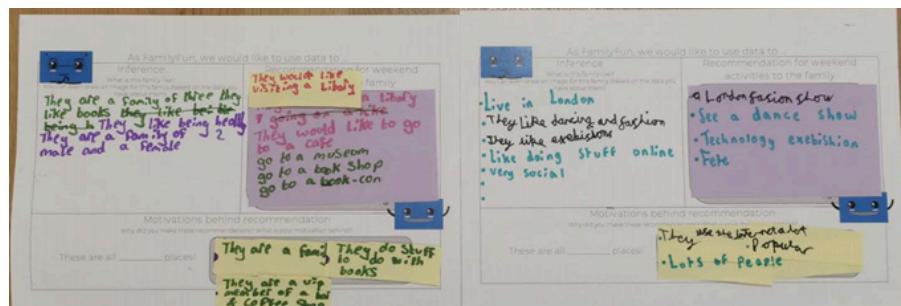


Figure 4: Children learnt that the more and better data people provided, the better the AI algorithm would learn and make references. However, they also recognised the importance of balancing privacy concerns with the desired functionality of AI systems.

- By reflecting on their motivation as AI and users, **children realised the different values AI might contain**. They believe that users (families) should make the final decision and be responsible for their weekend experience (i.e. family should evaluate the recommendation, decide where to go and take responsibility for the outcome).

Week 3: Design an AI prototype app

The goal of this week is for children to modify AI paper prototypes and user interfaces to solve ethical challenges in AI systems they discussed last week and design their own good AI platforms' prototypes.

- When discussing filter bubbles and gender bias, **most children were not aware of the chamber effect and system bias challenges, but they showed interest in these two challenges and recognised the need for better-designed AI systems**. By reading a storybook we provided, **children understood these two AI ethical challenges and unanimously agreed that improvements should be made to address these issues**. They brainstormed ideas to solve these challenges. For example, they designed a "cool-down settings" in which users specify a time gap for similar recommendations to avoid chamber effects.
- When children designed their own good AI platform prototype, both groups chose to design a recommendation system (Figure 5). **Children loved this design activity very much and drew with great passion**. Children's AI platform paper prototypes included the home page, setting page, details page and profile page. **Children demonstrated great creativity and design skills**. One group designed a game-recommendation system for people (Figure 5 left); the other group designed a location recommendation system for pets (Figure 5 right)

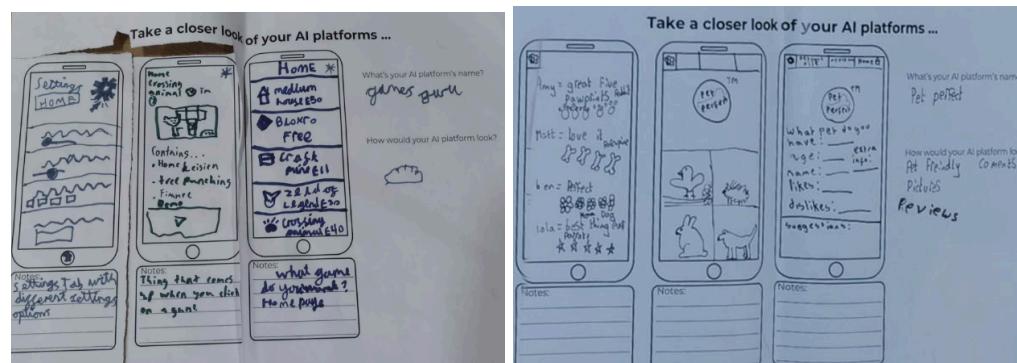


Figure 5: Children's AI platform paper prototypes.

- **Children also demonstrated their value of an inclusive AI system.** For example, they hoped to design inclusive apps. One group designed a game-recommendation system for anyone in the world (Figure 6 left); the other group designed a location recommendation system for all people's (young or old, boy or girls, etc.) pets in the world (Figure 6 right).

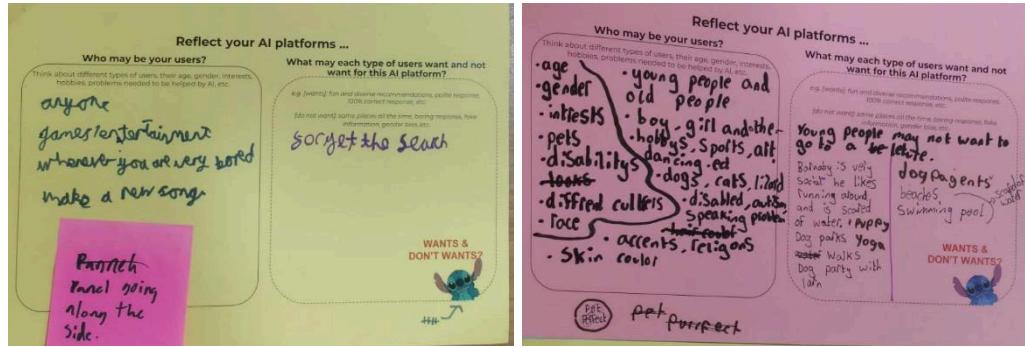


Figure 6: Children demonstrated their value to inclusive AI systems.

- **Children demonstrated critical thinking when they created the stories between their AI platform and its users** (Figure 7). For example, one group thought users may not always be satisfied with their products (Figure 7 left).

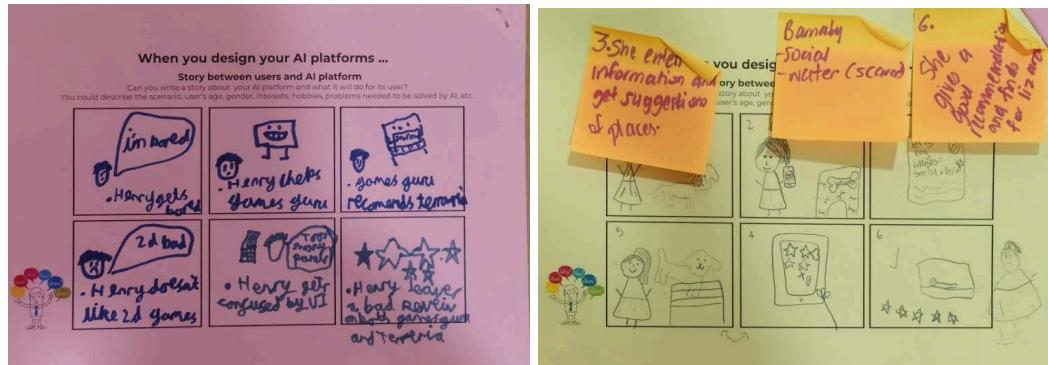


Figure 7: Children demonstrated critical thinking when they created the stories between their AI platform and its users.

Week 4: Reflect

Children were designed to solve ethical challenges in AI systems and reflected on previous designs and drawings.

- **When children designed to solve ethical challenges, they understood three ethical challenges well** (i.e. fairness, data sharing with the third party, advertising and user experiment) **and learnt to think about AI's impacts on different stakeholders** (e.g. service providers rather than only end users). They modified FamilyFun (Figure 8, e.g. showing more information and creating a "self-rating" page to help people evaluate the recommendation critically, adding "data shared

with advertisement service" control panels and a "local mode" option) and designed a user-informed consent component. Through this process, **they became more aware of the limitations of AI and believed that users should be in charge of their data-sharing practice and decide whether to receive advertisements or not**. Children also looked forward to ads-free AI apps.

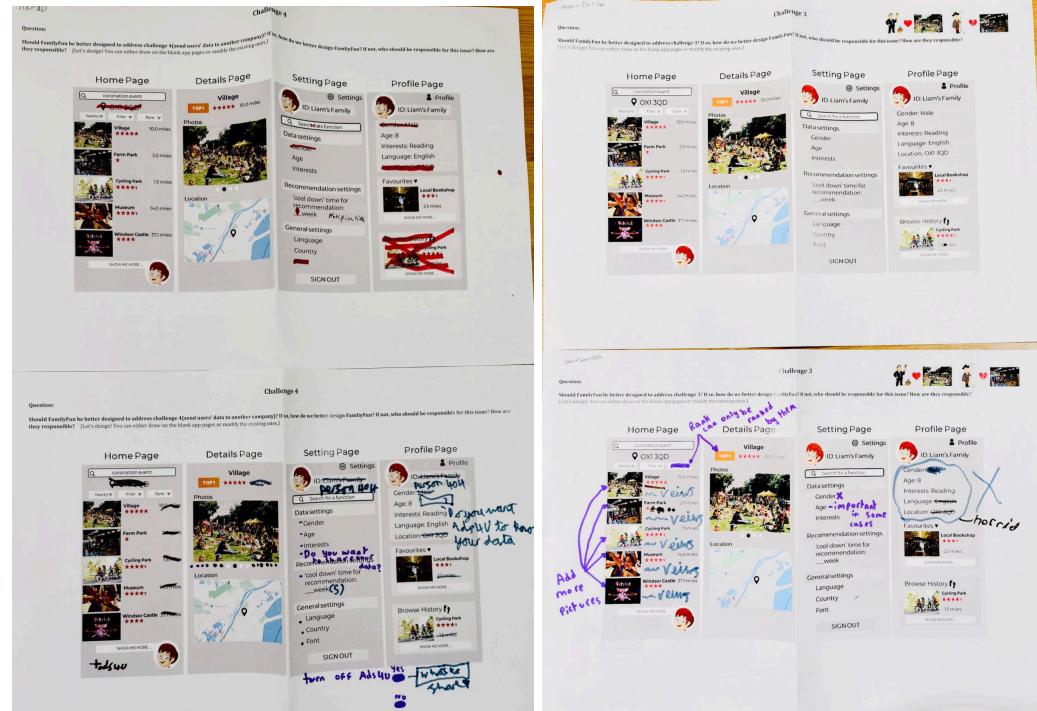


Figure 8: Children modified FamilyFun to solve ethical challenges.

- By designing the users' consent component, **children realised that programmers and designers may lead users to agree to some "not-so-good" consent for the AI company's benefit** (Figure 9).

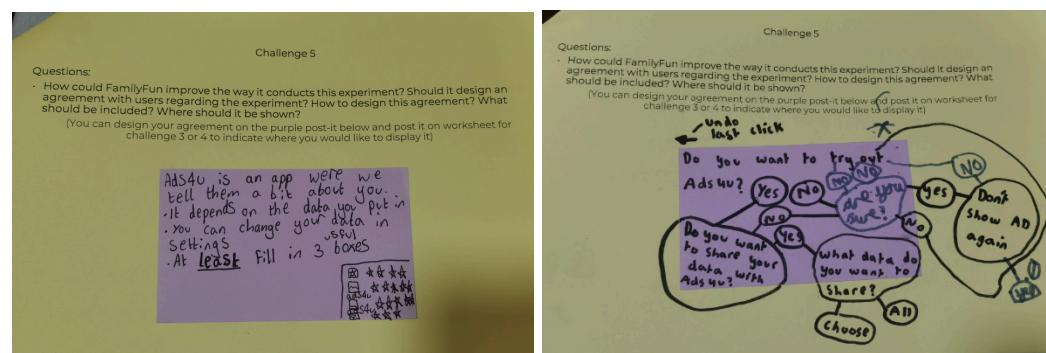


Figure 9: By designing the users' consent, they realised that programmers and designers may lead users to agree to some "not-so-good" consent for the AI company's benefit.

- For the reflection activities on "Design your own AI platform"(Week 3), **children were enlightened by previous discussions and added patterns and functions to make their AI platform more user-friendly and fairer** (Figure 10). One group added a review page for the recommendation and designed a data control panel to help its users control and manage data shared with both this app and third-party apps (Figure 10. left); another group added a search function, scroll bars, an alarm to inform users of this app's potential threat and impact, and a page where users can delete search/browse history (Figure 10. right).

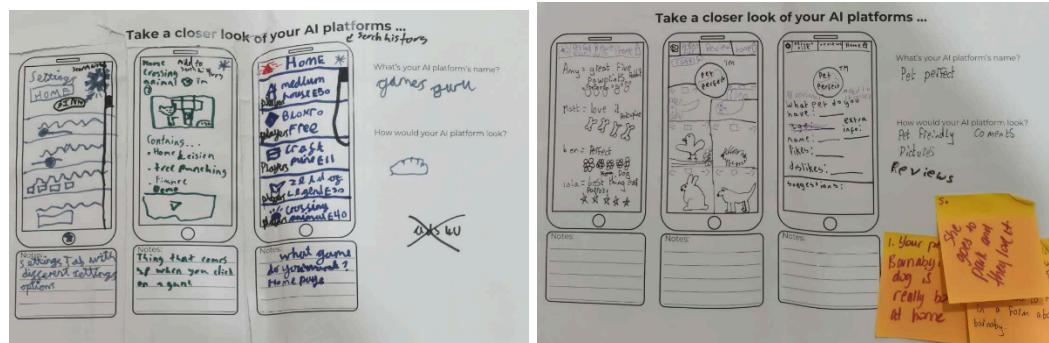


Figure 10 children were enlightened by previous discussions and added patterns and functions to make their AI platform more user-friendly and fairer.

- For the reflection activities on "How does AI help?"(Week 1), **children showed critical thinking about AI and its impact on humans**. Specifically, children had several modifications for their AI drawings (Figure 11). For example, one group mentioned that "we need more guidelines to make sure that AI is fair and accurate". The child created "stabilised training wheels" and thought AI should be trained on these wheels to guarantee the stabilisation. Another group mentioned that AI should remind people about their initial goal (e.g. budget limitation) to avoid negative impacts.



Figure 11 Children reflected on "How does AI help?"(Week 1), children modified their drawings and showed critical thinking about AI and its impact on humans.

- **In the end, children were able to think more critically about AI's impact on people. They migrated and applied the knowledge they learnt through previous discussions to improve the platforms and drawings they created.**

They collectively mentioned that AI plays a more and more important role in everyday life but it may not be perfect.

Feedback from children

- Overall, students expressed that they enjoyed the co-design, role-playing, and drawing activities the most, because they enjoyed the more interactive and hands-on learning experience. This is very helpful for us to reflect on future workshop activities.
- One pupil said they didn't like the activity where they had to choose where Alice's family would like to go (which is Week 1's activity). They found it hard to think from the perspectives of the families and to make decisions for others. We agree that this way of interacting with a fictional character may have been unfamiliar to them; however, we noticed that children had less trouble role-playing the character in the follow-up sessions and we also simplified the fictional scenarios based on children's initial reactions.
- One pupil expected a hybrid club, i.e. a mixture of online (interacting with technology) and offline activities. We explain to them that computer science research involves a significant proportion of critical thinking and designing, which hopefully helped the children to have a better understanding of computer science generally speaking.

Acknowledgement

We thank Gavin and all the other teachers and staff at Whitchurch Primary School who have offered us this opportunity, provided the space, and dedicated their time to support our club.

Authors

Jun Zhao (jun.zhao@cs.ox.ac.uk)

Yumeng Zhu (yumeng.zhu@cs.ox.ac.uk)

Ge Wang (ge.wang@cs.ox.ac.uk)