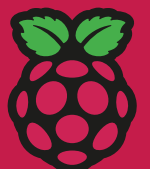**Mhairi Aitken** (The Alan Turing Institute)
Morgan Briggs (The Alan Turing Institute)

Engaging children with AI ethics

Engaging children with AI ethics

Mhairi Aitken (The Alan Turing Institute)
Morgan Briggs (The Alan Turing Institute)

Abstract

Internationally, there is growing interest in engaging children with artificial intelligence (AI) and data science. In this paper, we argue that rather than focusing solely on equipping children with skills to be the future AI workforce, we must also aim to equip children with skills to be the future — and current — critical public, which is needed to hold AI systems and their developers to account. As AI is impacting children's lives in ever more ways and increasingly shaping the future societies in which children will live and work, it is vital that children and young people are equipped to interrogate and understand the role of AI systems. This paper makes the case that education relating to AI must go beyond traditional STEM approaches to encompass ethical and social considerations relating to AI. This is important to ensure that children understand the role of AI in their lives (now and in the future) and are able to critically engage with AI to make informed choices about the ways in which they interact with AI. There are also substantial benefits for development and deployment of AI, since children's views and values need to be included in order to inform ethical practice.

Introduction

Internationally, there is growing interest in engaging children and young people with artificial intelligence (AI) and data science. This is considered important to build skills and capacities, and to equip the next generation to pursue careers in these fields. However, comparatively, little attention is directed at engaging children and young people with discussions of the ethical and social considerations around the ways that AI is designed, developed, and deployed. As AI is impacting children's lives in ever more ways and increasingly shaping the future societies in which children will live and work, it is vital that children and young people are equipped with the skills not just to develop the AI systems of the future but also to interrogate and understand the role of AI systems today. This paper therefore seeks to make the case that education relating to AI must go beyond traditional STEM approaches to encompass ethical and social considerations relating to AI.

Children interact with AI systems in myriad ways on a daily basis. Some of these interactions are intentional (e.g., playing with interactive toys or speaking with voice assistants), whereas others may be much less visible (e.g., in accessing tailored or personalised services, such as in education). AI is present in smart toys that "learn" and develop new skills when children play with them, and in smart home devices such as smart speakers and voice assistants, with which children increasingly interact. AI is also used to sort, filter, and target content online and may have a significant role in shaping children's views of the world, the information they receive, and the friendships they develop (e.g., through social media). AI is also used in ways that impact and shape children's lives through the provision or prioritisation of services in the public sector,

for example, through identifying which children, or families, are considered at risk and require interventions by social services. The significant impacts of such systems both for individual children and families as well as for wider society cannot be overstated.

There are a host of risks AI applications create for children. Not least among these are the potential transformative effects these technologies have on their development and their participation in the communities they belong to. Other big challenges that pose concerns are managing the privacy of children and their families in online settings in which data is constantly being collected about them.

There has been significant research conducted across many disciplines including psychology, education, healthcare/social services, etc. While there are different points emphasised across these literatures, there are some key points of overlap throughout. In the field of psychology, research has found that AI devices can alter young children's perceptions of their own intelligence (Druga et al., 2017; Howley, 2019; Williams et al., 2019). There is also a wider discussion surrounding the balance between protecting children and empowering them to learn and explore (Macenaite, 2017; Data Protection Working Party, 2009; Montgomery & Chester, 2015). Data is being collected about children and young people through what is called a 'data footprint' — all the data that is collected about an individual when they use online services (Kadho Inc., 2018; Lieber, 2018; Gibbs, 2015; Lupton & Williamson, 2017; Taylor & Michael, 2017; Harris, 2017). This footprint can be used to profile children and young people as well as to personalise ads and products, among other harmful uses. Another topic that is being widely discussed is the potential insufficiency of traditional forms of informed consent (Berman & Albright, 2017). The frequency with which parents and guardians have to sign consent

forms has given way to 'consent fatigue', in which details outlined in the consent form may be overlooked due to the high volume of consent forms present, along with the fact that parents and guardians may not be in the position to fully understand the best interests of the child (Macenaite, 2017). Furthermore, there is the overarching question of individualised notions of consent versus the average child dilemma. Should the age of consent to access certain online services be generalised, as it currently stands in GDPR, or should it be individualised to cater to different levels of maturity, development, and the unique needs of individual children?

One of the largest challenges in this field is the fact that often services are not designed with children in mind, but they are accessed by children (Barassi, 2018; Howley, 2019). For example, when a 10-year-old child asked Amazon Alexa for a challenge to do, Alexa responded with a challenge that placed the child's well-being and safety at risk (Segal, 2021). In fact, it was later found that Alexa pulled this so-called 'challenge' from a website in which parents were warning other parents about letting their children do an activity such as this (Segal, 2021). This instance exemplifies the possible harms that can occur when services are not designed with children in mind but are accessible to them.

The ongoing dialogue on children's rights as they relate to AI should be much more than an analysis of privacy concerns. While privacy is an important consideration, the best interests of the child must be considered. This is precisely why children and young people should be engaged on topics of the design, development, and deployment of AI systems that use their data. Children and young people have unique needs and considerations, and these should be not only taken seriously but incorporated into ongoing and future dialogues on this topic.

Introducing AI ethics

To explore how children can and should be involved in these processes, we must first better understand the landscape that enables an analysis of the ethical and social implications that AI technologies may have on society. This field is called Artificial Intelligence (AI) ethics. AI ethics is a growing field of research, which aims to mitigate the possible negative impacts of the uses of these technologies while maximising the value and benefits that AI can bring. It also aims to engage community members, policymakers, and AI developers to consider the effects that AI technologies may have.

There are a wide variety of ethical concerns expressed regarding AI and its role in society. Some of these concerns relate to the ways in which AI works, or how it has been developed, for example, whether AI has been trained on biased or incomplete data, which might lead to it reproducing or exacerbating inequalities in its outcomes. Other concerns relate to the impacts that AI has on society, for example, through producing unfair outcomes or changing the ways that services are delivered and accessed, leading to transformative impacts on society. Mittelstadt et al. (2016) labelled these broad categories as "epistemic" and "normative" concerns. While the two are interlinked, epistemic concerns draw attention to potential shortcomings in how AI is designed and developed, while normative concerns focus on the impacts AI has on society. AI ethics engages with both sets of concerns and notes that the ethical challenges associated with AI are interwoven with broader, long-standing social, political, and cultural factors (Aitken et al., 2021). AI ethics requires a combination of technical and social approaches that take account of the social, cultural, political, and economic dimensions of data and AI, and the ways in which these dimensions have shaped how AI is designed, developed, and deployed as well as the impacts it has. This entails broader consideration of the role these technologies play in society and the conditions under which they may be appropriate and acceptable (Aitken et al., 2021).

Importantly, ethics is not the same as legal compliance and there may be significant differences between what is legally permissible and what is ethically acceptable. Indeed, in many instances, ethics requires going substantially beyond legal requirements. While laws and regulation set out what we must or must not do (e.g., in terms of data protection, fair processing of data, or safeguarding of children), ethics grapples with the tricky questions of what we should or should not do (e.g., in what contexts or for what purposes should an AI system be deployed? How should the benefits of technologies be equitably distributed? What are the reasonable expectations users should have around privacy and consent?). Ethical questions typically do not have straightforward answers or clear-cut solutions, rather they require nuanced consideration and engagement with diverse perspectives to ensure that approaches taken align with societal values and expectations (Aitken et al., 2021).

An overview of AI ethics: Principles and concerns

Given the tricky nature of ethical considerations, ethical approaches are typically guided by principles rather than fixed rules, and, as the field of AI ethics has grown, a proliferation of principles and guidance have emerged to attempt to address these tricky questions and guide ethical practice (Aitken et al., 2020). These principles have been developed and adopted by a range of organisations including research institutes, policy bodies, and tech companies of all sizes. While this can be taken as an illustration of the significant interest — and investment — in this field, it has equally been criticised as enabling organisations to engage in 'ethics shopping' — selecting the set of principles that most closely aligns with their current practices, or which do not require them

to make substantial changes (Floridi, 2021). This is closely related to criticisms of 'ethics washing', which are often levied at organisations that make statements about their ethical commitments without taking meaningful actions or enforcing ethical practices (Floridi, 2021).

While there is a proliferation of sets of principles and guidance relating to AI ethics, there are generally common themes within these. Fjeld et al. (2020) carried out a review of existing sets of principles relating to AI, from a wide range of international organisations, and identified eight main themes that consistently emerged within these:

• Privacy
• Accountability
• Safety and security
• Transparency and explainability
• Fairness and non-discrimination
• Human control of technology
• Professional responsibility
• Promotion of human values

These principles highlight the relevance of both technical and social methods to underpin ethical approaches to AI.

In combination with principles, ethics also requires reflection on the values that underpin the innovation and deployment of technologies. At the Alan Turing Institute, we have produced a guidance document entitled *Understanding artificial intelligence ethics and safety: A guide for the responsible design and implementation of AI systems in the public sector* (Leslie, 2019). This guidance helps to lay the foundation for key principles related to AI ethics. There are four values that support, underwrite, and motivate responsible innovation, referred to as the SUM values; these were created to help researchers think about the possible impacts that using AI could have on society. This is also referred to as determining whether the use of AI is 'ethically permissible'. The four SUM values are: respect, connect, care, and protect.

Respect the dignity of individual persons
Connect with each other sincerely, openly, and inclusively
Care for the well-being of each and all
Protect the priorities of social values, justice, and the public interest

In order for an AI system to be 'ethically permissible', it is important that we consider how each of these four values are met, so that our uses of AI do not produce negative and harmful effects. Some additional considerations that fall under these four SUM values are things like ensuring everyone is free to make their own decisions about their own lives, making certain that diversity, participation, and inclusion are prioritised throughout the entire project, and thinking critically about how the use of AI could empower and advance the well-being of as many people as possible.

While principles are helpful to guide ethical practice, ascertaining how to maximise the benefits of AI and identifying the varied and unequal potential negative impacts of the technology requires engaging with diverse views and experiences to fully understand and anticipate the impacts of AI on society and to ensure that the ways in which it is developed and deployed reflect societal values. In particular, given the well-documented potential for AI to have inequitable impacts across society, it is important to engage and incorporate the views and interests of the most vulnerable groups. Children and young people are one such group who have so far been underrepresented in discussions of AI ethics.

As the field of AI ethics continues to expand, it is necessary and critical that the voices of children and young people are encouraged and heard. These voices are a critical piece of AI ethics work going forward. Next, we will explain why.

Why does this matter for children and young people?

While there are positive examples of AI being used to help better deliver public services and advance the well-being of individuals, there is unfortunately no shortage of examples of where the use of AI technologies has caused harm to people — including children and young people.

In 2020, due to the COVID-19 pandemic, all secondary education examinations were cancelled in the UK. In the absence of exam results, a solution was needed. While there were predicted grades available from teachers, there was concern that these may lead to inflated results due to over-optimistic or unrealistic estimations. Therefore Ofqual, the UK's Office of Qualifications and Examinations Regulation, decided to produce an algorithm that would determine the qualification grades for each student for that year. It was intended to moderate and standardise teachers' predictions of students' grades (Tennison, 2020). However, the algorithm resulted in skewed exam scores, and a pattern was detected where students from less-privileged schools were more likely to have their exam results downgraded, while those from private schools were more likely to receive the estimated grades given by their teachers (Bedingfield, 2020). This was in part because exam results for each school were moderated to reflect previous attainment levels for each school, and also because the size of the cohort played a significant role in the model's output — in schools with small class sizes (predominantly private schools), the algorithm could not be relied upon to moderate the results to the same extent as for larger cohorts (Bedingfield, 2020). There was a public outcry as it was discovered that higher exam scores were highly related to privately funded independent schools; thus, students from state schools were penalised. Ultimately, while the algorithm was intended

to address potential unfairness of relying on estimated grades, it, in fact, exacerbated existing inequalities in society leading to unfair outcomes. Following backlash from students and legal action on behalf of advocacy organisations, exam scores were reissued based on unmoderated teacher predictions. This example illustrates how algorithms that do not fully consider ethical and social implications can cause significant harm and discriminatory outcomes.

Another example in which an AI technology has caused harm is through AI-assisted chatbots. In 2018, the BBC conducted a study to test the effectiveness of chatbots in a mental health setting (White, 2018). After testing two chatbots, the researchers concluded that the applications failed to "properly handle children's reports of sexual abuse" even though this chatbot technology was designed with children in mind.

Unfortunately, there are countless examples across many sectors of ways in which AI technologies have caused harm, especially to children and young people. Children and young people have a unique set of needs, and it is important to note that if developed ethically and responsibly and with children's voices included and listened to, AI technologies could provide beneficial outcomes. For example, AI technologies have immense potential to improve the provision of public services in a variety of settings, such as education.

Within the education sector, there are several examples of how AI could be used to better support children, parents, guardians, and teachers. The use of translation tools to expand access to education for students across the globe is one way in which AI systems could provide benefits. AI systems can provide real-time translations into different languages as well as provide increased accessibility to the services for those with visual or hearing impairments, so that universal access to education is expanded.

There is also the potential for AI to be used to provide curated tutoring lessons for students' specific learning styles and ensure that they are not struggling with lessons. These are a few examples of how AI could be harnessed to improve the quality of education. However, in order to realise benefits such as these, AI technologies must account for children's needs and interests, which are informed by not only the potential risks but ethical principles — such as those outlined above. One way in which AI ethics principles have been framed with the unique needs of children placed at the forefront is a developing area of research called 'child-centred AI'.

Child-centred AI

Child-centred AI ensures that children are involved throughout all stages of the AI lifecycle in a meaningful and worthwhile way. A summary of the main components of child-centred AI can be found below:

- Helping children to make informed choices about their interactions with and uses of AI
- Enabling children and young people to play a role in discussions shaping future AI practices
- Ensuring the next generation of developers and policymakers are equipped with an understanding of the ethical considerations relating to AI and its uses
- Ensuring ethical mindsets of future developers and members of the tech industry

The United Nations Children's Fund (UNICEF) has been working on the topic of child-centred AI. In 2020, UNICEF and the Government of Finland co-authored a draft policy guidance entitled Policy Guidance on AI for Children (UNICEF, 2020). The draft contains an introduction to what is meant by the term AI and includes descriptions of the key opportunities and risks AI poses in the context of children's rights. UNICEF's nine requirements for child-centric AI are at the basis of this developing field of research. These are:

1. Support children's development and well-being
2. Ensure inclusion of and for children
3. Prioritize fairness and non-discrimination for children
4. Protect children's data and privacy
5. Ensure safety for children
6. Provide transparency, explainability, and accountability for children
7. Empower governments and businesses with knowledge of AI and children's rights
8. Prepare children for present and future developments in AI
9. Create an enabling environment

Putting child-centred AI into practice

Our team at the Alan Turing Institute was invited to test UNICEF's draft policy guidance and share our findings with the public about what worked and what did not. The goal of organisations participating in this programme was to improve child-centred AI moving forward. We interviewed 14 public sector organisations across the UK to gain perspectives on how they think about developing child-centred AI applications, their opinions on the UNICEF guidance and other data protection regulations, and how they wish to see children, young people, and parents and guardians involved in the design, development, and deployment of AI technologies that use their data.

Our main findings are summarised here and are discussed further in our full-length case study (Pauwels et al., 2021):

- Public sector organisations believe there are low rates of data literacy amongst the public.
- There is an overall lack of understanding and clarity surrounding the implementation of GDPR principles.
- There are many guidance documents being drafted on the topic of children's rights and

AI, and organisations are unsure which to use moving forward. Organisations wished to see synergies formed between existing and upcoming guidance documents.

- There is a desire to make the UNICEF *Policy Guidance on AI for Children* more actionable, to include more specific recommendations by sector, and to ensure the guidance is delivered in an age-accessible manner.
- Public sector organisations want to engage children and young people, but they stated that they did not know the best way to do this.

The findings from these interviews revealed public sector stakeholders' commitments to protecting children's rights and their enthusiasm to engage children in discussions relating to AI, but they also revealed many challenges associated with doing so. To illustrate this point, one interviewee stated:

There are lots and lots of ways and metrics that can be used to prove 'Ensure the inclusion of and for children' has happened without those children in the room actually being informed of what's going on. I'm not just talking about informed consent. I mean being fully appraised [sic] of the process and fully understanding.

It is clear that to address these challenges, children and young people must be involved in decision-making about the ways that AI is used in the public sector now and in the future. Our findings demonstrated that while public sector organisations wish to engage with children on these topics, they are not sure how to go about this in a meaningful way. In the next section, we will explore potential approaches to engaging children around AI ethics.

Approaches to engaging children with AI ethics

There are a number of reasons why an organisation developing or deploying AI might be motivated to engage with children or young people. These reasons in turn reflect different underpinning rationales, which can be normative, instrumental, and/or substantive (Fiorino, 1990; Wilsdon & Willis, 2004). First, a normative rationale leads to moral positions that suggest that if an organisation is developing or deploying an AI system that might impact on children they should engage with children as 'it's the right thing to do' (Wilsdon & Willis, 2004). Second, more practically minded approaches follow instrumental rationales, which view efforts to engage children as a means to achieve an organisation's own objectives (Wilsdon & Willis, 2004). Instrumental rationales might lead to a variety of potential approaches, including: efforts to build and maintain public trust in order to attract and retain customers; adopting ethical and transparent approaches to business to anticipate and respond to regulatory and policy developments; or efforts to demonstrate an ethical brand. However, following a purely instrumental rationale can lead to approaches that pay 'lip service' to public concerns through enacting purely cosmetic forms of engagement without genuine intentions to address concerns or reflect public values in an organisation's operation.

A final set of motivations are underpinned by substantive rationales that regard engagement as being aimed at creating wider positive outcomes across society.

From this point of view, citizens are seen as subjects, not objects, of the process. They work actively to shape decisions, rather than having their views canvassed by other actors to inform decisions that are then taken.
(Wilsdon & Willis, 2004, p. 39).

Following this approach, engagement with children offers opportunities to 'do things better' and maximise benefits not only for the organisation concerned but also for children

and wider society. This might lead to AI being developed in ways that not only reduce risks or potential harms of technologies, but that are also more appropriate and beneficial for children and young people. Here it is important to emphasise that engagement is not simply about avoiding or mitigating potential negative impacts but equally about maximising the benefits of AI.

When we think about these different rationales for engagement within educational contexts, they also lead to different approaches and priorities. An instrumental rationale would tend to emphasise more direct goals (e.g., increasing skills and knowledge relating to AI in order to prepare children for future careers in AI), while substantive rationales are more likely to emphasise indirect and less quantifiable outcomes. For example, a substantive rationale might underpin approaches to engagement that seek to engage children with discussions of AI ethics in order to enable them to understand the role and impact AI plays in their lives and their society and to equip them with the skills and understandings to be able to critique the ways that AI is designed, developed, and deployed. The main difference here is that rather than focusing on teaching children to equip them with skills to be the future AI workforce, we are also aiming to equip them with skills to be the future — and current — critical public, which is needed to hold AI systems and their developers to account.

The benefits of these substantive approaches are manifold. Firstly, children benefit from better understanding the role of AI in their lives (now and in the future) and by being able to critically engage with AI and make informed choices about this. Secondly, there are substantial benefits for the development and deployment of AI, since children's views and values need to be included and reflected in order to inform ethical practice. Quite simply, AI systems which impact — or have the potential to impact — children cannot be said to be developed or deployed ethically if children's experiences and

perspectives have not been reflected in the design and development processes. Thirdly, there are wider benefits for society through engaging with diverse stakeholders in relation to AI policy as well as in the design, development, and deployment of AI. Reflecting public values and interests in all these processes is essential to establish a social licence for AI (Aitken et al., 2020), which ensures that uses of AI reflect societal values and expectations.

Importantly, these approaches to engaging children with AI require going beyond one-way forms of communication and instead require engaging in dialogue with children. Previous studies in public engagement with science and technology have demonstrated the limitations of approaches aimed at gaining public trust through improving public understanding. Such approaches treat members of the public as "passive recipients of scientific knowledge" (Cunningham-Burley, 2006, p. 206), overlooking how members of the public critically assess, deconstruct, and evaluate claims to scientific knowledge in line with their own ideologies, experiences, and the contexts in which the information is received (Hagendijk & Irwin, 2006). Demonstrating technical competence or communicating the robustness of technical responses to ethical challenges will not automatically lead to public trust and support. Rather, technical approaches need to be combined with social responses that build relationships of trust through which claims to technical competence will be evaluated (Aitken et al., 2020). While scientific and technical expertise is important, "such expertise cannot resolve the moral and political aspects of policy-making" (Elstub et al., 2021) or ethical considerations relating to AI. As such, engagement and deliberation can play a role in establishing the trustworthiness of science and technology through efforts to address and reflect public values (Aitken et al., 2016; Wynne, 2006). Such deliberations do not require a detailed technical understanding of AI

technologies; an understanding of the contexts in which technologies will be applied and the lived experience of communities that may be impacted are valuable forms of expertise and knowledge within these deliberative processes (Aitken et al., 2021).

Recognising the importance of dialogue and deliberation, approaches to engaging children with AI should start by asking questions around what children currently know, and what they want to know. What are their concerns, interests, or priorities? What are the important issues in their lives to which AI may relate, or for which AI might have a positive or negative impact? A child-led approach to engagement, which does not begin with assumptions of what children already know, or what they *should* know about AI, is likely to be more fruitful in leading to discussions that can underpin ongoing engagement with AI ethics.

Conclusions and looking ahead

This paper has introduced the context of children's rights and AI, AI ethics principles, and the why and how of engagements with children surrounding the development of AI technologies that use their data. We have introduced principles that underpin the field of AI ethics and explained how these principles must be combined with meaningful participatory engagement to ensure the unique needs of children are met. UNICEF's nine child-centric principles have laid the foundation for child-centred AI, which places children's voices at the centre and ensures that the voices of children are taken on board throughout the entire AI lifecycle. Child-centred AI does not exist solely for harm reduction, but also to create new and beneficial approaches to the design, development, and deployment of AI technologies that involve children's data.

Child-centred AI and participatory engagement with children provide benefits for children by

equipping them with new understandings of the impacts that these technologies can have on their lives as well as the ability to critically engage with AI on a day-to-day basis. There are also possible benefits to the field of AI by providing a landscape that considers children's unique set of needs and circumstances, leading to better, more appropriate, sustainable, and beneficial use of technology.

While we advocate for expanding the focus of children's engagement with AI beyond the development of skills for the future AI workforce, it is also vital that those children who do later choose to follow careers in, or with, AI enter those careers with a sound understanding of the importance of ethics in AI and an appreciation of both the tremendous potential benefits of AI but also the risks. The future AI workforce needs a diverse mix of skills and expertise encompassing technical, social, ethical, legal, and policy dimensions. Engagement with children relating to AI must aim at addressing this interdisciplinarity and broad relevance of AI, recognising that AI is not purely a technical or scientific subject but one that touches on all aspects of our lives, and about which we should all have a voice.

Ultimately, realising the benefits of AI will require an engaged and critical public whose voices and experiences are taken on board in design, development, and deployment processes. Children and young people have so far been underrepresented in discussions of AI and AI ethics, but it is vital that their views and interests are taken on board to inform future approaches. This will be an important area of research and practice in the coming years.

In the Ethics Theme at the Alan Turing Institute, we are embarking on an exciting project to explore this further. Working in collaboration with the Scottish Children's Parliament and the Scottish AI Alliance, we are engaging children across Scotland in discussions about AI. This

research will explore what children currently know about AI and what they want to know about AI and how it is used; how children feel about the ways in which AI may be used to inform decisions about their lives (e.g., access to services); how they would like AI to be used in the future; what they think are the limits to how AI should be used in the future; and how children want to be involved in decision-making about future uses of AI.

We are excited about this next chapter in our research and to place children's voices at the heart of AI ethics.

References

Aitken, M., Ng, M., Horsfall, D., Coopamootoo, K.P., van Moorsel, A., & Elliott, K. (2021). In pursuit of socially-minded data-intensive innovation in banking: A focus group study of public expectations of digital innovation in banking. *Technology in Society, 66*(7), Article 101666. https://doi.org/10.1016/j.techsoc.2021.101666

Aitken, M., Toreini, E., Carmichael, P., Coopamootoo, K., Elliott, K., & van Moorsel, A. (2020). Establishing a social licence for Financial Technology: Reflections on the role of the private sector in pursuing ethical data practices. *Big Data & Society, 7*(1), Article 2053951720908892. https://doi.org/10.1177/2053951720908892

Aitken, M., Cunningham-Burley, S., & Pagliari, C. (2016). Moving from trust to trustworthiness: Experiences of public engagement in the Scottish Health Informatics Programme. *Science and Public Policy, 43*(5), 713–723. https://doi.org/10.1093/scipol/scv075

Barassi, V. (2018, September 18). *Home life data and children's privacy*. Child Data Citizen. Retrieved February 27, 2022, from http://childdatacitizen.com/cdc/wp-content/uploads/2018/09/'HOME-LIFE-DATA'-AND-CHILDREN'S-PRIVACY-1.pdf

Bedingfield, W. (2020, August 19). *Everything that went wrong with the botched A-Levels algorithm*. Wired. https://www.wired.co.uk/article/alevel-exam-algorithm

Berman, G., & Albright, K. (2017). Children and the data cycle. *Innocenti Working Papers*. https://doi.org/10.18356/7d33bc21-en

Cunningham-Burley, S. (2006). Public knowledge and public trust. *Public Health Genomics, 9*(3), 204–210. https://doi.org/10.1159/000092658

Data Protection Working Party (2009, February 11). *Opinion 2/2009 on the protection of children's personal data (General Guidelines and the special case of schools)*. https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2009/wp160_en.pdf

Druga, S., Williams, R., Breazeal, C., & Resnick, M. (2017). "Hey Google is it ok if I eat you?": Initial explorations in child–agent interaction. In *Proceedings of the 2017 conference on interaction design and children* (pp. 595–600). Association for Computing Machinery. https://doi.org/10.1145/3078072.3084330

Elstub, S., Thompson, R., Escobar, O., Hollinghurst, J., Grimes, D., Aitken, M., McKeon, A., Jones, K., Waud, A., & Sethi, N. (2021). The resilience of pandemic digital deliberation: An analysis of online synchronous forums. *Javnost-The Public, 28*(3), 237–255. https://doi.org/10.1080/13183222.2021.1969616

Fiorino, D.J. (1990). Citizen participation and environmental risk: A survey of institutional mechanisms. *Science, Technology, & Human Values, 15*(2), 226–243. https://doi.org/10.1177/016224399001500204

Fjeld, J., Hilligoss, H., Achten, N., Daniel, M.L., Feldman, J., & Kagay, S. (2020). *Principled artificial intelligence: A map of ethical and rights-based approaches*. Berkman Klein Centre for Internet & Society. https://dash.harvard.edu/bitstream/handle/1/42160420/HLS%20White%20Paper%20Final_v3.pdf?sequence=1&isAllowed=y

Floridi, L. (2021). Translating principles into practices of digital ethics: Five risks of being unethical. In: Floridi, L. (Ed.), *Ethics, Governance, and Policies in Artificial Intelligence* (pp. 81–90). Springer. https://doi.org/10.1007/978-3-030-81907-1_6

Gibbs, S. (2015, March 13). *Privacy fears over 'smart' Barbie that can listen to your kids*. The Guardian. https://www.theguardian.com/technology/2015/mar/13/smart-barbie-that-can-listen-to-your-kids-privacy-fears-mattel

Hagendijk, R., & Irwin, A. (2006). Public deliberation and governance: Engaging with science and technology in contemporary Europe. *Minerva, 44*(2), 167–184. https://doi.org/10.1007/s11024-006-0012-x

Harris, R. (2017, December 27). *72M data points collected on children in spite of COPPA*. App Developer Magazine. https://appdevelopermagazine.com/72m-data-points-collected-on-children-in-spite-of-coppa/

Howley, R. (2019). *The effects of artificial intelligence on the youth*. (Publication No. 22622916) [Master's thesis, Utica College]. ProQuest Dissertations Publishing. https://www.proquest.com/openview/d84a86cf41d62229ab84860328e20c92/1?pq-origsite=gscholar&cbl=18750&diss=y

Kadho Inc. (2018, May 18). *Kid's privacy in the age of artificial intelligence*. Medium. https://medium.com/@KadhoInc/kids-privacy-in-the-age-of-artificial-intelligence-4c113e642124

Leslie, D. (2019). *Understanding artificial intelligence ethics and safety: A guide for the responsible design and implementation of AI systems in the public sector*. The Alan Turing Institute. https://doi.org/10.5281/zenodo.3240529

Lieber, C. (2018, December 5). *Big tech has your kid's data – and you probably gave it to them*. Vox. https://www.vox.com/the-goods/2018/12/5/18128066/children-data-surveillance-amazon-facebook-google-apple

Lupton, D. & Williamson, B. (2017). The datafied child: The dataveillance of children and implications for their rights. *New Media & Society, 19*(5), 780–794. http://journals.sagepub.com/doi/full/10.1177/1461444816686328

Macenaite, M. (2017). From universal towards child-specific protection of the right to privacy online: Dilemmas in the EU General Data Protection Regulation. *New Media & Society, 19*(5), 765–779. https://doi.org/10.1177/1461444816686327

Mittelstadt, B.D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society, 3*(2), Article 2053951716679679. https://doi.org/10.1177/2053951716679679

Montgomery, K.C., & Chester, J. (2015). Data Protection for youth in the digital age: Developing a rights-based global framework. *European Data*

Protection Law Review, 1(4), 277–291. https://doi.org/10.21552/edpl/2015/4/6

Pauwels, E., Briggs, M., Aitken, M., & Leslie, D. (2021). *Case study: Understanding AI ethics and safety for children: The Alan Turing Institute*. UNICEF. https://www.unicef.org/globalinsight/media/2321/file

Segal, E. (2021). *A crisis is averted after Amazon's Alexa tells child about lethal viral challenge*. Forbes. https://www.forbes.com/sites/edwardsegal/2021/12/28/a-crisis-is-averted-after-amazons-alexa-tells-child-about-lethal-viral-challenge/?sh=4c32eca139b6

Taylor, E., & Michael, K. (2017, June 29). *Smart toys that are the stuff of nightmares*. IEEE Technology and Society. Retrieved February 27, 2022, from https://technologyandsociety.org/smart-toys-that-are-the-stuff-of-nightmares/

Tennison, J. (2020, August 16). *How does Ofqual's grading algorithm work?* RPubs. https://rpubs.com/JeniT/ofqual-algorithm

UNICEF (2020, September). *Policy guidance on AI for children*. UNICEF and the Ministry of Foreign Affairs of Finland. https://www.unicef.org/globalinsight/media/1171/file/UNICEF-Global-Insight-policy-guidance-AI-children-draft-1.0-2020.pdf

White, G. (2018, December 11). *Child advice chatbots fail to spot sexual abuse*. BBC News. https://www.bbc.co.uk/news/technology-46507900

Williams, R., Park, H.W., & Breazeal, C. (2019). A is for Artificial Intelligence: The impact of Artificial Intelligence activities on young children's perceptions of robots. In *CHI Conference on Human Factors in Computing Systems Proceedings* (pp. 1–11). Association for Computing Machinery. https://doi.org/10.1145/3290605.3300677

Wilsdon, J., & Willis, R. (2004). *See-through science: Why public engagement needs to move upstream*. Demos.

Wynne, B. (2006). Public engagement as a means of restoring public trust in science – Hitting the notes, but missing the music? *Community Genetics, 9*(3), 211–220. https://doi.org/10.1159/000092

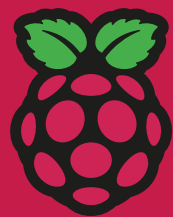### Mhairi Aitken (The Alan Turing Institute, UK)

Dr. Mhairi Aitken is an Ethics Fellow in the Public Policy Programme at The Alan Turing Institute, and an Honorary Senior Fellow at ACHEEV, University of Wollongong in Australia. She is a Sociologist whose research examines the social and ethical dimensions of digital innovation, particularly relating to uses of data and AI. She has also held roles as a Public Engagement Research Fellow in both the Farr Institute of Health Informatics Research and the Scottish Health Informatics Programme (SHIP).

Morgan Briggs (The Alan Turing Institute, UK)

Morgan Briggs currently serves as the Research Associate for Data Science and Ethics within the Public Policy Programme. Morgan works on a variety of projects at The Alan Turing Institute, including continued, ongoing work on AI explainability. In addition to her role at The Alan Turing Institute, Morgan serves as an Associate Teaching Fellow for the Global Leadership Initiative at the University of Oxford.