Data Autonomy in the Age of AI

Designing Autonomy-Supportive Data Tools for Children and Families

Ge Wang

Jesus College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2023

Acknowledgements

In the summer of 2018, when I first joined the KOALA project as a summer intern, I shared my genuine interest in AI, and expressed my desire to make a difference for the younger generation. Little did I know that the intersection of children's autonomy and AI would become the core of my thesis. This journey would not have been possible without the unwavering support I received from supervisors, mentors, colleagues, friends, and family. Their kindness, encouragement, and thoughtfulness have been instrumental at every step. It has been a true joy and privilege to have such backing, and I will try to pass it on to others as I move forward in my own career.

To Sir Nigel Shadbolt. My foremost thanks go to Sir Nigel Shadbolt. Throughout my DPhil journey, you have been a beacon of inspiration, consistently reminding me of the broader vision. Your astute guidance, whether on experimental intricacies or the nuances of conference reimbursements, has been invaluable. Your thorough reviews and feedback on my publications have sharpened my work, and I'm deeply grateful for your relentless efforts to secure funding, even when I delved into new domains. Your support has truly been remarkable. You have always emphasized the societal implications of my research, pushing me to ensure it was both innovative and transformative, and what our research truly means in the grander scheme of things – a sentiment I deeply hope to carry forward and pass on. It has been a true pleasure.

To Jun Zhao. I still remember the afternoon we first met, outside Gail's bakery in Jericho. I was expressing my interest in child-related work. Being just an undergrad student back then, I never anticipated the warmth and enthusiasm with which you'd welcome me into the fold. I remembered you once told me that guiding a DPhil student is like raising a kid, a sentiment that resonates deeply now. You were my guiding hand, handholding me every step of the way, from my first study, my first paper, my first research talk, to my first paper award. You didn't just show me the technicalities; you taught me how to think, how to ask questions, and how to navigate the intricate world of research with passion and purpose. For all this and more, I remain deeply grateful.

To Max Van Kleek. Throughout my DPhil journey, you've been an unwavering source of inspiration. Our brainstorming sessions were truly enlightening, nudging me to think beyond conventional boundaries. I still remember the first CSCW paper we did together, those caffeine-driven 2am chats, paired with the detailed Miro mindmaps we crafted together. You always provide critical advice that leads to those "aha!" moments, and your ceaseless stream of fresh ideas keeps our momentum alive. The time and effort we invested, the challenges we overcame together, all of it made working with you not just an educational experience, but also one filled with fun and profound inspiration.

To all my mentor, colleagues, and collaborators at Oxford. This work has been shaped by countless conversations with fellow researchers. A special thank you to Professors Reuben Binns and Victoria Nash, who not only served as my examiners for both my transfer and confirmation but also offered invaluable suggestions throughout. My gratitude extends to Professor Standa Zivny for his role as my college advisor, consistently providing me with academic and career guidance. Additionally, great gratitude to my collaborators and colleagues who made my time both memorable and productive: Anirudh Ekambaranathan, Claudine Tinsman, Daniel Omeiza, Hayoun Noh, Jake Stein, Konrad Kollnig, Lize Alberts, Martin Kraemer, Petr Slovak, Samantha-Kaye Johnston, Sarah Aldaweesh, Siddhartha Datta, Thomas Serban von Davier, Ulrik Lyngs, William Seymour, and Zhilin Zhang. And to all the students who collaborated with me on various research projects: your skills, enthusiasm, and creativity have not only amazed me but have also elevated our work tremendously.

To all my mentors, colleagues, and collaborators from beyond Oxford. I extend my deep gratitude for the doors you've opened for me in both academia and industry. I'm especially thankful to Kaiwen Sun; our shared research interests have sparked countless enlightening discussions. A sincere thank you to Grace C. Lin for her warm hospitality during my visit to MIT. My appreciation also goes to Ayca Atabey, Kruakae Pothong, Zhen Bai, and Jason C. Yip for joining hands with me in organizing the CCAI Workshop and Special Issue. I am profoundly grateful for the financial backing I received from Jesus College, the Oxford Martin School EWADA Programme, and the University of Oxford. Moreover, I am grateful to the ACM SIGCHI community for their invaluable research insights. Last but by no means least, my heartfelt thanks to all the children, parents, teachers, schools, and participants who have been pillars of support for my research. Your enthusiasm and unwavering backing are the bedrock upon which this research stands.

To my friends and family. Thank you to all who have been there in my life. Thank you to Jinsong, you have given me all the emotional support I needed. We all need a good pep talk every now and then, and you're the MVP in that department. And thank you to my little furry friends: Huda, Neinei, Heeya, and Nuonuo, who may never understand my words. Your nightly purrs and the enthusiastic hugs you give every time I come home are truly heartwarming. I will forever cherish these moments with you. My heartfelt gratitude goes out to my parents for their boundless love, unwavering kindness, and the steadfast support they've provided throughout my journey. You've always encouraged me to have faith in myself and to believe in my potential. Even when I was thousands of miles away from home since the age of 15, this conviction has been my unwavering anchor, guiding me through every challenge I faced. And a special thank you to my grandparents, especially my grandpa, my very first academic tutor, my personal goal-setting guru, and the one who taught me to stand firm in what I believe in.
春晖寸草，难以言报。椿萱并茂，棠棣同馨。

Abstract

The age of AI is a rapidly evolving and complex space for children. As children increasingly interact with AI-based apps, services and platforms, their data is being increasingly tracked, harvested, aggregated, analysed and exploited in multiple ways that include behavioural engineering and monetisation. Central to such datafication is online service providers' ability to analyse user data to infer personal attributes, subtly manipulating interests and beliefs through micro-targeting and opinion shaping. This can alter the way children perceive and interact with the world, undermining their autonomy. Yet, this datafication often unfolds behind the scenes in apps and services, remaining less noticed and discussed compared to the more straightforward data privacy issues like direct data collection or disclosure.

On the other hand, children are often seen as less capable of navigating the intricacies of online life, with parents and guardians presumed to possess greater expertise to steer their children through the digital world. However, the rapid evolution of AI technology and online trends has outpaced parents' ability to keep up. As they adapt to platforms like Snapchat or YouTube, children may already move to the next trend, a shift accelerated by rapid datafication that heightens the challenge of effectively guiding children online. Consequently, there's a mounting call for a child-centred approach, which shifts from just protecting or limiting children with parents in charge, to actively guiding and empowering children to take a leading role. In this shift towards a child-centred approach, there's growing consensus on fostering children's autonomy in the digital space, encompassing the development of their understanding, values, self-determination, and self-identity.

Given that data is the cornerstone of AI-based platforms' vast influence, this thesis uniquely focuses on the key concept of *data autonomy* for children. This exploration follows a structured four-step methodology: 1) Landscape analysis to comprehend the present scope of AI-based platforms for children and the prevalent challenges they encounter; 2) Conceptual review to elucidate the meaning of autonomy for children in the digital realm; 3) Empirical investigation focusing on children's perceptions, needs, and obstacles concerning data autonomy; and 4) Technical evaluation to assess the impact of technical interventions on children's sense of data autonomy.

Synthesising the research presented in this thesis, we propose the pivotal concept of *data autonomy* for children in the age of AI, aiming to address their online wellbeing from a unique data perspective. This work not only lays the foundation for future research on data autonomy as a novel research agenda, but also prompts a rethinking of existing data governance structures towards a more ethical data landscape.

Contents

List of Figures

List of Tables

Glossary

Children . . . According to the United Nations Convention on the Rights of the Child (UNCRC), a 'child' is defined as anyone under the age of 18 *General comment No. 25 (2021) on children's rights in relation to the digital environment* 2020. In the scope of this thesis, we primarily focused on children in their 'early adolescence', ranging in age from 7 to 14.

Child Rights . According to the General Comment 25 by UNCRC, child rights in the digital domain refers to 1. Non-discrimination: Ensure all children have equal access to meaningful digital experiences; 2. Best interests of the child: Prioritize children's welfare in all decisions and actions affecting them; 3. Right to life and development: Protect children from digital threats like violent content, harassment, and exploitation; emphasize technology's influence during early childhood and adolescence and educate caregivers on safe usage; 4. Respect for the child's views: Promote children's expression on digital platforms, integrate their input into policies, and ensure service providers respect their privacy and thought freedom.

Digital Space The virtual environment where electronic data is stored, managed, and exchanged. It encompasses online platforms, websites, databases, and other venues where digital activities occur. In this thesis, we also used other similar terms including digital realm, digital environments.

Online . . . . . Web-based or app-based venues that enable users to engage in various digital activities, such as socializing, shopping, learning, and more. Examples include social media sites, e-commerce websites, and online educational platforms.

AI . . . . Machine-based systems designed to perform tasks that typically require human intelligence. Using a set of human-defined objectives, these systems can make predictions, recommendations, or decisions that influence real or virtual environments.

Social Media . Web-based or app-based platforms that facilitate the creation and sharing of content, ideas, career interests, and other forms of expression among communities and networks. Examples include Facebook, Twitter, and Instagram. In the scope of this thesis, for 'social media', we are typically referencing the major centralized platforms that currently dominate the market. Emerging decentralized social media platforms like Mastodon, Blockstack, and Steemit are not included in the scope of this thesis.

Algorithmic . Pertaining to a set of rules or procedures that are used for problem-solving or to perform a task. In digital contexts, it usually refers to the logic used by computers to process data and make decisions.

Datafication . The process of transforming various aspects of modern life into data, which can then be stored, analysed, and utilised. In the digital age, this often relates to how human behaviors, interests, and interactions are converted into quantifiable data points, often for the purpose of analysis, prediction, and monetisation.

Autonomy . . Autonomy refers to the ability and freedom to make one's own choices and decisions without external coercion or interference. It emphasizes self-governance, independence, and personal agency. In the scope of this thesis, we focus on *personal autonomy* as the ability of individuals to shape their choices according to their own understanding or values, without being manipulated or controlled by others.

Agency . . . . Agency is the ability of individuals to act, decide, and influence outcomes, even amidst external constraints. Unlike *autonomy*, which centers on freedom from external impositions, agency stresses the potential to enact choices. One can exhibit agency without absolute autonomy, as choices may be influenced by both internal and external factors. Though agency is vital for understanding individual behaviors, this thesis prioritizes autonomy to emphasize acting in alignment with one's true beliefs and values, unswayed by outside forces.

Control . . . . Throughout this thesis, the term "control" is occasionally used to make the research more accessible to a wider audience and study participants. For instance, *"How do you want to feel more in control with your data online?"*.

Statement of Authorship

This thesis is the result of my original research work conducted under the supervision of Sir Nigel Shadbolt and Jun Zhao at the Department of Computer Science, University of Oxford. Previous forms of much of the work in this thesis have been published as co-authored papers. On each of these papers I was the principle author, and contributed the majority of the writing that appears in them. This also applies to the rest of the research process, with me being either wholly or primarily responsible for planning studies, collecting data, and performing analysis. Where content in this thesis has previously appeared in one of these papers, this is clearly marked at the beginning of the chapter.

ge wang

Ge Wang
22/09/2023

1 Introduction

Contents

1.1 Motivation

AI-based platforms, which are digital services or environments that seamlessly integrate artificial intelligence technologies and utilise complex algorithms to analyse vast data sets, are assuming various roles in the digital ecosystems of children, being embedded in the connected toys, apps and services that children interact with daily (UNICEF 2020; *Generation AI: Establishing Global Standards for Children and AI* 2019). Such platforms offer many advantages to children, such as personalised teaching and learning found in intelligent tutoring systems (Ramachandran, Huang, and Scassellati 2017; Fadhil and Villafiorita 2017), or online content monitoring and filtering algorithms that proactively identify potentially harmful content or contexts before they are experienced (Tahir et al. 2019; Kumaresamoorthy and Firdhous

2018). Algorithms in games and entertainment services provide personalised content recommendations (Frikha, Zlitni, and Bouassida 2020), while social robots power interactive characters that are engaging and human-like (Castellano et al. 2014; Kochigami et al. 2015).

Going forward, AI-based platforms will, in all likelihood, become altogether even more pervasive in children's everyday lives simply due to their sheer capabilities in creating compelling, adaptive, and personal user experiences (*The state of AI in 2020* 2020). Yet, despite its enormous potential, the algorithmic nature of AI comes with new kinds of risks, some of which are of particular concern for children. Examples include potential bias against certain groups (Hutchinson and M. Mitchell 2019; Corbett-Davies et al. 2017), which may cause systems to treat those in different socioeconomic or ethnic groups differently and bring psychological or social impacts to children disproportionately in their formative years (UNICEF 2020; *Generation AI: Establishing Global Standards for Children and AI* 2019). Similarly, inscrutability and unpredictability could inadvertently cause children to be exposed to harms in content filtering systems in ways that are difficult to anticipate or predict, such as those crafted by malicious adversaries (*Generation AI: Establishing Global Standards for Children and AI* 2019). Moreover, children are among those at greatest risk of privacy-related harms due to the fact that data collected about them could affect them throughout the lifetime they have yet to live (*Age appropriate design code* 2020).

Datafication, a process which involves recording, tracking, aggregating, analysing, and capitalising on user data, is prominently evident in the way AI-based platforms capture, monitor, and amalgamate children's data and activities. By analysing and exploiting children's data, these services have the capacity to influence and potentially manipulate children's online behaviour, engagement, or content consumption (Mascheroni 2020; Mejias and Couldry 2019; Zuboff 2019). This contributes to the emerging phenomenon termed as the datafied childhood (Mascheroni 2020). At the core of this *datafication* is platforms' ability to make *data inferences* about children, by analysing their data, supported by algorithms, with the aim to evaluate certain personal aspects relating to a natural person, including predicting their performance at work, economic situation, health, preferences, interests, reliability, behavior, location, or movements (Mascheroni 2020; *What is automated individual decision-making and profiling?* 2020a). In particular, such practice could be utilised to predict aspects concerning natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements (*What is automated individual decision-making and profiling?* 2020a). The pressing concern is that as algorithms hungrily consume and exploit user data, AI-based platforms gain an unprecedented power to fine-tune and manipulate the individual nuances of users' beliefs, interests, and actions. Such meticulous harnessing of personal traits not only paves the way for micro-targeting but can also subtly nudges users towards specific opinions and fosters an escalating dependence on these platforms. The ramifications of this are immense, potentially reshaping the very fabric of how our young ones perceive and interact with the world around them.

Meanwhile, children were traditionally seen as less capable of navigating the intricacies of online life, with parents and guardians presumed to possess greater expertise to steer their children through the digital world. This position is reflected in most existing measures focusing on creating barriers and protections (G. Wang, J. Zhao, et al. 2021). These often involve parents limiting the kind of content children can see, setting rules for how much screen time they get, or keeping a close watch on their online activities. However, with the rapid evolution of AI technology and online trends, this parental-oriented protective approach is increasingly seen to be less effective. Children have grown up in the digital era in a way that interacting with digital environment is almost second nature to them, and for most of them, their parents could possess far less knowledge in this domain (Fletcher and Blair 2014; Shifflet-Chila et al. 2016). The rapid evolution of online trends has made it even harder for parents to keep up. By the time they adapt to platforms like Snapchat or YouTube, children may already move to the next trend. Meanwhile, the rising datafication on AI-based platforms complicates parents' efforts to monitor and guide their children effectively (Mascheroni 2020). In response to this, recent work increasingly calls for a **child-centred** approach (Danby 2017; Vidal-Hall, Flewitt, and Wyse 2020; G. Wang, Sun, et al. 2023), which shifts from just protecting or limiting children with parents and caretakers in charge, to actively guiding and empowering children to take a leading role. Amidst this shift towards a child-centred approach, there is a growing recognition on the importance of fostering children's autonomy in the digital space, which includes developing their understanding, values, self-determination, and self-identity (Danby 2017; Vidal-Hall, Flewitt, and Wyse 2020; G. Wang, Sun, et al. 2023).

The intertwined dynamics of datafication practices and traditional parent-oriented approaches, though seemingly designed for protection, are challenging children's online opportunities and autonomy in several ways. Continuous surveillance of children's behaviors, preferences, and movements pigeonholes them into restrictive data-defined categories, potentially skewing perceptions and interactions. Further, parent-oriented methods, largely rooted in monitoring and restrictions, curtail children's chances to learn, build defenses, and develop autonomy in the face of such intricate datafication practices. In our rapidly advancing age of AI, to have **data autonomy** is thus paramount. It not only empowers individuals with control over their personal information but also fosters a sense of responsibility and understanding about the implications of shared data. For children, in particular, achieving data autonomy means navigating the digital world with confidence, self-awareness, and an understanding of their rights, thus preparing them for a future where data is an integral part of personal identity and agency.

In the age of AI, **how to take a child-centred approach to support children's autonomy and development within the prevalent datafication landscape** presents several complex open challenges:

- **Firstly**, the current comprehension of the AI landscape as it relates to children is insufficient, especially when considering the intricate datafication practices involving them.

- **Secondly**, there's a discernible gap in understanding children's viewpoints, requirements, and the challenges they encounter in preserving their autonomy over data on these platforms.

- **Lastly**, the significance of data autonomy for children has been somewhat overlooked in much of the existing research. Data autonomy is not merely about control; it's about granting children the agency to determine how their personal data is used, accessed, and shared. In a world dominated by digital interactions, ensuring children have such autonomy becomes paramount. This agency affords them not just protection, but also the empowerment and confidence to navigate the digital world responsibly, understanding the weight and worth of their personal information.

1.2 Thesis Questions

In this thesis, we adopt a unique and distinct perspective on children's well-being in the age of AI, one that diverges fundamentally from previous approaches. Central to our argument is the pivotal role of data in shaping children's online experiences. Recognising that data is the linchpin behind AI platforms' vast influence, **we seek to empower children to take control of their own data.** Motivated by the challenges listed above, this thesis seeks to answer fundamental questions about the nature of data autonomy for children in the age of AI. Specifically, we answer the following thesis questions (TQ):

- TQ1: In the current AI landscape, what insights can be gathered about children's understanding, perceptions, needs, and the obstacles they face regarding having autonomy over their data? (Chapter 4 & Chapter 5)

- TQ2: How can we take a child-centred design approach to develop technical interventions for children; and how might these technical interventions influence, support, and enhance children's autonomy over their data within AI-based platforms? (Chapter 6 & Chapter 7)

- TQ3: With these findings in mind, how can we define *data autonomy* for children in the age of AI, and what open challenges remain to be addressed? (Chapter 8)

1.3 Contributions

Through answering these questions, this thesis makes a number of direct contributions relevant to the emerging discipline of 'child centred computing' (Hourcade et al. 2016; G. Wang, Sun, et al. 2023) as a sub-field of Human Centred Computing (HCC). In summary, the contributions of this research are:

A Landscape Analysis of Existing AI-based Platforms for Children - In Chapter 2, we embarked on a comprehensive analysis of the current landscape of AI-based platforms designed for children, pinpointing the predominant challenges they face in this realm. Additionally, we scrutinised existing tools aimed at safeguarding children within these AI-based platforms, identifying the prevailing approaches adopted in recent research. This investigation brought to light a critical research gap, fostering a growing emphasis on the necessity of designing initiatives that encourage and support children's autonomy in the digital realm.

A Conceptual Understanding of Autonomy in the Digital Space - In Chapter 3, we delve into how the concept of autonomy has been acknowledged, conceptualised, and facilitated by current research and tools. This analysis provided a comprehensive depiction of what autonomy embodies for children in the digital space, pinpointing existing gaps in the conceptualisation and research surrounding children's autonomy. This laid down a solid groundwork for our further exploration into the nuances of children's data autonomy within AI-based platforms.

An Empirical Investigation on Children's Perceptions, Needs and Obstacles Regarding Autonomy of Their Data - In Chapter 4 and Chapter 5, we embarked on a detailed series of empirical studies to delve deeply into the children's current perceptions, expectations, and the challenges they encountered concerning data autonomy within AI-based platforms. These insights pave the way for a pioneering understanding, fostering a blueprint for crafting solutions that enhance the sense of autonomy for children regarding their data.

Technical Prototypes as Design Interventions for Children's Data Autonomy - In Chapter 6 and Chapter 7, we designed, developed and critically assessed two technical prototypes with the goal of bolstering children's autonomy concerning their data within AI-based platforms. This process yielded vital insights, offering a robust foundation and guiding principles for future researchers and designers in crafting more child-centred and autonomy-enhancing strategies, revolutionizing the approach to safeguarding children's autonomy in the digital realm.

Data Autonomy for Children in the Age of AI as New Research Agenda - Expanding upon the collective insights gathered, this thesis proposes a nuanced conceptual framework that delineates what *data autonomy* entails for children in the age of AI, positioning itself as a seminal work addressing this vital yet underexplored concept. We elucidate how data autonomy is contextualised within the intricate dynamics of children and families, and pinpoint the open challenges that mark the frontier of this burgeoning research agenda, paving a pathway for further in-depth exploration and discourse in this pivotal area.

1.4 Thesis Outline

Chapter 2 lays out the prior research this thesis expands upon. First, we presented an overview of AI-based platforms designed for children and the current standards and policies on ethical AI, including a comparison and critique of the ways in which a child's best interests are currently taken into account. Second, we delved into the concept of *datafication* and its potential negative implications for children's wellbeing and autonomy. Finally, we presented an analysis of existing practices of digital tools for children to navigate datafication. Through this comparison and analysis of existing work, this chapter highlights the pressing need for a child-centred approach, emphasizing the significance of promoting children's data autonomy within AI-based platforms.

Chapter 3 examines the extent to which children's autonomy on digital platforms has been recognised and supported by existing research and designs. By amalgamating a series of design mechanisms highlighted in existing research efforts with current conceptualisations, we delved deeply into the multifaceted concept of autonomy for children in the digital realm. The chapter culminates in proposing a nuanced working definition of data autonomy, outlining potential facets encompassed in children's autonomy with data, a concept which we further scrutinise and contemplate through empirical investigations and technical evaluations in subsequent sections.

Chapter 4 explores how children perceive and understand the datafication practices on their AI-based platforms. Through semi-structured interviews, we found that not only do children already possess rudimentary conceptual understanding of datafication, but they also demonstrate a significant willingness to take action to shape it to their desire.

Chapter 5 takes this desire for autonomy from children more deeply into design considerations. Through co-design sessions with children, we explored children's preferences for coping with everyday datafication practices. Our findings offer pivotal insights into designing mechanisms that support children's autonomy in navigating datafication practices, especially through developmentally-appropriate transparency mechanisms, repositioning children as active participants, and emphasising care and respect in the design.

Chapter 6 details the design, development and evaluation of *KOALA Hero Toolkit*, a multi-component hybrid toolkit designed to support children and their parents to work together to gain more understanding about datafication risks online, particularly those associated with the use of mobile apps. The user study results indicated that families engaged in deeper critical reflection about datafication concerns, and engaged in more democratic family dynamics. This work demonstrated that a child-centred approach might be more effective than the traditional parent-guided method.

Chapter 7 details the design, development and evaluation of *CHAITok*, an Android mobile application that enables children to view, manage and control datafication practices on their social media platforms. The results of user studies showed how the app, with its novel affordances, prompted children to become aware of, reevaluate, and establish their own control strategies on data. This chapter offers crucial insights into the expectations of children regarding data autonomy online, and provides vital insights into how we can empower children's data autonomy through a socio-technical journey.

Chapter 8 amalgamates the insights gleaned from the preceding chapters, offering a comprehensive discussion on the collective contributions of this thesis. It revisits previously gathered findings, carefully sculpting a nuanced definition of what *data autonomy* signifies for children in the age of AI, and discusses open challenges and opportunities for future investigations.

1.5 Publications

The following publications arose from this thesis. First authorship implies efforts for research planning, data collection, data analysis, and write-up. Where content of a chapter has previously been published, this is highlighted at the start of the chapter.

1. **Ge Wang**, Jun Zhao, Max Van Kleek, and Nigel Shadbolt. 2022. Informing Age-Appropriate AI: Examining Principles and Practices of AI for Children. In Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22). Association for Computing Machinery, New York, NY, USA, Article 536, 1–29. `https://doi.org/10.1145/3491102.3502057`
   Best Paper Honourable Mention Award

2. **Ge Wang**, Jun Zhao, Max Van Kleek, and Nigel Shadbolt. 2021. Protection or Punishment? Relating the Design Space of Parental Control Apps and Perceptions about Them to Support Parenting for Online Safety. Proc.

ACM Hum.-Comput. Interact. 5, CSCW2, Article 343 (October 2021), 26 pages. https://doi.org/10.1145/3476084. Also presented at Symposium on Usable Privacy and Security (SOUPS'22).

3. **Ge Wang**, Jun Zhao, Max Van Kleek, and Nigel Shadbolt. 2022. 'Don't make assumptions about me!': Understanding Children's Perception of Datafication Online. Proc. ACM Hum.-Comput. Interact. 6, CSCW2, Article 419 (November 2022), 24 pages. https://doi.org/10.1145/3555144
Best Paper Honourable Mention Award **Impact Recognition Award**

4. **Ge Wang**, Jun Zhao, Max Van Kleek, and Nigel Shadbolt. 2023. 12 Ways to Empower: Designing for Children's Digital Autonomy. In Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23). Association for Computing Machinery, New York, NY, USA, Article 91, 1–27. https://doi.org/10.1145/3544548.3580935
Best Paper Honourable Mention Award

5. **Ge Wang**, Jun Zhao, Max Van Kleek, and Nigel Shadbolt. 2023. 'Treat me as your friend, not a number in your database': Co-designing with Children to Cope with Datafication Online. In Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23). Association for Computing Machinery, New York, NY, USA, Article 95, 1–21. https://doi.org/10.1145/3544548.3580933 **Best Paper Award**

6. **Ge Wang**, Jun Zhao, Konrad Kollnig, Adrien Zier, Blanche Duron, Zhilin Zhang, Max Van Kleek, and Nigel Shadbolt. 2024. KOALA Hero Toolkit: A New Approach to Inform Families of Mobile Data Risks. In Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (CHI '24).

7. **Ge Wang**, Jun Zhao, Samantha-Kaye Johnston, Zhilin Zhang, Max Van Kleek, and Nigel Shadbolt. 2024. CHAITok: A Proof-of-Concept System

Introducing Data Autonomy for Children. In Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (CHI '24).
Best Paper Honourable Mention Award

8. **Ge Wang**, Jun Zhao, Max Van Kleek, and Nigel Shadbolt. 2024. Challenges and Opportunities in Translating Ethical AI Principles into Practice for Children. Nature Machine Intelligence, pp.1-6. `https://doi.org/10.1038/s42256-024-00805-x`

1.6 Statement of Positionality

I acknowledge that, for qualitative research, the research perspectives and approaches are shaped by our own experiences and positionality. Living and working in the UK, my primary research focus is on human-centered AI and its impact on children's experiences with data-driven AI algorithmic decision-making. My engagement with children, parents, educators, and AI practitioners, especially on the digital impact on childhood, shapes the foundation of my work. At the heart of my research lies a critical and reflective stance on children's use of AI-based social media, a subject of significant attention and debate. This perspective is not born out of a general distrust of technology; rather, it is fueled by a belief in the potential of AI to improve lives and drive societal progress. However, through comprehensive research and collaboration with both the academic and professional communities in education and technology, I have come to adopt a cautious approach towards the unmoderated interaction of children with social media platforms.

This positionality is pivotal in highlighting the potentially harmful effects of AI-based social media on children. While these platforms offer avenues for learning and socialization, their unregulated usage can significantly impact the mental and emotional well-being of young users. Research has consistently shown correlations between excessive social media use and issues such as decreased self-esteem, cyberbullying, anxiety, and depression among children and adolescents. Furthermore, the addictive nature of social media platforms, driven by algorithms designed to maximize user engagement, can lead to excessive screen time and unhealthy digital habits from a young age. The constant exposure to curated and often unrealistic representations of life on social media can foster feelings of inadequacy and negatively influence self-perception and body image among young users. Moreover, the lack of transparency and control over data collection practices on these platforms raises concerns about privacy and security, leaving children vulnerable to exploitation and manipulation by third parties. Given these risks, it is imperative to approach the role of AI-based social media in children's lives with caution and vigilance. While acknowledging the potential benefits, it is essential to prioritize measures that mitigate harm, empower children with digital literacy and resilience skills, and advocate for policies that ensure their safety and well-being in online environments.

This acknowledgment of my positionality is not merely a reflection of my viewpoint but also serves as a critical lens through which the role of AI-based social media in children's lives is examined. It underscores the need for a careful balance between embracing technological advances and ensuring the protection and nurturing of the younger generation in our increasingly digital world.

Publications arising from this chapter:

- *Chapter 2.1 was published as "Ge Wang, Jun Zhao, Max Van Kleek, and Nigel Shadbolt. 2022. Informing Age-Appropriate AI: Examining Principles and Practices of AI for Children. In Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22). Association for Computing Machinery, New York, NY, USA, Article 536, 1–29. `https://doi.org/10.1145/3491102.3502057`".*
  Best Paper Honourable Mention Award
- *Chapter 2.3 was published as "Ge Wang, Jun Zhao, Max Van Kleek, and Nigel Shadbolt. 2021. Protection or Punishment? Relating the Design Space of Parental Control Apps and Perceptions about Them to Support Parenting for Online Safety. Proc. ACM Hum.-Comput. Interact. 5, CSCW2, Article 343 (October 2021), 26 pages. `https://doi.org/10.1145/3476084`".*

2
Literature and App Review

Contents

The thesis questions introduced in the previous chapter motivate the literature review presented in this chapter. Within the interdisciplinary realm of HCI, this literature review scrutinizes three central aspects that provide the background for this thesis: 1). we delineate the landscape of AI-based platforms concerning children; 2). we explore the phenomenon of datafication, highlighting the algorithmic decision-making processes behind these platforms and the potential risks they pose to children; 3). we scrutinise and assess contemporary digital tools designed to mitigate datafication and other digital risks for children.

Based on these literature review, this chapter highlights gaps in the existing research and studies, and pointing to the thesis chapters that explore corresponding research opportunities.

2.1 AI Landscape

This chapter presents a thorough analysis of AI-based platforms designed for children, coupled with a detailed examination of prevailing ethical AI principles and frameworks. Through a comparative analysis of these domains, we identify existing gaps, particularly the insufficient incorporation of child-centred perspectives in current AI offerings.

The structure of our analysis in this section is threefold:

1. We begin by cross-comparing the principles of current ethical, safe, and trustworthy AI frameworks with those designed specifically for children. From this comparison, we discern significant areas of convergence and identify ten core principles. These principles are then interpreted in the context of child-centred AI.

2. We proceed with a landscape analysis, employing a systematic literature review of AI-based platforms designed for children documented in existing research efforts. This analysis offers a high-level summary of the main systems, shedding light on their target audience, application domains, computational techniques, and data types processed.

3. Finally, we delve deeper into the AI-based platforms reviewed, scrutinising their theoretical and empirical underpinnings, multi-stakeholder involvement, and evaluation methodologies. We then map these insights to the design principles established by existing frameworks.

Section 2.1 summarises by highlighting the current gaps and challenges in the ethical principles for AI-based platforms tailored to children. It also underscores the essential elements that future research should address to advance child-centred AI.

2.1.1 Scope and Definition

To establish the scope of our investigation we first aim to define what we mean by "AI for children". Starting with "AI", the OECD's *Recommendation of the Council of Artificial Intelligence* defines *AI* as "machine-based systems that can, given a set of human-defined objectives, make predictions, recommendations, or decisions that influence real or virtual environments" (Yeung 2020). This definition is cited by the proposed EU *Artificial Intelligence Act* (EUAIA) (*EU AI Act* 2021), as well as the *2020 UNICEF Policy Guidance for AI for Children* (UNICEF 2020) and offers a convenient definition that remains independent of particular implementation or application. With respect to "AI-based", there is some divergence among common use; some papers use the term to mean the particular algorithm or subsystem that enables a particular AI-associated capability (e.g. learning, inference, or recommendation), while others consider a broader context, namely those components that enable a specific application-specific capability (e.g., voice recognition, face recognition, video content recommendation). In other contexts, an "AI-based platform" refers to end-user systems that have such capabilities embedded within them (e.g. intelligent tutoring systems (Graesser, Hu, and Sottilare 2018)). In the scope of this thesis, we adopt the concept of *AI-based platforms*, considering those platforms and systems embedding with algorithmic capabilities as entire end-user systems could be directly used by users. Within the scope of this thesis, we embrace the concept of *AI-based platforms* as follows:

AI-based platforms. Machine-based platforms and systems, embedded with algorithmic capabilities, that are designed to serve as comprehensive end-user systems for direct interaction. They can influence real or virtual environments by making predictions, recommendations, or decisions based on human-defined objectives.

With respect to "children", we adopt the definition according to the United Nations Convention on the Rights of the Child (UNCRC) (*General comment No. 25 (2021) on children's rights in relation to the digital environment* 2020):

Children. A human being below the age of 18 years unless under the law applicable to the child, majority is attained earlier.

While this is a broad definition, it is directly adopted by numerous major ethical AI guidelines, often treating 'children' as a singular category without further exploration. Only specific frameworks, like UNICEF's Policy Guidance on AI for Children (UNICEF 2020), delve into distinctions such as 'adolescence' (10 to 18) and 'youth' (15 to 24). Meanwhile, the UN Committee on the Rights of the Child's endorsement of General Comment 25 provides authoritative guidance on children's rights within the digital domain, outlining four key principles essential for upholding these rights in a digital context:

1. **Non-discrimination**: Ensure all children have equal access to meaningful digital experiences;

2. **Best interests of the child**: Prioritize children's welfare in all decisions and actions affecting them;

3. **Right to life and development**: Protect children from digital threats like violent content, harassment, and exploitation; emphasize technology's influence during early childhood and adolescence and educate caregivers on safe usage;

4. **Respect for the child's views**: Promote children's expression on digital platforms, integrate their input into policies, and ensure service providers to respect their privacy and thought freedom.

With respect to "AI-based platforms for children", the aforementioned UNICEF policy guidance identifies three potential scopes: systems that were explicitly *designed for children* (but not necessarily have to be used by children), systems that *children interact with* (not specifically designed for children but could be accessed by them), and most broadly, systems that may *impact children* (UNICEF 2020). In this thesis, we focus primarily on systems *designed for* children, as we see this as the best starting place for discussing what it means by better AI design *for* children.

2.1.2 Ethical AI Frameworks

There has been increasing effort made on attempts to regulate for more responsible AI. Regulatory frameworks have emerged that attempt to systematically characterise risks relate to AI technologies and establish methods by which risks might be identified and mitigated. A UNICEF review of 20 national AI strategies in 2018 has shown that very little attention has been explicitly given to safeguarding the rights of children in an algorithmic-oriented economy and society (*What do national AI strategies say about children?* 2020). Meanwhile, a separate branch of work has focused on more specifically guiding AI (or digital) technologies for children including ICO's age appropriate design code (**icoaadc**), UNICEF's Policy guidance on AI for children (UNICEF 2020). The two branches of work were separate but related, and sometimes they touched on similar topics but in different ways. There is therefore this confusing smorgasbord of different frameworks and guidelines relating to somewhat overlapping concerns. This can make it difficult for designers and practitioners to effectively establish concrete design suggestions and standards (*Generation AI: Establishing Global Standards for Children and AI* 2019). In order to gain some clarity on how these two sets of framework developments relate to AI-based platforms for children, we present a review comparing the two.

Analysis Methods

Three AI regulatory frameworks are selected for this purpose as they are the only ones across the world with (or close to) regulatory powers at the time of writing:

- **EUAIA** - *The EU Artificial Intelligence Act* (*EU AI Act* 2021): In 2021, the EU proposed artificial intelligence (AI) regulation that applies to any company that develops or wants to adopt machine-learning-based software with EU market exposure. It sets up a series of escalating legal and technical obligations depending on whether the AI product or service is classed as low, medium or high-risk.

- **GRAIA** - *A White House Memorandum on "Guidance for Regulation of Artificial Intelligence Applications"* (*Guidance for Regulation of Artificial Intelligence Applications* 2020): In 2020, the White House issued a *memorandum* providing guidance to federal agencies to consider when developing regulatory approaches to artificial intelligence (AI) applications.

- **ATI** - *Alan Turing Institute's Guide for the responsible design and implementation of AI-based platforms in the public sector* (Leslie 2019): Although not strictly with regulatory status, this framework is recommended by the UK government as an important guidance on the ethics and safety in the development and implementation of AI tools (*AI in the UK: No Room for Complacency* 2020).

Apart from the AI frameworks, we also include three leading frameworks on regulating and guiding AI/digital technologies for children, two of which as legally-binding documents:

- **UNICEF** - *Policy guidance on AI for children* (UNICEF 2020): In 2019, the UNICEF developed a draft policy guidance on AI for children. This document remains a voluntary guidance for self-enforcement at the moment, and aims to promote children's rights in government, whilst informing private sector AI policies and practices, and raising awareness of how AI-based platforms can uphold or undermine these rights.

- **AADC** - *Age Appropriate Design Code* (*Age appropriate design code* 2020): In 2020, the UK ICO published its final Age Appropriate Design Code – a set of 15 standards that online services should meet to protect children's privacy. The legally-binding document sets out the standards expected of those responsible for designing, developing or providing online services that are likely to be accessed by children.

- **UNCRC-** *General comment No. 25 (2021) on children's rights in relation to the digital environment* (*General comment No. 25 (2021) on children's rights in relation to the digital environment* 2020): In 2021, the UN Committee on the Rights of the Child officially launched their new General Comment No. 25 on Children's rights in relation to the digital environment. The adoption of this General Comment signals the first time that children's digital experiences are mentioned within the UN Convention on the Rights of the Child (CRC) as a ***legally-binding*** statement.

We applied a qualitative analysis process to these six frameworks. Each framework can be organised in different ways, by subsections, principles or chapters, which provides a broad indication of their structures. We read through each subsection/chapter to identify specific and distinctive principles (similar to 'themes' in the thematic analysis (Braun and V. Clarke 2006) methodology) discussed in each framework. We recognised that these principles may be named and defined differently in each regulatory framework. By reading through the definitions from each framework and consulting our in-house legal scholar, we consolidated the diverse definitions and terminologies across the different frameworks into 10 common principles, with different degrees of engagement of children's issues. We then reviewed and discussed the principles, drawing on their experience on responsible AI research and regulations of children's technologies. For each identified common design principle, we then concretely identified and extracted the issues specifically around children (Figure 2.1).

Ten Ethical AI Principles

Our result shows that (see Figure 2.1) although there are overlaps amongst the frameworks from various legislation background, limited consideration has been given to child-specific challenges and needs from the AI frameworks. This is in line

COMMON PRINCIPLES	CHILD-SPECIFIC CONSIDERATIONS					
	AIA	GRAIA	ATI	UNICEF	AADC	UNCRC
1.Fairness & Non discrimination	-	-	"parents and guardians should be treated fairly"	"Support marginalised children" "Develop diverse child datasets" "Eliminate prejudicial bias"	-	"All children to have equal access" "prevent discrimination of the basis of sex, disability, ethinic..."
2.Accountability	"Specific assessments on systems could be assessed by children"	-	-	"Review, update and develop AI-related regulatory frameworks" "Mechanisms for redress"	"DPIAs"	"Regular audits and accountability measures"
3.Sustainability	-	-	-	"Support children's long-term development and well-being"	"Support children's physical, psychological and emotional development"	"Special attention on the effects of technology in children's earliest years of life"
4.Transparency	-	-	-	"Use child-friendly language" "Support caregivers to understand"	"Information east to find and accessible for children" "a child-friendly way"	"support and training opportunities on how to effectively articulate the AI system"
5.Privacy	-	-	-	" A responsible data approach" "Promote children's data agency" "Privacy-by-design"	"Default privacy settings on services could be assessed by children" "Data minimisation" "Do not disclose children's data"	"Any digital surveillance of children should not be conducted without the child's knowledge"
6.Safety/Do not Harm	-	-	-	" Children can use digital services and apps in unanticipated ways"	-	-
7.AI for safeguarding /Protect from Harm	"Develop AI to address crimes against children"	-	-	" Leverage AI to protect children"	"AI to protect children from potential harms"	"Protect children from harmful and untrustworthy content"
8. Exploitation and manipulation	"AI could exploit vulnerabilities of children"	-	-	" Continuously assess and monitor AI's impact on children"	"Do not use children's data in ways that are detrimental to their wellbeing"	"Extra attention on systems that affect or influence children's behaviour or emotions"
9.Ensure inclusion of and for all	-	-	-	" Ensure all children can use AI regardless of age, gender, geographic and cultural diversity" "Support child participation"	"Agency that allow children to form their own views and have them heard"	"Promote means for children express their view, and support for children to participate on an equal basis with adults"
10.Meet developmental needs	-	-	-	" AI informed by the unique developmental stage of children"	"Appropriate for children's use and meet their development needs"	"Consider the evolving capacities of children"

Figure 2.1: Ten common principles derived from a thematic analysis of the 6 frameworks for trustworthy and ethical AI (EUAIA, GRAIA, ATI), and age-appropriate design (UNICEF, AADC, UNCRC). The child-specific considerations for each identified principle mentioned in each guidance are indicated. The cells in light grey indicate that the principle has been discussed for the broad population but nor for children specifically; The cells in white indicate no discussion about that principle in that framework at all.

with the previous UNICEF AI strategy review (*What do national AI strategies say about children?* 2020). Below, we present the ten most commonly proposed ethical AI principles found in current policy and legislative efforts.

I. Fairness and non-discrimination. It is recognised that there can be different ways to characterise or define fairness in the design and use of AI-based platforms. The ATI proposes that the *principle of discriminatory non-harm* should be a minimum required threshold of fairness - requiring designers and users of AI-based platforms to ensure that the decisions and behaviours of their models do not generate discriminatory or inequitable impacts on affected individuals and communities (p15. (Leslie 2019)). This also corresponds to one of the fundamental principles outlined in the EUAIA for prohibiting any systems that may involve *"detrimental or unfavourable treatment of natural persons or whole groups that is disproportionate or unjustified"* (p44. (*EU AI Act* 2021)). Although ATI makes an explicit mention of "parents and guardians" (p72. (Leslie 2019)), EUAIA and GRAIA make no specific mention. The UNICEF urges explicitly to *"support marginalized children"* and suggests to *"develop datasets so that a diversity of children's data are included"* and to *"eliminate any prejudicial bias against children"* (p30. (UNICEF 2020)). The UNCRC makes a similar request to the state parties to ensure that *"all children have equal and effective access to the digital environment"*, regardless of their *"sex, disability, socioeconomic background, ethnic or national origin, language or any other grounds"*(p2. (*General comment No. 25 (2021) on children's rights in relation to the digital environment* 2020)).

II. Accountability. The principle is defined in the ATI as built upon two sub-components: the answerability - *"to establish a continuous chain of human responsibility across the whole AI project delivery workflow"*, and "auditability" - *"able to justify the answers to questions of how the designers and implementers of AI-based platforms are to be held accountable"* (p24. (Leslie 2019)). The EUAIA is the only general AI framework that emphasises the importance of extra attention paid to systems that could be assessed by children, in particular, the importance of impact assessments (p9. (*EU AI Act* 2021)). Both the UNICEF and UNCRC urge for *"constant review, update and refine to integrate child rights"* (p34 (UNICEF 2020)). In the AADC, a series of data protection impact assessments (DPIAs) are set for auditing and mitigating possible risks to the rights and freedoms of children. The DPIAs are *"a key part of service providers' accountability obligations under the GDPR"* (p28. (*Age appropriate design code* 2020)), and help service providers to effectively assess and document their compliance.

III. Sustainability. Sustainability is formally defined in ATI as *"designers and users of AI-based platforms should remain aware that these technologies may have transformative and long-term effects on individuals and society"* (p26. (Leslie 2019)), and usually refers to "environmental sustainability" (p36. (*EU AI Act* 2021)) in EUAIA. However, when put under child-specific context, the concept is contextualised by UNICEF as systems that have long-term effects on children, supporting their *"long term development and well-being"* (p28 (UNICEF 2020)). The UNICEF calls for *"prioritising AI-based platforms that can benefit children, and make use of existing well-being frameworks and metrics as a primary success criterion"* (p28 (UNICEF 2020)). Similar statements are also found in AADC and UNCRC. Specifically, the UNCRC mentions how special attention should be paid to *"the effects of technology in children's earliest years of life*, and to support relationships with parents and caregivers, which is crucial for *"shaping children's cognitive, emotional and social development* (p3. (*General comment No. 25 (2021) on children's rights in relation to the digital environment* 2020)).

IV. Transparency. This principle is defined in the GRAIA as allowing *"non-experts to understand how an AI application works"*, and *"technical experts to understand the process by which AI made a given decision"* (p6. (*Guidance for Regulation of Artificial Intelligence Applications* 2020)). As for child-specific transparency, all three children frameworks brought forward the idea of using *child-friendly language* to improve the understandability of information given. The UNICEF also discussed the importance of enabling *"caretakers of children and those around them"* to understand how systems would have impact on children (p33 (UNICEF 2020)). The AADC adds on to the principle by suggesting making the information *"easy to find and accessible for children"* (p38 (*Age appropriate design code* 2020)). The UNCRC further urges to provide children with *"training opportunities on how to effectively articulate the AI-based platforms"* (p10. (*General comment No. 25 (2021) on children's rights in relation to the digital environment* 2020)).

V. Privacy. All three AI frameworks urge providers and developers to ensure respect for users' privacy, some through *technical means as pseudonymisation, or encryption where anonymisation may significantly affect the purpose pursued"* (p48. (*EU AI Act* 2021)). The frameworks also emphasised on *"improper consent of collected data, and improper handling of personal data* (p5. (Leslie 2019)). As for child-specific data privacy, the AADC made extensive effort on regulating data practices for children, including *"default privacy settings on systems that could be* *access by children"*, retaining only *"the minimum amount of children's personal data that is needed"*, and not sharing personal data of children if can *"reasonably foresee that sharing with third parties could lead to detrimental effects on children"* (p7. (*Age appropriate design code* 2020)). The UNCRC also mentions that *any digital surveillance of children should not be conducted routinely, indiscriminately or without the child's knowledge"* (p13. (*General comment No. 25 (2021) on children's rights in relation to the digital environment* 2020)).

VI. Safety/Do No Harm Here safety refers to the "accuracy, reliability, security and robustness" of an AI system (p30. (Leslie 2019)), and that *"developers should prioritise the safety and the mental and physical integrity of people when deploying AI applications"* (p11. (*EU AI Act* 2021)). As for the children frameworks, the Safety of AI is contextualised with children's unique characteristics. The UNICEF explicitly elaborates that *"children are biologically and psychologically distinct from adults*, and more importantly *"can use digital services and apps in unanticipated ways"* (p32. (UNICEF 2020)). This requires AI-based platforms to put in special considerations on the specificities of children when ensuring the safe use of AI.

VII. AI for safeguarding/Protect from Harm This principle states that AI-based platforms should be designed and applied to proactively (and semi-automatically) protect users from harms, be they online and offline. This principle is extensively brought up in the frameworks for children; the AADC, for instance, notes down a list of potential harms that AI-based platforms could protect children against, including *"physical harm, online grooming, access to harmful or inappropriate content, excessive screen time"* etc (p30. (*Age appropriate design code* 2020)). The UNCRC further elaborates on how the digital environment can include *"gender-stereotyped, discriminatory, racist, violent, pornographic and exploitative information"* (p9. (*General comment No. 25 (2021) on children's rights in relation to the digital environment* 2020)). The UNCRC urges for development of AI-based platforms that could protect children from these harmful and untrustworthy content, and ensure their right to freedom and expression under sufficient protection.

VIII. Avoidance of exploitation and manipulation in targeting and personalisation Both sets of frameworks discuss the application of AI to data-driven personalised targeting techniques and the potential harms such techniques could have when applied in various contexts. For instance, the EUAIA explicitly prohibits AI-based platforms that target individuals to cause *physical or psychological harm*, including systems that *deploys subliminal techniques beyond a person's consciousness* or *exploits their vulnerabilities in order to materially distort a person's behaviour* (p12. (*EU AI Act* 2021)). UNICEF urges for *"continuous assessment and monitoring on AI's impact on children"*, especially for those involving personalised targeting, even while the same AI-based platforms may be beneficial to other groups (p32. (UNICEF 2020)). The UNCRC also calls for increased scrutiny of AI-based platforms that could *"affect or influence children's behaviour or emotions"* (p11. (*General comment No. 25 (2021) on children's rights in relation to the digital environment* 2020)). Apart from these, the AADC lists some concrete examples of behavioural manipulation to avoid, including *"using personal data in a way that incentivises children to stay engaged, such as offering children personalised in-game advantages"*, nudging children to continue to play or keep engaging by *"suggesting that children will lose out if they don't"*, or profiling children's personal data to *"make inference about them by exploiting their vulnerabilities"* (p46. (*Age appropriate design code* 2020)).

IX. Ensure inclusion of and for all. This principle is defined in ATI as *"encourage all voices to be heard and all opinions to be weighed seriously and sincerely"* (p10. (Leslie 2019)), and is more often brought up in the children frameworks. The UNICEF argues for ensuring the *"diversity amongst those who design, develop, collect and process data, implement, research, regulate and oversee AI-based platforms"*, and supporting the *"meaningful child participation, both in AI policies and in the design and development processes"* (p28. (UNICEF 2020)). The AADC also encourages *"agency that allow children to form their own views and have them heard"* (p24. (*Age appropriate design code* 2020)). Similarly, the UNCRC requires state parties to *"promote means for children express their view, and support* *for children to participate on an equal basis with adults"* (p3. (*General comment No. 25 (2021) on children's rights in relation to the digital environment* 2020)).

X. Meet developmental needs. The UNICEF states that "the developmental stages and different learning abilities, need to be considered in the design and implementation of AI-based platforms" (p6. (UNICEF 2020)). The term "developmental stage" is also stressed in the UNCRC. The framework encourages state parties to respect the *"evolving capacities of children as an enabling principle that addresses the process of their gradual acquisition of competencies, understanding and agency"* (p4. (*General comment No. 25 (2021) on children's rights in relation to the digital environment* 2020)). The AADC gives some explicit examples on how "developmental needs" could be achieved under the context of transparency. Systems are required to *"tailor the content and presentation of the information according to the age of the user"*, the framework also argues against a *"one-size-fits-all"* approach that does not recognise that children have different needs at different stages of their development (p39. (*Age appropriate design code* 2020)).

2.1.3 AI Landscape in the Literature

Now, we endeavor to explore the landscape of AI-based platforms for children, establishing the foundation for our research. In this process, we conducted a literature review using the PRISMA (Page et al. 2021) methodology, to review the AI-based platforms that have been designed and developed for children in existing research efforts.

We used and combined the following primary and secondary terms for our search. The primary terms were "application" or "system" or "AI" or "artificial intelligence" or "algorithm". The secondary terms were "child" or "children" or "kid" or "teens" or "teenager". The primary and secondary terms were combined and searched in abstracts. We carried out the same search queries in ACM digital library and IEEE Xplore. These two libraries are defined to be the most relevant databases as our purpose was to focus on the technical details of each AI applications. We chose to focus on literature that was published over the last 10 years. Applications of AI for children is a fast-changing area, and so are the regulations in this space. However, research and discussion around it have started since 2011 when UNICEF published its 'Save the Children: Children's Rights and Business Principles Report' (Unicef et al. 2012), emphasising children's fundamental rights when designing for them. Therefore, we believe 10 years is a reasonable time span to reflect on both the more recent practices as well as the more established AI practices and their underlying design ideologies. We only included full peer-reviewed research papers. This resulted in 870 papers from ACM digital library and 1027 papers from IEEE Xplore. We then conducted a more thorough manual elimination process to only select the papers that were specifically describing an application of AI for children in detail. Our screening criteria were as follows:

- Publications that were about AI-based platforms for children.
- Publications that engaged in a technical discussion about how AI-based platforms have been designed and evaluated in different parts of children's everyday lives.

We carried out the manual screening throughout all the search results by skimming through each of the papers. Out of the 1897 collected papers, 483 of them did not talk about a specific application/system developed for children, 702 of them were not AI-enabled, another 341 of them did not engage in a technical discussion, 157 papers were found to not have done evaluation process for the introduced AI system, and another 26 papers were found not to be full papers. The final dataset yielded 188 papers (87 papers from ACM, 101 papers from IEEE).

We analyzed this set to ascertain: 1) the manner in which AI technologies are tailored for children, including facets like target audiences, application domains, computational techniques, and data types used, and 2) the degree to which these technologies are congruent with design principles we delineated earlier in our study (by examining elements such as theoretical underpinnings, stakeholder participation, speculative approaches, and evaluation methods).

Target audience

The review of papers from the past ten years showed that an extensive number (84 papers out of 188) of AI-based platforms were developed for the children, without specific target user groups (Table 1). On the other hand, there still have been various types of AI-based platforms developed specifically for children of different age groups and of different needs. The most represented age groups among the reviewed papers were preschoolers (2-5) and young children (6-12), whereas fewer systems have been targeted at infants (0-1) or teens (13-18). In terms of children with special needs, a considerable number of paper (23 papers) aimed at developing AI-based platforms for children with physical special needs, which includes children with speech/hearing impairment (12 papers), children with visual impairment (1 paper), children with motor disabilities (7 papers), and children with other health issues (3 papers). Even more papers targeted children with developmental special needs (40 papers), among which more than half of them were dealing with children with Autism Spectrum Disorder (25 papers). Other papers include children with learning disabilities (2 papers), depression (2 papers), and mental disorders (8 papers). Apart from special needs of these two categories, another 6 papers focused on children under social disadvantages including being put under the welfare system (4 papers), from low income families (1 paper) and under social risks such as bullying (1 paper).

Age	Nr.
0-5 (Preschoolers)	22
6-12 (Young children)	19
13-18 (Teens)	6
Unspecified	141

Special Needs	Nr.
physical special needs	23
developmental special needs	40
socially disadvantaged	6
Unspecified	119

Table 2.1: Target user groups for the reviewed papers.

Application domains of AI and computational methods used

In this section, we discuss how AI technologies have been applied for children under a variety of different contexts. We organised the applications of AI into 9 domains along *personalised tutoring/intervention, medical diagnosis, harms & safety, social robotics, personalised entertainment, public services, speech recognition, emotion* *recognition* and *age recognition*. An application of AI is often built up upon multiple computational methods. We believe by reporting the specific methods in additional to the application domains, our landscape analysis would provide readers with a more comprehensive overview on how AI has been applied in children's everyday lives - in what application domains that AI is used and their underlying technologies. Therefore, we introduce theses computational methods jointly with the application domains of AI. The computational methods were organised along *classical machine learning, deep learning, reinforcement learning, inferential statistics* and *rule-based models* (Figure 2.2).

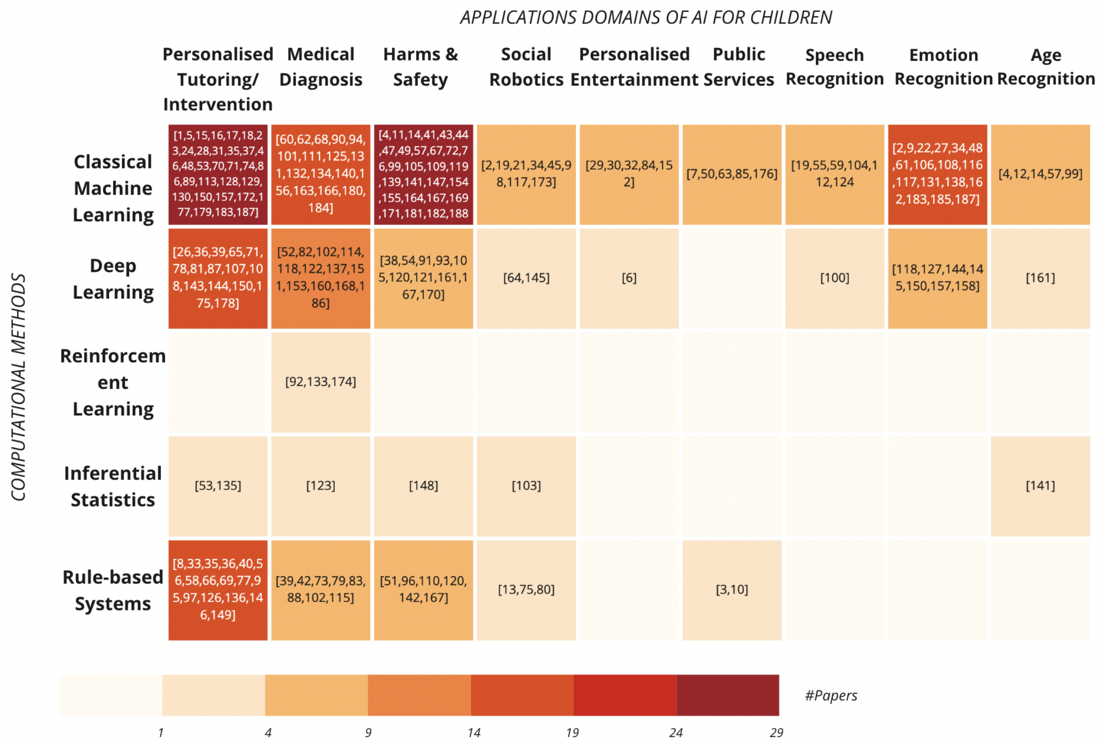


Figure 2.2: Application domains of AI and Computational Methods Used.

Here *classical machine learning* refers to algorithms that parse and learn from data, and make informed decisions based on what it has learned (e.g. supervised learning, unsupervised learning) (Jordan and T. M. Mitchell 2015); *deep learning* involves structuring algorithms in layered formations to construct an "artificial neural network". This network is capable of learning from data and autonomously making intelligent decisions (Goodfellow, Bengio, and Courville 2016); *reinforcement* *learning* is a type of machine learning technique that enables learning through trial and error, by using feedback from its interaction with the environment (Sutton and Barto 2018); *inferential statistics* make use of data analysis to infer properties of an underlying distribution of probability. Its main difference with machine learning is that instead of making predictions, its main purpose is to infer about the relationships between variables (Lowry 2014). Finally, *rule-based systems* apply human-made rules to store, sort and manipulate data (Hayes-Roth 1985).

As shown in Figure 2.2, the most dominant application domain of AI for children was *personalised tutoring/intervention* systems (35%) - with 27 papers on AI-based platforms of pure educational purpose such as generating personalised learning contents for children e.g., (Al-Razgan and Alshaarri 2019; Sobel et al. 2017; Pera and Ng 2013) or assessing children's learning outcomes e.g., (Spaulding and Huili Chen 2018; Spaulding, Gordon, and Breazeal 2016); 14 papers on AI-based platforms that support physical well-being of children e.g., (Yulina and Hajar 2017; Hazman and Idrees 2015); and another 26 papers on AI-based platforms that support cognitive development, such as scheduling personalised strategies to promote children's physical/cognitive development e.g., (Ramachandran, Huang, and Scassellati 2017; J. T. d. Souza et al. 2020; Pacurucu-Pacurucu et al. 2016). Classical machine learning methods were found in the majority of these applications (45%), followed by deep learning (15 papers) and rule-based systems (17 papers).

The second most dominant area of AI for children was the *medical diagnosis* systems (22%), with 36 papers making early diagnosis of disabilities and cognitive disorders, and the remaining 6 papers identifying children at risk of diseases. A striking amount of work (21 out of 36 papers) has been done specifically for early diagnosis of autism spectrum disorder (ASD) e.g., (F. Zhang, Cui, and H. Wang 2017; Mohanta and Mittal 2020; Moghadas and Moradi 2018); and other work includes identifying early signs of disorders such as speech disorder e.g., (Azizi, Towhidkhah, and Almasganj 2012; Shahin, Zafar, and Ahmed 2020; Ringeval et al. 2011) and learning disorders e.g., (He Chen et al. 2017; Khayat, Mabrouk, and Elmaghraby 2012; P. V. C. Souza et al. 2019). Classical machine learning methods (17 papers) were the most utilized computational methods, followed by deep learning (11 papers) and rule-based systems (8 papers).

The third most dominant area of AI for children was AI applications that protect children online and offline (21%). Throughout these research efforts, different kinds of online harms were targeted. The most addressed online harm was online inappropriate content (16 papers) e.g., (Tahir et al. 2019; Kaushal et al. 2016), followed by cyber bullying and harassment (12 papers) e.g., (Dadvar and Jong 2012; Rafiq et al. 2018; Adikara, Adinugroho, and Insani 2020), and detecting crimes against children (8 papers) e.g., (Macedo, Costa, and A. dos Santos 2018; Ayyappan and Matilda 2020). The remaining 9 papers focused on privacy (4 papers) e.g., (Lin et al. 2019; Zakaria et al. 2011) and safeguarding children physically e.g., (F. Zhang, Cui, and H. Wang 2017; Fasching et al. 2012). Again, classical machine learning methods were most exploited (25 papers), followed by deep learning (9 papers) and rule-based systems (6 papers).

A fair amount of research effort has been made on building social robotics (13 papers) - AI driven systems that were designed to engage with or build a relationship with children (Duffy et al. 1999). Systems were designed to generate personalised chat topics thus to communicate with children e.g., (Boteanu and Chernova 2013; Wöllmer et al. 2011), or react differently according to the emotional states of children e.g., (Castellano et al. 2014; Abe et al. 2012). Others were designed to actually physically interact with the children e.g, (Kochigami et al. 2015), or to convince the children that they were another human (peers, students or teachers) through imitating human behaviours such as learning or gaming e.g., (Jacq et al. 2016; Chandra et al. 2018; Hood, Lemaignan, and Dillenbourg 2015; Zheng et al. 2014). Again, classical machine learning methods were used the most.

While perhaps being the most dominant AI use on the market (*AI and the future of media* 2021), there has been relatively less research effort (7 papers) made on personalised entertainment systems for children. The majority of the papers (6 papers) focused on making personalised recommendations for children using classical machine learning methods e.g., (Zeniarja et al. 2018; Santarcangelo and X.-P. Zhang

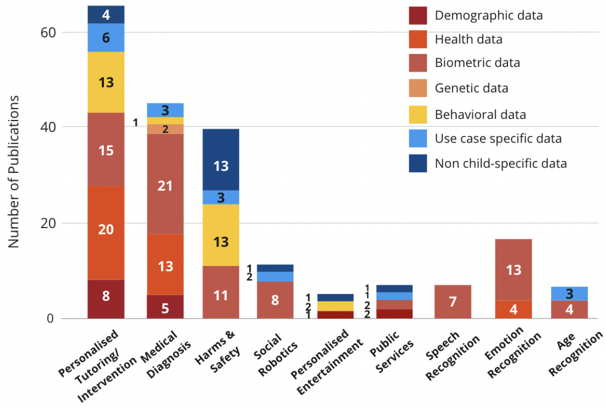
2014). Similarly, we did not find much work on AI-enabled public services for children (only 7 papers). Public service systems here refer to the systems that exploit AI algorithms for social work around children. Typical examples include child welfare systems that decide whether a family needs support or whether a child needs to be taken away, find suitable fostering families e.g., (Andreswari, Darmawan, and Puspitasari 2018; Rodriguez, DePanfilis, and Lanier 2019), or recognise children under welfare risks e.g., (R. Zhao et al. 2019; Ansari et al. 2018).

Recognition systems accounted for a fair amount of work when applying AI for children. *Emotion recognition* systems were among the most common (12%), with a typical context of use being classrooms, educational settings, and online platforms for assessing children's attention and interests e.g., (Baldovino, Vergonio, and Tomas 2019; Castellano et al. 2014). An important additional application was in support of children with autism (Bairavi and Sundhara 2018; Alnajjar et al. 2019). The remaining 14 papers focused on designing and developing *speech recognition* systems (7 papers) and *age recognition* systems (7 papers) for children for distinguishing child and adult users of systems e.g., (Nezhad and Mehrnezhad 2018; Cheng et al. 2020; Rasheed, Rextin, and Nasim 2019).

Types of data processed across application domains of AI

In this section, we examine the types of data that were processed in each of the identified AI-based platforms. The inputs here refer to the type of data needed when actually implementing the AI-based platforms (not just training and testing). We coded the type of data processed into two major categories - personal data and non-personal data (see Figure 2.3).

Personal data. We define *personal data* as *"any information relating to an identified or identifiable natural person"* (*What is personal data?* 2021), which contains "special category data" as defined under the GDPR (data concerning health, racial or ethnic origin, genetics or biometrics) (*What is personal data?* 2021), and behavioral data such as children's searching histories and chat records.



Types of Data	PI	SC	Examples
Demographic data	✓	✓	*name, gender, ethinic*
Health data	✓	✓	*medical history, treatment records, developmental records, test results*
Biometric data	✓	✓	*face, video, voice, fingerprints*
Genetic data	✓	✓	*gene data*
Behavioral data	✓		*searching history, watching history, chat records, location data, previous, user preference*
Use case specific data			*input to the system, demonstration given by children*
Non child-specific data			*videos scraped from video platforms, posts and ads on social media platforms*

PI - personal data **SC** - special category data

Figure 2.3: Types of data processed across application domains of AI for children and our definition of personal data (PI) and special category of data (SC). Applications shaded in different shades of red were identified to use *special category data*, including demographic data, health data, biometric data, and genetic data; applications shaded in yellow used *behaviour data*; those shaded in lighter blue used *use-case specific data*; and those in dark blue did not use any child-specific data. Both *use-case specific data* and *non child-specific data* belong to category non-personal data.

For *personalised tutoring/intervention* systems, 85% made use of children's personal data (with 43 papers using special category data), mainly for building child records to personalise their learning experience (V. Robles-Bykbaev, Velásquez-Angamarca, et al. 2018; Yuhana et al. 2017). Unsurprisingly, the majority of *medical diagnosis systems* required processing personal data; 93% papers we analysed made use of personal data, mainly health data (21 papers) such as medical records (Maharani et al. 2015; Sendari, Widyaningtyas, and Maulidia 2019) and test results (Ivanov et al. 2015; Sangi, B. Thompson, and Turuwhenua 2015). Recognition systems (*speech recognition, emotion recognition, age recognition*) was another domain in which personal data was heavily used, most commonly in the form of biometric data such as images/scans of children's faces (Nezhad and Mehrnezhad 2018) or samples of their voices (Samonte et al. 2019; Rahman et al. 2014).

Non-personal data. The non-personal data refers to the data that "does not contain any information that can be used to identify a natural person" (Finck and Pallas 2020). In our review, it was found to contain two types of data, use-case specific data and non child-specific data. The use-case specific data refers to children's live interaction with the specific application e.g. children inputting a drawing into the system. The non child-specific data refers to the data that is not related to children themselves, such as posts and ads on social media platform. In general, it is not common for AI-based platforms to only make use of the non-personal data (only 37 papers). Harms & safety systems accounted for most of such cases (16 papers), and this is probably because research mostly focused on the content consumed by the children, such as filtering harmful videos for children on video platforms (Tahir et al. 2019; Tahir et al. 2019). Therefore, the data of children themselves is not much needed in these systems.

2.1.4 Mapping AI Landscape with Frameworks

In this section, we discuss how existing AI-based platforms for children have addressed the design principles identified in Section 2.1.2, focusing on the methods and approaches used.

Of the total 188 papers, only 84 papers were found to have addressed any of the the design principles. We found that a fair number of papers (71 papers) have considered to base their design on some kind of ***theoretical*** or ***empirical grounding***. Fewer papers (39 papers) included some form of ***multi-stakeholder involvement*** in their design and evaluation. We did not find any paper that adopted a *speculative aspect*. This is perhaps not surprising as all the papers we included in our review were of actual functioning systems. Finally, in terms of ***evaluation***, we found that while all 188 papers performed some kind of evaluation, only 68 of them involved actual user-centred evaluation (i.e., lab or field studies with children), while the remaining measured system performance on representative datasets (e.g. speech recognition accuracy (Shahin, Zafar, and Ahmed 2020).

Table 2.2 depicts the mappings between the ten design principles identified from the frameworks, and if and how each design principle is considered throughout application domains of AI for children. In general, we found that papers were focused on a small number of specific concerns related to a small number of principles that were most relevant in each application domain. We now describe below: For each application domain, which design principles were seen as most relevant; and for each of such design principles, what kind of ***theoretical/empirical grounding, multi-stakeholder involvement***, and ***evaluations*** have been made to address it.

		APPLICATIONS DOMAINS OF AI FOR CHILDREN								
		Personalised Tutoring/ Intervention	Medical Diagnosis	Harms & Safety	Social Robotics	Personalised Entertainment	Public Services	Speech Recognition	Emotion Recognition	Age Recognition
10 THEMES FROM AI FRAMEWORKS	1.Fairness & Non discrimination	-	-	-	-	-	1	-	**3**	**4**
	2.Accountability	**8**	**7**	-	-	-	1	-	-	-
	3.Sustainability	**15**	-	-	**5**	-	-	-	-	-
	4.Transparency	1	2	-	-	-	-	-	-	-
	5.Privacy	-	-	3	-	2	-	-	-	-
	6.Safety/Do No Harm	-	**42**	-	-	-	-	-	-	-
	7.AI for safeguarding /Protect from Harm	-	-	**39**	-	**4**	-	-	-	-
	8.Exploitation and manipulation	-	-	-	-	-	-	-	-	-
	9.Ensure inclusion of and for all	**19**	-	-	1	2	-	-	-	1
	10.Meet developmental needs	**4**	-	-	-	**5**	-	-	-	-

Table 2.2: Crosstabulation between application domains of AI and ten design principles identified from AI & children frameworks. The number in each cell represents the number of papers from each domain that have addressed each principle. Each cell is not mutually exclusive as a paper can address multiple principles. The most addressed principles from each application domain (with the greatest number of papers) were highlighted in bold.

Personalised Tutoring/Intervention. This domain is perhaps the only domain in which the ten principles were more evenly addressed. Starting from *accountability*, while originally articulated in the frameworks as encouraging audits and continuous monitoring, we found that in practice, this principle was mostly achieved through ***multi-stakeholder involvement***, such as conducting expert consultations co-design workshops, and evaluation with experienced teachers (Al-Razgan and Alshaarri 2019; V. Robles-Bykbaev, Andrade-Prieto, et al. 2018). As for *sustainability*, this principle refers to supporting children's long-term development under child-context. and is predominantly achieved through ***theoretical groundings***. Such groundings include educational theories such as Zone of Proximal Development (Schön, Ebner, and Kothmeier 2012), early childhood theories (Choi 2019; Thida et al. 2020; Mitra, Mostafiz, and Rashid 2017), psychological-based theories such as motivational design theory (Tsiakas et al. 2020), and social support theory (Sobel et al. 2017). All these theories were consulted so as to ensure the system will have a positive long-term effect on children, instead of purely focusing on short-term learning gain. The principle is also achieved through ***evaluations***, to assess children's self-motivation (Jacq et al. 2016) and whether children's learning habits were cultivated (Jacq et al. 2016; Schön, Ebner, and Kothmeier 2012). As for *ensuring inclusion of and all*, the principle is most supported through ***multi-stakeholder's involvement***, which means that children as well as parents were invited to the design (Schön, Ebner, and Kothmeier 2012) and evaluation process (Tsiakas et al. 2020; Chandra et al. 2018). One paper based its design on the ***theoretical grounding*** of self-determination theory, to support student-centered learning (Sobel et al. 2017). Finally, *meet developmental needs* is achieved through both ***empirical groundings*** and ***evaluations***. Children's developmental needs at different ages were consulted during the design stage (Redrován-Reyes et al. 2019; V. E. Robles-Bykbaev et al. 2016). Multiple types of assessments were made including evaluating the ease of use of a system for a given age (Choi 2019; M. and C. 2019), its suitability to children's developmental needs (Redrován-Reyes et al. 2019), and the completeness for covering all the developmental areas of the child (Redrován-Reyes et al. 2019).

Medical Diagnosis. The *medical diagnosis* domain stands out as one where only a few design principles were addressed, focusing narrowly on specific concerns. In fact, the only two principles that were seen as important for this domain are *accountability* (7 papers) and *safety/do not harm* (42 papers). Starting from *accountability*, the principle is mostly achieved through ***multi-stakeholder engagement***, which means that professional therapists (Samonte et al. 2019; Redrován-Reyes et al. 2019) and expert medical doctors (Wanglavan et al. 2019; Gamaethige et al. 2017; Alnajjar et al. 2019; Maharani et al. 2015) were invited in the design and evaluation to assess the reliability and quality of the generated results. As for *safety/do not harm*, it is not surprising to see that this principle is considered to be of huge importance for medical diagnosis systems because any mistakes could potentially lead to tremendous detrimental impact on children. Interestingly, this principle is predominantly achieved through ***evaluation***, and almost all assessments were conducted through pure technical evaluations, measuring the accuracy, precision and recall of the results generated by the systems (Dillhoff et al. 2019; Ivanov et al. 2015) using open-sourced children's data. Out of these 42 papers, 35 of them were found to have only carried out evaluations on results generated by machine learning classifiers (Mohanta and Mittal 2020; Moghadas and Moradi 2018), without any focus on how such results were generated and how they may affect the users; while the other 7 papers were grounded upon more principally collected ***empirical groundings*** (Dillhoff et al. 2019; Ivanov et al. 2015).

Harms & Safety. Similar to *medical diagnosis* systems, this application domain is also one of those that only addressed a small number of design principles, focusing very locally on some specific concerns. The most addressed principle, not surprisingly, is *AI for safeguarding* (39 papers). This principle states that AI-based platforms should be designed and applied to proactively and (semi-automatically) protect users from harms, be they online and offline - which is exactly the design goal of all systems from the domain of *harms & safety.* This principle is achieved mainly from two aspects: first through consulting on ***empirical groundings*** on the types of online and offline harms that could be there for children - such empirical groundings include categories of online inappropriate content out there for children (Tahir et al. 2019; Kaushal et al. 2016), and offline risks such as previously identified risk factors in classrooms (Fasching et al. 2012); second, through conducting ***evaluations*** to assess whether the systems have successfully mitigated such risks and harms (Fasching et al. 2012; F. Zhang, Cui, and H. Wang 2017). *Privacy* is another principle addressed in this domain, which is achieved mainly through consulting on ***empirical groundings***, such as existing findings suggesting children's demand for their privacy, such as online activities and chat records, to be respected by their parents (Lin et al. 2019; Fuertes et al. 2015).

The design principles were less addressed in the other six remaining application domains. That said, we can still see that some principles were considered very locally in certain application domains. In **Social Robotics**, the most well-addressed design principle is *sustainability* (5 papers). Researchers made ***evaluations*** on how learning with robots have impact on children (Hood, Lemaignan, and Dillenbourg 2015; Leite et al. 2016), and how playing with robots impacts children and changes their perceptions of the robots (Spaulding, Gordon, and Breazeal 2016; Chandra et al. 2018; Castellano et al. 2014). In **Personalised Entertainment** systems, the principles were addressed more evenly: *AI for safeguarding* is considered through conducting ***evaluations*** on how successful a system is in shielding harmful content from children; and *meet developmental needs* is achieved through ***evaluations*** on the suitability, readability and understandability of generated contents (Madrazo Azpiazu et al. 2018; Pera and Ng 2013; Bilal 2013). We did not find many papers addressing design principles from the domain of **Public Services** and **Recognition** systems in general. However, the principle of *fairness & non-discrimination* is considered to be most relevant in the **Emotion Recognition** and **Age Recognition** systems, and is addressed through ***evaluations*** on bias among different age groups, genders and races (Cheng et al. 2020; Srinivas et al. 2019; Michalski, Yiu, and Malec 2018; Redrován-Reyes et al. 2019). This is not surprising as both domains heavily relied on the use of children's facial expression data.

2.1.5 Summary

Our analysis underscores a critical necessity to embed a child-centred perspective – to always have their best interests in mind, in the current designs of AI-based platforms for children, for the following reasons:

First, our landscape analysis of the literature showed that while designers and researchers did address some of the common principles in their systems, these considerations were localised to a small subset of principles we identified. Such subsets most often corresponded to areas of greatest concern for the application context; for instance, medical systems focused most often on issues of *safety*, and slightly less often on *fairness/bias*, and much more rarely on any other principles, such as *privacy* - which is particularly concerning as we have previously identified that as high as 93% of medical diagnosis systems made use of children's personal data. Bringing a child-centred perspective not only serve as a reminder of the other principles, but to highlight risks and issues that may be no less important than the ones that are most immediately relevant for comprehensively ensuring the best interests of children.

Second, a child-centred perspective for AI may create a space for researchers to assemble expertise on the ways different kinds of AI-related harms affect children differently from adults, and a corresponding body of theory and empirical findings for mitigation. In our landscape survey, we noted that various aspects of the designs were grounded in, or guided by, theoretical and empirical work. However, such expertise was localised to a few domains, such as education and learning, or fairness/non-discrimination. The remaining principles had much less supporting theory, which may be due in part to the lack of availability of relevant theory or empirical evidence.

Finally, we assert that adopting a child-centred approach to AI could guide standards and regulatory bodies in securing the intricate, multidimensional, and often unpredictable long-term safety needs of children, especially as tech startups swiftly advance AI innovations to the market (*PwC's Global Artificial Intelligence Study: Exploiting the AI Revolution* 2021). Our findings reveal a significant discrepancy, with a notable majority of papers (64%) in our survey neglecting to explicitly address ethical or AI safety risks to children, instead focusing merely on performance metrics like accuracy, rather than potential impacts.

2.2 Datafication

Having defined the scope of AI-based platforms for children, Section 2.2 aims to zoom in with a special lens on a central facet of AI for children: the algorithmic processing of their data, often termed *datafication*. We initiate by elucidating the concept of datafication and then explore its implications on children's online digital experience. We conclude this section by underlining the urgent necessity to empower children against the myriad detrimental datafication practices found on AI-based platforms.

2.2.1 Datafication as Problematic Use of Data

To establish the scope of our investigation, we first aim to define what we mean by *datafication* online, which is broadly referred to as process that children's actions are pervasively recorded, tracked, aggregated, analysed, and exploited by online services in multiple ways that include behavioural engineering, and monetisation (Mejias and Couldry 2019; Zuboff 2019; Mascheroni 2020).

To be more specific, we would like to draw on Livingstone's digital data types framework (Sonia Livingstone, Stoilova, and Nandagiri 2019), in which digital data is categorised into three types: *data given*, which is the data contributed by individuals during their participation online; *data traces*, which is the data left by participation online and captured via data-tracking technologies such as cookies, web beacons or device/browser fingerprinting, location data and other metadata; and *inferred data*, the data derived from analysing data given and data traces, often by algorithms, possibly combined with other data sources. The *Data Inference* process for inferred data is core to datafication online, and refers to the process of analysing data, supported by algorithms, with the aim to evaluate certain personal aspects relating to a natural person (Sonia Livingstone, Stoilova, and Nandagiri 2019), in particular to analyse or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements (*What is automated individual decision-making and profiling?* 2020b).

The algorithmic prowess of AI-based platforms is fundamentally anchored in datafication practices, and such practices are becoming increasingly common in the online world today, and in fact, can be found on almost any online platform (Büchi, Fosch-Villaronga, Lutz, Tamò-Larrieux, Velidi, and Viljoen 2020; Rao, Schaub, and Sadeh 2015; Kazai, Yusof, and D. Clarke 2016). Through a study done on three different companies - Bluekai, Google and Yahoo, Rao et al. (Rao, Schaub, and Sadeh 2015) found a wide range of data including demographic data, data on interests and attitudes and more have been used to make inferences about individuals or groups, which includes predictions about their future actions and inactions, general characteristics and specific preferences. Facebook has been found to make inferences on its users to form an 'interested reading' of their digital trace data (Rieder 2017), so as to create interest classifications that produce sales for advertisers and maintain user engagement on the news feed (Thorson et al. 2021). Research on Instagram also showed that there have been profiling practices on its users, and users were nudged towards certain content such as idealised images which could have negative impacts on the body satisfaction of young girls (Draft Online Safety Bill 2021). Similarly, there has been evidence on YouTube conducting inference on users to maximise their engagement on the platform, which could be particularly problematic for the minors (Draft Online Safety Bill 2021).

Meanwhile, various regulations have attempted to protect children from such practices. For example, COPPA (*Children's Online Privacy Protection Rule ("COPPA")* 1998) protects against the online collection of personal information of children under 13 years of age, and GDPR (Commision 2021) sets restrictions on profiling on children. However, the above mentioned platforms managed to find exemptions for themselves by claiming that there was a minimum age requirement of 13 year old for children to use their service. That said, there has been clear evidence that children, especially those under 13s are still on these platforms, and in fact, have quite heavy usage on such platforms (THORN 2021; Ofcom 2021). A recent report on 2,002 US children showed that 45% of kids under 13 were already on Facebook, and 40% already used Instagram (THORN 2021). YouTube in particular, although has a 'YouTube Kids' version that was claimed to be for under 13s, the most recent Ofcom report still showed that YouTube remained to be the most popular video-sharing-platform among 8-12 year-olds, and more than 85% of preschoolers were found to most commonly used YouTube to watch content (Ofcom 2021).

2.2.2 Negative Impact on Autonomy

By making inferences about individual's lives, datafication has been seen as more than a violation of privacy, or issue of data protection (Büchi, Fosch-Villaronga, Lutz, Tamò-Larrieux, and Velidi 2021); instead, it is more appropriately seen as a potential threat to human autonomy (Wachter 2020) brought by increasingly sophisticated dataveillance techniques (Sax 2016). A number of scholars have broadly categorised the impacts of social media on autonomy through its control over user data, attention, and behaviour (Sahebi and Formosa 2022; Peterson-Salahuddin and Diakopoulos 2020; Furnham 2019). The vast scale of data collection of social media platforms respond first and foremost to a business logic. This raised concerns on exploitation of users' data by social media companies (Zuboff 2019). Companies profit from user data without fair compensation, while platforms may claim that free access to their services justifies data exchange, with critics like Fuchs (Fuchs 2011) assert that this does not equate to a fair trade. Alongside this, questions are raised about the genuine ability of users to opt-out of these platforms due to social or professional obligations (Kuss and Griffiths 2017). Additional complexities include informed consent issues, such as the "transparency paradox" (Nissenbaum 2011), where terms of service are either too simplistic or too convoluted for users to fully grasp.

Beyond mere data exploitation, research has found that social media algorithms critically shape users' beliefs, interests, and actions, including political dialogue and personal values, often creating echo chambers or personalising content in ways that can promote radicalisation (Worden 2019; Mittelstadt 2016). This can be particularly concerning when comes to children. Winpenny et al.'s (Winpenny, Marteau, and Nolte 2014) findings revealed that Facebook exposed 89% of males and 91% of females, aged 15-24, to alcohol marketing every month in the UK, including those under the legal drinking age. Susser et al.'s (Susser, Roessler, and Nissenbaum 2019) research highlighted how social media platforms can micro-target vulnerable teenagers, such as by showing them specific advertisements during moments of insecurity or low self-worth, such as promoting a new watch to an insecure teenager to superficially boost self-esteem.

Addiction to social media is another factor that diminishes autonomy, fueled by intermittent rewards (Bhargava and Velasquez 2021), the fear of missing out (Kuss and Griffiths 2017), and the lack of natural stopping cues (Bhargava and Velasquez 2021). It is a global phenomenon particularly affecting young people. A recent study suggests that 12.5% of 10-year-olds in the UK are losing about one night of sleep per week because they're waking up in the middle of the night to check notifications (University 2022). Meanwhile, a 2019 study with 6595 US adolescents found that teens who spent more than three hours a day on social media faced double the risk of experiencing poor mental health outcomes, including symptoms of depression and anxiety (Riehm et al. 2019).

2.3 Digital Tools for Navigating Datafication

The literature review on datafication practices discussed in the previous section emphasises the pressing necessity to enhance safeguards for children utilising AI-based platforms. Datafication practices are profoundly ingrained in the daily AI-based platforms that children engage with regularly. Identifying and addressing methods to manage such datafication is crucial for their well-being.

Concurrently, given that datafication is not a broadly discussed phenomenon yet, we opted to explore more prevalent approaches to preserving children's safety online. Traditionally, it has been assumed that parents, guardians, and caretakers possess the expertise to guide their children in the digital sphere. This notion has spurred the development of predominantly parent-oriented online safety tools, also known as *parental control apps*. These tools are predominantly crafted to assist parents in monitoring their children's online activities and shielding them from potential online hazards (Zaman and Nouwen 2016).

In this chapter, we conducted an analysis of 58 popular Android parental control apps, analysing their features and user reviews, to understand how current market features are realised and their design principles. Our analysis was twofold:

1. We conducted an analysis of 58 top Android parental control apps designed for the purpose of promoting children's online safety, finding three major axes of variation in how key restriction and monitoring features were realised: granularity, feedback/transparency, and parent-child communications support.

2. To relate these axes to perceived benefits and problems, we then analysed 3264 app reviews to identify references to aspects of the each of the axes above, to understand children's and parents' views of how such dimensions related to their experiences with these apps.

2.3.1 Scope of App Analysis

In July through August of 2020, we sought to identify popular Android parental control applications ("apps") that promote children's online safety on Google Play. We performed keyword searches using the terms "online safety", "family safety," "child safety," "child online safety", "parental controls," "parental monitoring," "child monitoring," "cyberbullying," and "sexting". We also included all 'similar apps' that were suggested by the search results, and read each app description & app screenshots to ensure returned results met the inclusion criteria, which required the app to be designed for parental mediation of children's online activities.

We generated an initial list of 241 apps. During our process of finding and filtering through these apps, we found that the popularity (in number of downloads) of these apps followed a long tail distribution: nearly 80% of these 241 apps had fewer than 20 downloads, while the more popular apps had multi-thousands of downloads. Since our intentions were to examine features and design elements in the most popular apps, we kept the top 66 apps with at least 10K+ downloads. Among these 66 apps, we further removed the following: three that required subscriptions to use, two that required SIM cards (which was not installed on the test device), and three focused on parental control of IoT devices. Our final app list included 58 apps.

We first generated descriptions of app features by applying a walkthrough method (Light, Burgess, and Duguay 2018), first role-playing as a parent and then as a child using the app. This method is similar to the more commonly used cognitive walkthroughs (CWs) in usability evaluation (Lewis and Wharton 1997), in that it is an exploratory screen-by-screen navigation through each app, role-playing as a particular kind of user. However, there are two important differences: unlike cognitive walkthroughs, our objective was not to identify usability problems, but to identify and descriptively characterise the features of each app. Second, instead of starting with extremely concrete tasks, we started with the goal of approaching each app as a methodical new user might, trying all features provided and systematically exploring their options. Each feature was examined to identify all options and functionality. Then, a short textual description was generated for each.

To do this, we installed each app on our devices, in turn. For parenting apps that came in pairs (corresponding to a parent child version), we installed the parent app on an Android tablet, and the corresponding child app on an Android mobile phone. For those apps that were singular apps, we set up the app in parent mode on the tablet, and installed set up the app in child mode on the mobile phone. The tablet used for analysis was a Huawei MediaPad M5 with 32GB of storage running Android 9 and the phone we used was a Samsung Galaxy S20 with 16 GB of storage running Android 9.

2.3.2 App Features

When looking at how features are specifically realised we identified that they fell into three axes of variation, including *granularity*, which refers to the level of control an app enables parents to have or the level of information given to the parents, *feedback/transparency*, which refers to the different designs that support varied level of information given to the children, and finally *parent-child communication support*, which reflect how apps supported or stimulated discussions between parents and children about their online activities. We identified 6 feature *designs* along granularity, 8 feature *designs* along feedback/transparency, and 8 feature *designs* along communications support, as summarised in Figure 2.4.

AXES	LEV	FEATURE DESIGN CODE	%APPS(N=58)
Granularity	Coarse	ALL-OR-NOTHING	40%
	Med	CONTROL-CATEGORY/AGE	32%
		MONITOR-CATEGORY	18%
		CONTROL/MONITOR-CONTACT-LIST	14%
		MONITOR-SUSPICIOUS	16%
	High	CONTROL/MONITOR-PER-APP	34%

AXES	LEV	FEATURE DESIGN CODE	%APPS
Feedback/ Transparency	Low	RULES-NO-SHOW	14%
		MONITOR-NO-SHOW	32%
		ONLY-RULES-AT-MOMENT	20%
	Med	RULES-PAGE	12%
		MONITOR-PAGE	8%
	High	RULES-EXPLAIN	6%
		MONITOR-SELF-TRACK	14%
		SHOW-KIDS-INFO	6%

AXES	LEV	FEATURE DESIGN CODE	%APPS
Communications Support	Low	NO-MEANS	54%
	Med	SEND-MESSAGE	18%
		REWARD	12%
		BOTH-DEVICE	2%
		APP-INFO	4%
		AGE-GROUP-INFO	4%
	High	CO-CONFIG	4%
		PARENTING-TIPS	10%

Figure 2.4: Three axes of variation along which app feature designs varied.

Axes 1: Granularity

We observed that the level of control/information provided by apps for the parents spanned from very coarse level of granularity, e.g. control/monitor based on an **ALL-OR-NOTHING** filtering — parents having to either block all contents or nothing at all; to highly granular feature designs including allowing parents to configure control/monitor on a per app basis (**CONTROL/MONITOR-PER-APP**).

1. (*Coarse*) ALL-OR-NOTHING : This is a most widely supported design (appearing in 40% of all apps). Parents had to either block all contents and request access to every detail (e.g., every video watched, website visited, the apps opened, even logging google search &keyword queries.) about children's online activities, or gaining no access about children's activities at all, leaving no middle ground in between.

2. (*Medium*) CONTROL-CATEGORY/AGE: In comparison to the all-or-nothing approach, this category allows parents to control children's access based on app/website categories. Such categories were usually derived directly from app store listings or website ratings, and seen to afford parents the convenience of setting broad policies without concerns over specific apps or websites.

3. (*Medium*) MONITOR-CATEGORY: Provide only high-level summaries of activities children performed on the phone, such as a list of top contacts, apps used, and time spent on the device per day. These high-level summaries were sometimes grouped by app category.

4. (*Medium*) CONTROL/MONITOR-CONTACT-LIST: At this level, parents are given a chance to provide customised control by providing a list of a pre-approved contact (or "suspicious" contact list); and when children contact people on that list, parents will be informed (or showed with texts etc.)

5. (*Medium*) MONITOR-SUSPICIOUS: Similarly, this control of suspicious content can be automatically achieved by apps, which report and alert parents only based on "suspicious" or "dangerous" activities, messages, or content.

6. (*High*) CONTROL/MONITOR-PER-APP: At the highest granularity of control, parents are required to configure settings on a per-app (or website) basis.

Axis 2: Feedback/Transparency

By feedback/transparency, we refer to the designs (specifically of monitoring features) that support variation of the level of information given to children. The variation of these designs ranges from very low feedback/transparency - providing no information on the screen rules and things being monitored, to high feedback/transparency - supporting children with resources (e.g. expert reviews, ratings) to let them decide for themselves what to use.

1. (*Low*) RULES-NO-SHOW: A considerate amount of apps (14%) provided no indication about the restrictions enforced on children's phones, leaving them with no idea of the things they could still do on their phones.

2. (*Low*) MONITOR-NO-SHOW: Nearly a third of the apps in our data set did not inform children about how their information were being monitored by their parents.

3. (*Low*) ONLY-RULES-AT-MOMENT: This design shows prompts *at the moment* e.g. when children were attempting to access a website on blocking list. However, children weren't informed with the screen rules in advance.

4. (*Medium*) RULES-PAGE, provides a means for children to view/inspect the restrictions policies, giving them more transparency than the designs above.

5. (*Medium*) MONITOR-PAGE, provides some rudimentary information to children about which activities were being monitored, such as browsing history, app use history, device use time, or messaging.

6. (*High*) RULES-EXPLAINED, identified in a small number of apps reviews (6%), offers a clear explanation for children when an activity or action exceeded or violated a restriction and why that might be bad for children (rather than

simply terminating the activity or giving a generic system error or a blank screen).

7. (*High*) MONITOR-SELF-TRACK, provides children feedback about their activities as they used the device, such as how much time they had left (total screen time or on an app).

8. (*High*) SHOW-KIDS-INFO, again identified in only 6% of all apps, offers children with detailed information of each app, including expert reviews and ratings, enabling children to decide what apps to use for themselves.

Axis 3: Communications Support

This dimension pertains to the ways apps supported or stimulated discussions between parents and children about their online activities. This dimension portrayed the variation of such design details. This was done in two ways: first, through features that encouraged communication around the restriction and monitoring policies, and through coaching using discussion aides.

1. (*Low*) NO-MEANS: More than half of the apps (54%) offered no means for children to negotiate screen rules with their parents, and their only choice was to accept and obey.

2. (*Medium*) SEND-MESSAGE: In comparison to the complete lack of communication support, these designs make it simple for children to send a message to their parents asking for permission to perform a particular restricted activity, or to grant an exception or extension to a particularly restrictive policy. These features have been identified in 18% of our apps.

3. (*Medium*) REWARD: Provide extrinsic reward incentives for children to earn additional screen time, app categories and content, or access to ordinarily restricted activities.

4. (*Medium*) BOTH DEVICE: We found one app that coaches or prompts parents and children to spend time with each other, e.g., using a family time lock screen feature for both parents and children's phones to lock both devices out for a duration of time per day to encourage both to spend offline time without the distraction of a screen.

5. (*Medium*) APP-INFO: Provide specific information to help parents ascertain whether an app was suitable for their children's use, to understand potential risks, and other issues about the apps they might want to discuss.

6. (*Medium*) AGE-GROUP-INFO: Help parents to compare their children's online activities (usually screen time) in comparison with other kids of the same age group.

7. (*High*) CO-CONFIG: Provide explicit support that encourages parents and children to set screen rules together, which included interfaces designed to serve as boundary negotiating artifacts (C. P. Lee 2007) for a joint resolution of activity restrictions and monitoring policies.

8. (*High*) PARENTING-TIPS: Offer advice to parents, include strategies ranging from how to apply restriction/monitoring policies to their children, to how to talk about sensitive issues such as bullying, stranger danger, online pornography, and sexting.

2.3.3 Users' Perceptions

The second analysis pipeline focused on app user reviews; to complement our app analysis, we sought to collect user reviews of each of the 58 apps to understand users' opinions. We scraped all reviews for 58 apps (including the *child app* versions) using open source scraping tools ("google-play-scraper 0.1.1" 2020). This resulted in an initial data set of 93,404 reviews from all 58 apps. Duplicate reviews were then removed, yielding 81,488 reviews.

Due to the immense number of reviews being spams, irrelevant, or simple statements without justification like "It's good", we sought to keep only user reviews that expressed a view about specific features with some sort of justification. Thus, we developed the NLP pipeline to achieve this. Our pipeline was based on that by Guzman et al. (Guzman and Maalej 2014). An initial data processing procedure was done to only keep the nouns, adjectives and verbs in the reviews. We then used the bigram finding algorithm provided by the NLTK toolkit (Bird, Klein, and Loper 2009) for extracting user reviews around specific features (A bigram is a two-word phrase that co-occurs unusually often). We filtered the bigrams by only considering those that appeared more than five times and had less than three words distance between them. We then clustered bigrams whose pairs of words were synonyms using Wordnet (G. A. Miller 1995). After that, we manually reviewed each bigrams cluster to map it to each feature identified. For example, both bigrams clusters <limit time> and <limit screen> would belong to the SCREENTIME-BLOCK feature, which is about limiting screen time. This gave us an indication whether an app feature was indeed discussed in a review, and hence made it a potentially 'meaningful' review.

We then traced back to the original reviews where these bigrams that can be mapped to an app feature were extracted from, and went through all of these reviews to manually verify that the corresponding feature was indeed mentioned in the review, remove any inconsistent ones and the ones not about a specific feature. We ended up with a review data set consisting of 3,264 reviews in total.

We used certain keywords and phrases to identify posts by parents and children respectively. Examples indicating a child's perspective included: "my parent", "my mom", "my dad", "I'm a kid", "I'm xx years old". We manually reviewed the automatically generated results. Out of the 3,264 reviews, we identified 746 child reviews and 2,518 parent reviews. We now have a data set consisting of review around each app features, from children, and parents, respectively.

This section presents an overview of the primary themes pertaining to the reasons children and parents, respectively, liked or disliked feature designs provided by parental control apps, and how their perspectives varied across the axes of different ways in which features were realised.

Coarse Granularity

We first start with perceptions around granularity - how different feature designs varied in granularity influenced users' perceptions of features. For *coarse* granularity, we specifically refer to the features that are based on **ALL-OR-NOTHING** filtering. Parents had to either block all contents/shown with every detail (e.g., every video watched, website visited, the apps opened,even logging google search & keyword queries.) about children's online activities, or having no access at all, leaving no middle ground in between. Both parents and children, children in particular, expressed grievances on these features:

Disliked being overly surveilled and restricted [C] - The most common grievance of children regarding **ALL-OR-NOTHING** was the view that these designs not only enabled but nudged parents to set up overly restrictive controls and excessive surveillance. Kids expressed resentment at the extent of both the restrictions and their surveillance, and reflected on the effects lives, welfare, and activities:

> This literally blocks everything, I can't even read e-books on my phone.

> This is insane. Now they can see EVERYTHING! From my browsing history to what apps I downloaded, even my texts! Worst app ever!

Kids reflected on a variety of secondary effects of excessive restrictions and surveillance. One child discussed that they felt that, beyond violating their privacy, it was particularly wrong that their parents' access to their messages would compromise their friends' privacy as well:

> they can eves drop on your convos and stuff that you dont want them to hear [...] not only is it a violation of my privacy that i didnt permit, but it is of friends too that parents dont know about

The pervasiveness and constancy of surveillance made it feel to one as if these apps enabled their parents to "stalk" them:

> I hate this app my mom is like stalking my life!!

Even when not restricted, the perception that their social communications were being surveilled by their parents had a chilling effect that indirectly forced them to cease communicating and be cut off from their friends.

> I can't talk with my friends anymore, everything will be recorded!

Several comments discussed longer-term effects such restrictions were having on their well-being. One view was that restrictions were directly and immediately harmful because they broke essential lines of social support.

> This is stupid. absolutely awful. this will ruin people's lives. I had severe depression and the only thing keeping me from killing my self was my friend who I could only talk to online. Now I'm fine but if I lose contact with that friend I will most likely get my depression back. Horrible and stupid app.

Beyond cutting off lines of social support, restrictions were seen to prevent kids from using apps and activities that they normally used to cope with boredom and isolation, further undermining well-being:

> I can't play a lot of games, and I can't watch YouTube. I've sat in my room for weeks doing nothing and practically getting depressed because there is literally NOTHING I can do!

Beyond the social and emotional aspects of children's lives, a few comments connected these restrictions to developmental and educational harms-namely, how such restrictions impeded learning by depriving them of experiences and learning opportunities:

> The internet is where kids discover and learn new things. And by restricting it, you're denying them that ability.

We don't want to know everything [P] - Interestingly, the most common comments regarding **ALL-OR-NOTHING** from parents were the complaints that the app is showing them *too many* things while leaving them with *too few* choices. Parents expressed resentment in terms of wanting to be able to do more tailored controls based on children's individual needs:

> You either completely stop your kids from using their phone, or absolutely no rules. It's just so silly, don't they know kids these days need to do homework ONLINE?

With regard to monitoring features, parents showed confusion and some reported feeling lost in the vast amount of information given by the apps:

> I don't want to know everything! What's the point of showing all these location info to me? I'm not a control freak! I just want to protect him from the crazy stuffs online, that's all!

We do want to know everything [P] - On the other hand, some parents cherished being able to know every single detail of their children's lives, and that led on to them expressing multiple opposing views around children's rights to privacy - whether children deserved any rights to privacy at all. One stated that since parents were paying for the phone, they should be able to set the rules:

> Don't listen to these spoiled children. I pay the bills. I control the phone. You want to control you pay the bills! Very very simple equation! It's a shame how kids here actually can fathom the thought of "rights" while living in their parents' home. Who made that joke up?

In some cases, parents also talked about how their duty as parents to keep kids safe outweighed any claim to rights:

> When it comes to social media, kids don't need privacy. It's not even about the child as much as it is about others preying on them. I would hope none of the negative reviews are from parents. You should monitor everything your children do. That is our job as parents.

A more lenient version of this was the view that rights to autonomy/privacy should be an earned privilege, not a fundamental right:

> If you want more trust privacy, prove you can handle it with good choices to show your parents you are trustworthy!

Others viewed their children as being too young to make decisions for themselves until they become adults, regardless of their age:

> It's not that we don't trust them, but studies show they can't make decisions or assess risk like adults can until age 25 or so.

High Granularity

The other extreme along the granularity axis is the feature designs that are highly granular - for this, we specifically refer to the designs that allow parents the freedom to configure apps, websites, videos on a "per item" basis (**CONTROL/MONITOR-PER-APP**). Parents were now able (sometimes even required) to refine the specific controls/monitoring they want.

Lost in choices due to lack of support [P] - As we observed earlier, many parents brought up how they'd like to be able to do more refined configurations when they mediate their children's online activities. However, when actually offered with these choices, several of the parents' comments expressed confusion and disappointment regarding these designs. One common complaint was that they simply don't have the time to go through each settings one by one, and they don't have the time to review each app to see if it needs to be banned:

> I'm a working mum with a full time job with three kids. Although I appreciate the app designers' efforts in letting us make the decision. It's just not practical for us working parents to go through all the apps one by one.

In some cases, parents talked about how frustrated they were as they felt they were indirectly accused as irresponsible parents, and they were "nudged" to go back to banning everything:

> I don't see the point in letting us choose which videos for children to watch. One, I can't sit through everything they watch, I have a job. Two, 5-year-olds quickly get bored of the old ones and they want more. This so-called "refinement" just gave me two choices: either let them watch whatever they want (God knows what's on there), or bans the whole platform.

In most cases, parents expressed their needs for supporting resources that help them make restriction/monitoring rules:

> I really wish there's something like an app version of the TV age guide.

Medium Granularity

Feature designs of *medium* granularity offer a middle ground for parents to mediate their children's online activities, without being too coarse or requires too much effort. Designs like this includes control/monitor based on app/website categories or age ratings; or control/monitor of contacts based on a parent-pre-approved contact; or apps reported and alerted parents only based on "suspicious" or "dangerous" activities, messages, or content.

Protection, not punishment: achieving a successful middle ground [P] - Parents were generally positive about the apps that offer feature designs of medium granularity, and they saw achieving a successful middle ground being essential for striking a balance between protection and respect for their children. This was seen not only as convenient for parents, but also supporting setting of boundaries in a more flexible manner. Similarly, both parents and children appreciated how some apps enabled different children, especially older ones, with different level of freedom:

> I love how this app allows us to reach balance. The app only alerts us when it detects something unusual, we can adjust the things he could access as he gets older.

And when such balance was achieved, parents generally saw the parental control apps as effective at helping them support their primary goal, which was to keep their children safe. These parents felt such apps gave them "peace of mind":

> I love this app. I have a 10 year old son that I just recently found out was doing inappropriate things on the Internet with his phone, like viewing porn [...] but with this app I Have some peace of mind. I have full control of what he downloads and views [...] all form [sic] my device.

Reasonable safe zone [C] - Similarly, when app designs managed to help strike a balance between protection and punishment, children regarded them as effective and reasonable at protecting them from dangers online:

> Not too bad, I guess a bit boundary is necessary, at least I still have access to things I love.

> This is just great. It allows age-appropriate control, so giving us more freedom as we get older.

Low Feedback/Transparency

A second axis of feature design variation we looked at is feedback/transparency, and how users', especially children's perceptions varied across different designs along this axis. The designs came in various details along their ability of supporting children to have sufficient feedback and transparency, thus to allow them to learn and understand the screen rules, as well as learn more about their own activities online. For *low* feedback/transparency, we refer to the designs that support very little or no means for children to become informed of these things. Typical designs feature apps that do not signal the restrictions placed on children's phones, leaving them unaware of what they are allowed or not allowed to do. Furthermore, these apps fail to notify children about the information parents can see about them. Additionally, while apps may display prompts when children try to access certain content, they do not proactively inform children about screen time rules or restrictions ahead of time.

Insecure, and not respected[C] - Both children and parents, children in particular, expressed their dislike of feature designs that supported no feedback/transparency, which were sadly the most common feature designs. Rules without prior acknowledgment were sometimes presumed to be annoying by children. For example, both parents and children brought up that prior warnings on time remaining are important. Otherwise, children will feel upset due to the sudden cut-off. Children voiced complaints about certain apps failing to communicate the screen time rules, leaving them without clear guidelines. As a result, they often had to resort to trial and error to understand the limitations imposed by these apps:

> I constantly get these error messages with blank pages, is that part of the screen rules or just an error?

In terms of monitoring, children felt insecure when not knowing what their parents can see about them:

> Can they see all my texts too? That would be creepy.

High Feedback/Transparency

Feature designs that were of *high* feedback/transparency includes designs that not only provided children with means of viewing/inspecting the restriction policies but also offered a clear explanation on why the screen rules were made and why accessing restricted content might be bad for them. Regarding monitoring, features that enhance feedback and transparency are those that give children insights into their own activities, enabling self-monitoring. These features also supply detailed information about each app, including expert reviews and ratings, empowering children to make informed decisions about which apps are suitable for their use:

Keeping them safe and productive [C] - Both parents and children liked feature designs that supported children's understanding of their online boundaries. In terms of restriction, both parents and children brought up how they liked apps offering children with clear screen rule pages so their children would be better informed instead of suffering from being turned down at the moment when they were trying to do some online activities. In particular, many kids pointed out how they liked being offered with explanations on why a website/app was blocked, which made them feel less confused and more respected as a consequence:

> Not the best but definitely better than the previous one. Now I know why the websites are blocked. They give you reasons and things like that. I mean, I disagree with them all the time, but at least they tried to show some respect!

When children understood why the apps were for, they reported more positively about the apps. Several of the comments by children pointed out that the apps helped to keep them safe online from inappropriate things:

> I know this may sound crazy from the kids view, but I love this now! It fits me and my phone perfectly, and my mom knows that I am safe on my phone without having to go to any other horrible apps.

Other children who suffered from addiction online appreciated how the apps dragged them out of that cycle:

> This is AWESOME! I'm a kid so I got this app in order to keep myself in check on my screen time because I am an internet addict. I'm so much happier now! It makes it absolutely impossible to get around!

Beyond the positive impacts online, children also appreciated how the apps helped with their time offline. Some children commented that they were now able to become more productive and spend more time with their families:

> When my mom and dad put this on my phone and tablet at first I hated it. but then I realised with a limited amount of time I spent more time with my family and do actual work. I hope this inspires you to limit yourself with the amount of time you spend staring at useless junk.

Supporting them to make own decisions [C] - In particular, children reported positively about designs that offer them with detailed information (including expert reviews and ratings) of each app. They talked about how they cherish being respected to make their own decisions on what is best for themselves:

> They tell you what others said about this app, but let you decide to use it or not – it's my call.

No/Little Communications Support

A third axis of feature design variation was *parent-child communications support*. Specifically, we looked at how features were perceived differently due to the variation of communications support designs they came in with. For *little* communications support, we refer to the designs that first, did not offer any explanation helping children to understand their screen rules, and also, offered none or very little means for children to negotiate screen rules with their parents, leaving them with the only choice - to accept and obey.

Role of Parenting Apps: Unnecessary, Punishment, or Lazy Parenting [C] - When the purposes of these apps were not communicated with children, several of the children's comments expressed confusion or questioned the role these apps in keeping them safe. One child viewed that these apps were fundamentally redundant because they were already old enough or competent enough to keep themselves safe:

> I'm a 4.0 student. Who is able to manage her school work and screen time by herself thank you very much. I'm old enough to know what's good and bad, I can't change the settings and there no way to let my mom know that.

Other kids lamented that these apps were unnecessary because kids could be told what to do and trusted to keep what they were told. In a sense, the use of apps to force represented a failure of trust.

> How bout you try and talk to them about your phone usage first, see if they'll make a change for you first, then go from there. Think twice before you destroy you kid's trust like this. I understand there's kids who need to get their head straight. But for those like me who are focused in school and a well-rounded kid, I feel all you need to do is talk to them.

The perception that parents were not trusting them made children question the point of these apps. In some cases, children regarded the use of these apps as purely for punishment.

> If you are a parent that wants this app I would reconsider getting this and punish your child a different way. Instead of us being punished, parents should be blamed for dictating their children to use this app.

Some children continued to criticise the bad parenting styles that these apps nudged their parents into. Apart from parents being overly-protective (as we previously reported), children also accused parents of being lazy and using the apps as substitutes for parenting.

> Yes, it's good to be one step ahead, but having an app to do it for you? You might as well call child protective services if you're that lazy.

High Communications Support

Designs that were of *high* communications support mainly came in two ways: first, through features that encouraged communications around the restriction and monitoring policies and through coaching using discussion aides. These includes designs such as apps enabling children to negotiate their rules with their parents, apps providing explicit support that encouraged parents and children to set screen rules together, and apps offering advice to parents - strategies ranging from how to apply restriction/monitoring policies to their children to how to talk about sensitive issues.

Feeling of being respected [C] - When children were given the chance to communicate and negotiate with their parents, they felt they were part of the decision and they generally respected the rules more. In particular, they loved the co-configuration designs that allowed them to sit down with their parents to reach mutual agreement on their boundaries:

> My parents sat down with me to go through this "setting rules together" thing, and I can send a request to them whenever I felt the restrictions are unfair. Love it!

Supports, rather than enacts parenting[P] - Meanwhile, parents also reported their favour towards feature designs that supported or coached them on conversations with their children. And they found designs like parenting tips particularly useful. Some parents felt that they were supported by the app without being hijacked by it.

> This is a great tool. Easy to communicate. Easy to adjust. I use this tool to help, but in noway does it replace being a good parent watching over their child. It should not be the only arrow in your quiver, and you should not expect it to do your parenting for you.

2.3.4 Summary

What do our findings reveal? Our findings showed that the majority of online safety tools to date for children are predominantly parental-guided, restrictive and surveillance-based; offering very little support on fostering children's own agency in managing online safety concerns. Such tools are generally disfavored by both parents and children, as they undermine children's autonomy and infringe on their privacy. This underscores a significant gap in current research, which often operates under the prevailing assumption that parents, guardians, and caregivers play the primary protective role, with children merely at the receiving end of such protection. As indicated by our findings, this traditional approach seems to be resonating less and might benefit from some updates. Both parents and children are now gravitating towards tools that not only monitor, but more importantly, empower and support children in the digital realm.

2.4 Conclusion

This chapter has delved deeply into literature pertinent to the subject of this thesis. In doing so, it has meticulously examined and shed light on the current state of AI-based platforms designed for children. This chapter also underscores concerns about datafication practices in AI-based platforms, particularly their questionable use of data and its potential impact on children's autonomy. This chapter further underscores a prevailing trend: the dominance of restrictive and surveillance-focused mechanisms when considering children's digital well-being and online safety. A crucial gap emerges from this observation, emphasising the absence of a perspective that champions a child-centred approach and supports their autonomy development.

Statement of Authorship for joint/multi-authored papers for PGR thesis
To appear at the end of each thesis chapter submitted as an article/paper

The statement shall describe the candidate's and co-authors' independent research contributions in the thesis publications. For each publication there should exist a complete statement that is to be filled out and signed by the candidate and supervisor **(only required where there isn't already a statement of contribution within the paper itself).**

Title of Paper	Informing Age-Appropriate AI: Examining Principles and Practices of AI for Children.
Publication Status	Published
Publication Details	Ge Wang, Jun Zhao, Max Van Kleek, and Nigel Shadbolt. 2022. Informing Age-Appropriate AI: Examining Principles and Practices of AI for Children. In Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22). Association for Computing Machinery, New York, NY, USA, Article 536, 1–29. https://doi.org/10.1145/3491102.3502057

Student Confirmation

Student Name:	Ge Wang		
Contribution to the Paper	Ge Wang is the principal author who contributed the ideation, data collection, data analysis, and paper write-up. Jun Zhao, Max Van Kleek, and Nigel Shadbolt critically reviewed the manuscript and offered valuable feedback.		
Signature		Date	22/09/2023

Supervisor Confirmation

By signing the Statement of Authorship, you are certifying that the candidate made a substantial contribution to the publication, and that the description described above is accurate.

Supervisor name and title: Dr Jun Zhao, Senior Research Fellow			
Supervisor comments			
Signature		Date	19 March 2024

This completed form should be included in the thesis, at the end of the relevant chapter.

Statement of Authorship for joint/multi-authored papers for PGR thesis

To appear at the end of each thesis chapter submitted as an article/paper

The statement shall describe the candidate's and co-authors' independent research contributions in the thesis publications. For each publication there should exist a complete statement that is to be filled out and signed by the candidate and supervisor **(only required where there isn't already a statement of contribution within the paper itself).**

Title of Paper	Protection or Punishment? Relating the Design Space of Parental Control Apps and Perceptions about Them to Support Parenting for Online Safety.
Publication Status	Published
Publication Details	Ge Wang, Jun Zhao, Max Van Kleek, and Nigel Shadbolt. 2021. Protection or Punishment? Relating the Design Space of Parental Control Apps and Perceptions about Them to Support Parenting for Online Safety. Proc. ACM Hum.-Comput. Interact. 5, CSCW2, Article 343 (October 2021), 26 pages. https://doi.org/10.1145/3476084

Student Confirmation

Student Name:	Ge Wang		
Contribution to the Paper	Ge Wang is the principal author who contributed the ideation, data collection, data analysis, and paper write-up. Jun Zhao, Max Van Kleek, and Nigel Shadbolt critically reviewed the manuscript and offered valuable feedback.		
Signature		Date	22/09/2023

Supervisor Confirmation

By signing the Statement of Authorship, you are certifying that the candidate made a substantial contribution to the publication, and that the description described above is accurate.

Supervisor name and title: Dr Jun Zhao, Senior Research Fellow			
Supervisor comments			
Signature		Date	19 March 2024

This completed form should be included in the thesis, at the end of the relevant chapter.

Publications arising from this chapter:

- *Ge Wang, Jun Zhao, Max Van Kleek, and Nigel Shadbolt. 2023. 12 Ways to Empower: Designing for Children's Digital Autonomy. In Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23). Association for Computing Machinery, New York, NY, USA, Article 91, 1–27.* `https://doi.org/10.1145/3544548.3580935`. ***Best Paper Honourable Mention Award***

3
Autonomy in Digital Spaces

Contents

A primary motivation for this thesis is to take a child-centred approach to support children's autonomy and development within the prevalent datafication landscape. The literature review in Chapter 2 underscores that while growing research has been made on supporting children to become more autonomous in the digital environment around them, there has been little consensus regarding the conceptualisation of what autonomy means for children in digital spaces within the HCI community and how best they can be supported. In this chapter, we explore the notion of *Digital Autonomy* for children. Through a systematic review of autonomy-supportive designs within HCI research, we identified:

1. A landscape overview of the existing conceptualisation of *Digital Autonomy* for children within HCI;
2. A framework of 12 distinct design mechanisms for supporting children's digital autonomy, clustered into 5 categories by their common mechanisms;

3.1 Background and Motivation

A traditional supposition has been that parents and carers would have the greater expertise and skill set than their children, to guide their children's navigation of the digital space around them and help them learn (Shifflet-Chila et al. 2016). In today's digital age, this may not always be true. Children have grown up in the digital era in a way that interacting with digital environment is almost second nature to them, and for most of them, their parents could possess far less knowledge in this domain (Fletcher and Blair 2014; Shifflet-Chila et al. 2016). This potential reverse in expertise calls for attention on the recent line of work promoting a gradual shift from a parent/teacher-led perspective to a child-centred approach (Wood 2007; Langford 2010). Such an approach shifts philosophically from the process of *instructing*, to *supporting* children's experiences, including forms of play and exploration, which are seen as integral elements of children's development (F. P. Hughes 2021; Fagen 2010). Along this line of work, there has been a growing consensus on supporting children to develop autonomy online, including the ability to have and exercise a critical understanding of their digital environments, and to make their own informed choices when interacting with digital technologies and services (Kafai, Proctor, and

Lui 2020; Mullin 2014). However, in the HCI community, what digital autonomy for children means and how best they can be supported is yet well-defined. We argue that this clarity is critical for the current attention on developing better support for children, helping with their skill and autonomy development. Furthermore, it is crucial for us to recognise the landscape of how digital autonomy for children is currently supported, and identify any design patterns or gaps of attention.

In this chapter, we examine existing HCI literature discussing definitions and designs for children's digital autonomy. Our aim is to contribute an understanding of how digital autonomy for children is positioned in the current HCI community, and to identify how specific kinds of designs have been explored to support digital autonomy development in children. In order to do so, we conducted a systematic review of the use of autonomy-supportive design mechanisms in HCI research, with the goal of laying out its design space, specifically answering two research questions as follows:

- RQ1: *How does the HCI literature conceptualise digital autonomy for children?*
- RQ2: *What autonomy-supportive design mechanisms have been explored in apps and systems for children?*

3.2 Methods

In order to investigate the current landscape of designing for children's digital autonomy, a systematic literature review is conducted to identify how digital autonomy has been conceptualised in existing HCI literature within the last 10 years; and how technological interventions and designs have been drawn in to support children's digital autonomy. To achieve this, we followed the PRISMA statement (Page et al. 2021) (see Figure 3.1). We started with identifying a group of keywords to be used for the literature search, the sources for our literature search, and the inclusion/exclusion criteria.

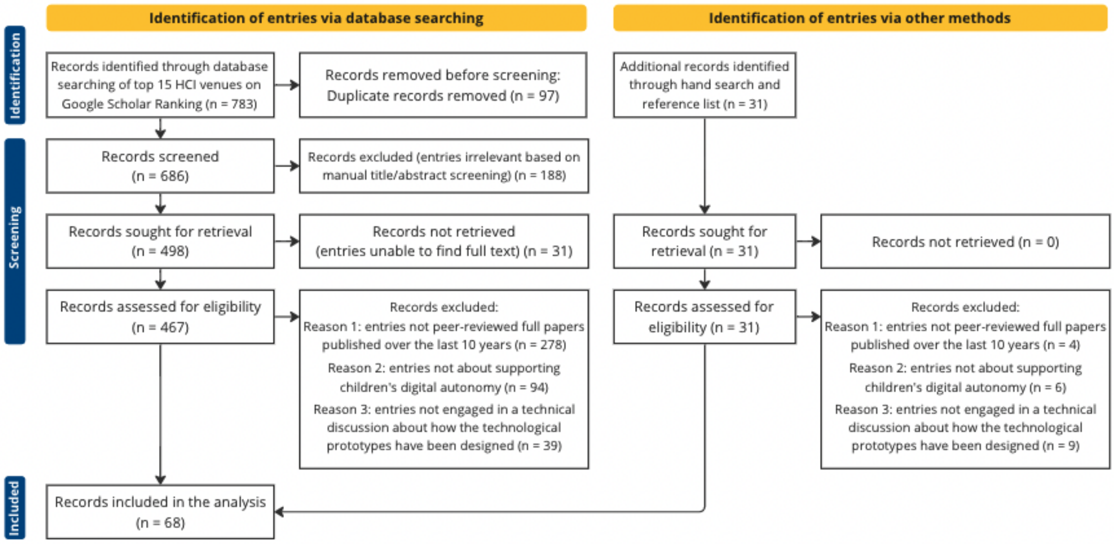


Figure 3.1: PRISMA flowchart of the article selection process.

3.2.1 Data Collection

The unit of analysis for this systematic literature review was peer-reviewed full papers from the top fifteen HCI journals and conferences on Google Scholar Ranking[1] (CHI, CSCW, HRI, Ubicomp, IJHCS, IEEE HMS, IMWUT, PACMHCI, UIST, DIS, IUI, BIT, IJHCI, TOCHI). We also included IDC proceedings due to its strong relevance to the scope of our review.

At the start, we experimented with different keywords combinations related to our research topic, and we identified the final set of keywords which gave us the best matching set of literature for further analysis. We used and combined the following terms for our search: (design OR tool OR app OR program OR game) AND (support OR empower OR inform OR autonomy OR regulation OR self- OR learn OR control) AND (child OR children OR kid). The terms were combined and searched in abstracts. We carried out the same search queries in ACM, Springer, IEEE, Taylor & Francis, and Elsevier. Duplicate records, records unable to find full text, and records irrelevant to the topic of our research were removed, this resulted in 467 papers. We then conducted a more thorough manual elimination process to only select the papers that were most relevant. To be eligible

[1]List retrieved from *Google Scholar Ranking* in April 2022.

for our analysis, the publications must present a novel design solution to digital autonomy and fulfill the following three criteria:

- Articles must be peer-reviewed full papers that were published over the last 10 years. Designing for children is a fast-changing area, however, we believe 10 years is a reasonable time span to reflect on both the more recent practices as well as the more established practices and their underlying design ideologies.
- Articles should focus on supporting children's digital autonomy. We employed a widely accepted definition of adolescent autonomy by Spear et al., as a starting point (Spear and P. A. Kulbok 2001). This includes *Cognitive Autonomy*, signifying independent thinking and the ability to critically evaluate and form personal beliefs; *Behavioural Autonomy*, denoting the capacity for self-regulated action and decision-making based on personal judgments; and *Emotional Autonomy*, which involves managing emotions independently, distinguishing one's feelings from others', and maintaining emotional stability. Meanwhile, we were looking for papers around supporting children's autonomy when interacting with digital technologies. The papers supporting children's autonomy in other fields *using digital technologies* were not included, e.g., healthy lifestyle, road safety, writing skills and etc.
- Articles must be engaged in a technical discussion about how the technological prototypes have been designed. We did not include any papers that were not specifically designed for children, without a discussion of a technical design/implementation, or with a more specific focus on supporting children with special needs.

We carried out the manual screening by skimming through the abstract of each of the papers. Out of the 467 papers assessed for eligibility, 278 of them were not peer-reviewed full papers published over the last 10 years; 94 of them were not specifically about children's autonomy in a digital setting but about scenarios such as healthy eating habits; and another 39 papers did not engage in a discussion on the design of a technological prototype. The final dataset includes 68 papers, from which 12 papers were added following Wohlin guidelines for snowballing in systematic reviews (Wohlin 2014). Of the 68 papers, 47 are conference publications (all full-length proceeding papers), 21 are journal articles.

3.2.2 Data Analysis

We then conducted an analysis on the final filtered set of papers, addressing our two research questions: R1). *How does the HCI literature conceptualise digital autonomy for children?* and R2). *What autonomy-supportive design mechanisms have been explored in apps and systems for children?*

Specifically, for R1, we applied a qualitative analysis process through an open-coding approach (Hsieh and Shannon 2005) by first looking for any existing definition or conceptualisation on the term "autonomy". If the paper does not specifically mentioned the term autonomy, we read through the whole paper to identify what goals or purpose they were trying to support in relation to helping children navigate the digital environment. During this process, we paid particular attention to their research questions, stated contributions, and sentences such as "the focus/goal of this paper is...". We identified a diverse set of positioning of autonomy or the goals towards autonomy, we then consolidated these into three groups of conceptualisation around digital autonomy.

For R2, we aimed to identify the design mechanisms used for supporting children's digital autonomy. Similarly, we also applied a qualitative analysis process through an open-coding approach (Hsieh and Shannon 2005). The analysis was conducted in a *bottom-up* manner (Auerbach and Silverstein 2003). For each paper in our database, we carefully read through all sections related to the actual design of their proposed technological prototypes together with diagrams of the design (if presented), and paid extra attention to descriptions around the incentives, motivation, theoretical and ethnographic groundings behind the design. In particular, for each technological prototype reviewed, we noted down the design details through which it tried to support autonomy (e.g., use of gaming elements, introducing peer collaboration elements, setting default goals for children). The prototypes with similar design details were then grouped into clusters, which were then compared and consolidated, based on the ways in which children's digital autonomy is meant to be supported (e.g., through providing external information input, through providing social context). This process led to a total of five high-level categories.

3.3 Autonomy in Digital Spaces

Our analysis identified three groups of conceptualisations about digital autonomy from the existing HCI literature, namely: digital autonomy as *the ability to develop intrinsic motivation and self-regulation*, digital autonomy as *the ability to make critical thinking and informed decisions*, and digital autonomy as *computational thinking and literacy development* (see Figure 3.2).

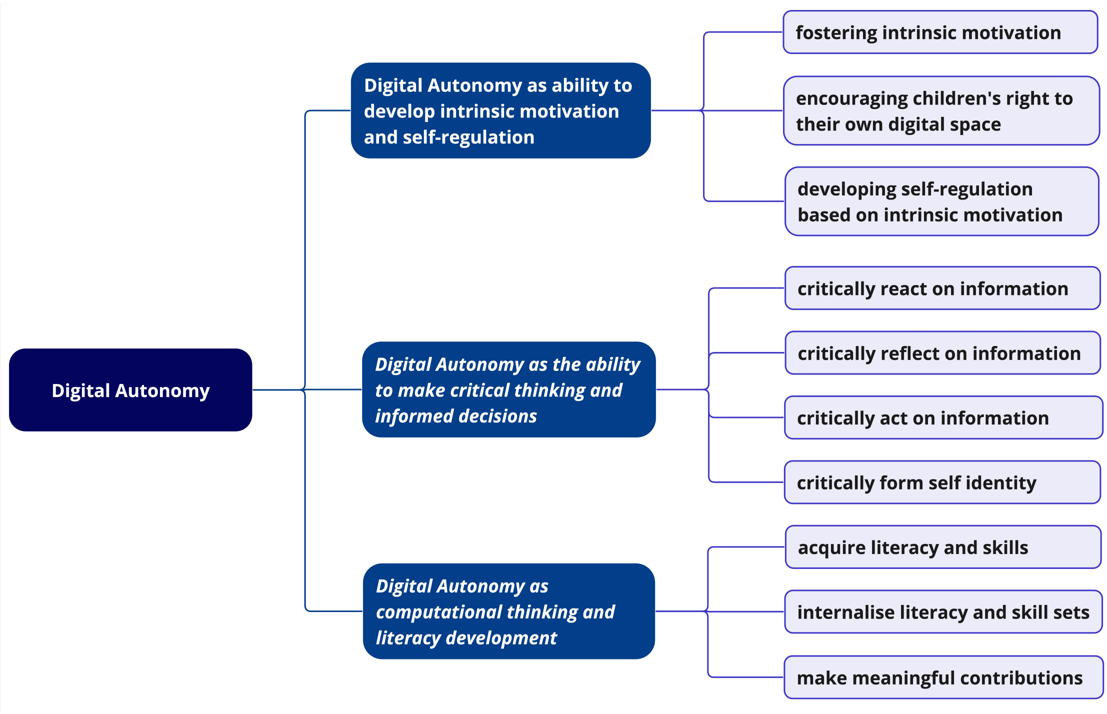


Figure 3.2: A taxonomy on *Digital Autonomy*.

3.3.1 Digital Autonomy as the ability to develop intrinsic motivation and self-regulation.

From about a third of the papers, we identified that digital autonomy has been positioned as developing children's intrinsic motivation and self-regulation. This conceptualisation was usually brought up and discussed in the context of supporting children's self-regulation on screen time (Hiniker, Heung, et al. 2018; Hiniker, B. Lee, et al. 2017; Shin and Gweon 2020; Yim, D. Ko, and W. Lee 2021) and regulating their own online activities (Ghosh, C. E. Hughes, and P. J. Wisniewski 2020; McNally et al. 2018; M. Ko et al. 2015; Sangal et al. 2021; Akter et al. 2022). Through analysis of how autonomy was defined in these papers, we identified three specific focuses related to this type of digital autonomy, including *encouraging children's right to their own digital space*, *fostering intrinsic motivation*, and *developing self-regulation based on intrinsic motivation.*

Supporting *children's right to their own digital space* has been mainly discussed in the context of a child's right to privacy in relation to parental controls and negotiating the power dynamics between children and their parents, as highlighted by work from Ghosh et al. (Ghosh, C. E. Hughes, and P. J. Wisniewski 2020), Mcnally et al. (McNally et al. 2018), Ko et al. (M. Ko et al. 2015), as well as Sangal et al. (Sangal et al. 2021). These research highlighted the importance of children's *autonomy* against parental monitoring and restrictions, which often exercise excessive data surveillance of children's online activities without their consent or knowledge. As a result, these research explored mechanisms to encourage communication of privacy and respect between parents and children so as to support children's development of digital autonomy.

A step further is to support children to *internalise external rules and goals into self motivations*, often drawn from self-determination theory (Hiniker, Heung, et al. 2018; Hiniker, B. Lee, et al. 2017; Shin and Gweon 2020), to emphasise the importance of helping children to translate norms into intrinsic motivations. For example, in a paper by Hiniker et al. (Hiniker, B. Lee, et al. 2017), autonomy is defined as *"for children to regulate their behaviours and develop intrinsic motivation"*, and a similar definition was also found in a paper by Shin et al. (Shin and Gweon 2020) — *"Supporting autonomy as triggering intrinsic motivation, which in turn helps people to internalize rules and show a change in behavior"*. This internalisation could be supported through making children aware of the consequence of their choices (Dempsey et al. 2022), encouraging meaningful discussions with their parents on relevant topics (McNally et al. 2018; M. Ko et al. 2015), as well as supporting children to identify their own intrinsic interests and take ownership of constructing their own plans (Hiniker, B. Lee, et al. 2017).

Finally, built on this internalisation, some research further emphasised the importance of *encouraging actual action development by fostering children's intrinsic motivation*. For instance, Hiniker et al. (Hiniker, Heung, et al. 2018) explored supporting children self-regulating their screen time, defining autonomy as *"the ability to self-regulate: plan, set goals, and choose their own actions with intention"*. Similarly, Ko et al. (M. Ko et al. 2015) explored ways for children to self-manage their online activities, defining autonomy as *"the quality children need to develop into self-dependent adults"*, and *"the ability to self-regulate through making responsible choices"*. This aspect is different from the previous aspect by its emphasis on children's ability to take actions, driven by their intrinsic motivations. Efforts on supporting children to take actions include the use of notifications and alerts to remind children of their goals (Hiniker, B. Lee, et al. 2017; Hiniker, Heung, et al. 2018), as well as encouraging children to self-monitor and self-reflect on their plans and goals from time to time (Sangal et al. 2021; M. Ko et al. 2015).

3.3.2 Digital Autonomy as the ability to make critical thinking and informed decisions.

Another group of research, which although have not explicitly defined digital autonomy in most of the cases, addressed digital autonomy through their goals towards supporting children's ability to make critical thinking and informed decisions. This conceptualisation was usually brought up and discussed in the context of supporting children to interpret online information (e.g., adverts, stereotypical content) (Parker et al. 2013; Ballagas et al. 2013; Rubegni et al. 2022; Hou et al. 2015; Gauthier et al. 2022; Maqsood, Mekhail, and Chiasson 2018), or to cope with more specific risks such as online privacy risks (José Alemany et al. 2019; Kumar et al. 2018; Yap and J.-J. Lee 2020; Dowthwaite et al. 2020; Jose Alemany, Del Val, and Garcia-Fornes 2020; Alemany Bordera, Del Val Noguera, and García-Fornes 2020; J. Zhao, Duron, and G. Wang 2022; M. Williams, Nurse, and Creese 2019), online harmful contents (Hashish, Bunt, and Young 2014; Poblet et al. 2017; Badillo-Urquiola et al. 2019; Baciu-Ureche et al. 2019), and cyberbullying (Piccolo, Troullinou, and Alani 2021; Ashktorab and Vitak 2016). The goals of these research mainly focused on children's ability to *critically act on information* and *form self-identity*.

Research on supporting children to critically act on information emphasised on the importance of providing sufficient information for children, so that they can sense-make the current situation, be more aware of associated risks, and consequently self-reflect and take actions based on the information. In some papers, autonomy is defined as *"for a child to make their own informed decisions about what information to disclose online"* (Kumar et al. 2018) or *"to make decisions and follow through at their own pace"* (Jose Alemany, Del Val, and Garcia-Fornes 2020). Some exemplar research include Parker et al. (Parker et al. 2013) proposed that children should be made more aware of how marketing information could be conveyed in online food adverts and investigated various ways to achieve this. Zhao et al. (J. Zhao, Duron, and G. Wang 2022) developed a prototype which enables children to be made aware of the data tracking associated with their mobile apps and how their data could be transmitted to platforms and companies without them knowing. Rubegni et al. (Rubegni et al. 2022) supported children to be more aware of online gender stereotypes by encouraging them to self-reflect on their own possible choices when creating a digital story. And similarly, in a privacy game developed by Maqsood et al. (Maqsood, Mekhail, and Chiasson 2018), children were instructed to go through an everyday life of a game character Jo, and form judgement on what is the best decisions for Jo to make at various privacy decision points.

Apart from the line of work around supporting children to interpret online information and cope with online risks, there has been another interesting line of work around cultivating children's critical thinking around their self-representation online and thus form self identity. An example is Hou et al.'s work (Hou et al. 2015) on supporting children's online identity in international communities. Another paper (Speer et al. 2021) explored supporting children to be aware of their own emotional states when interacting with online content, thus reflect on their self identity online.

3.3.3 Digital Autonomy as computational thinking and literacy development.

Digital autonomy was not explicitly defined in the third group of research in most cases. However, these papers presented ways to 'acquire the knowledge and understanding to evaluate thoughts and make decisions', i.e., the cognitive-level autonomy — through supporting children to develop an ability to *internalise the literacy and skill sets*, and to *make meaningful contributions*. This conceptualisation was usually brought up and discussed in the context of supporting children's literacy development such as computational thinking skills (Dietz et al. 2021; Dasgupta and Hill 2017; Rode et al. 2015; Brazauskas et al. 2021; Dasgupta 2013), coding skills (Ofer et al. 2019; Deng et al. 2019; Cabrera, Maloney, and Weintrop 2019), as well as various forms of digital literacy including algorithmic literacy (Hitron et al. 2019; Nasi et al. 2019; Litts et al. 2019; Zhu et al. 2016), data literacy (Bowyer et al. 2018; Wolff, Wermelinger, and Petre 2019), AI literacy (Druga and A. J. Ko 2021; Van Brummelen, Tabunshchyk, and Heng 2021) and etc.

To have the ability to *"internalise the knowledge gained"* (T. Y. Lee et al. 2014) is positioned as crucial for children to develop computational thinking in a way that connects computer-based problems with their personal everyday scenarios as well as broader social issues and challenges, and then apply reasoning and actions upon them. Several papers explored how to support this internalisation process. For example, Rode et al. talked about enabling this internalisation process through problem-solving (e.g., *"generalizing and applying this problem solving process to other kind of problems"* (Rode et al. 2015)); Other papers also worked on encouraging children to relate new knowledge gained to their everyday life context (e.g., *"apply their new knowledge to everyday life context, including personally meaningful applications"* (Bekker, Bakker, et al. 2015)), as well as to the deeper and larger context (e.g., *"Children were not only able to store and access data online, but they also get to explore larger and powerful ideas like privacy, scale, etc."* (Dasgupta 2013)). Related to this internalisation process, to *make meaningful participation* emphasised on children gaining the ability to voice and form their own opinions and conduct meaningful discussions based on literacy gained. For instance, Bowyer et al. (Bowyer et al. 2018) showed how children could develop voices towards civic data issues (related to their data autonomy) through a card game activity and discuss their concerns with their parents.

3.4 Autonomy Design Mechanisms

How has digital autonomy been supported? In this section we present an analysis of the 12 design mechanisms for supporting children's digital autonomy, clustered into 5 overall categories by their common mechanisms.

3.4.1 Scaffolding

The first category of design mechanisms was synthesised based on their mutual standing on providing *external information inputs* for children to develop digital autonomy. More specifically, such external information inputs were typically provided in a scaffolding manner, which originates from Vygotsky's Zone of Proximal Development (ZPD) (Wertsch 1984). Vygotsky believed that when a student is in the zone of proximal development for a particular task, providing the appropriate assistance will give the student enough of a "boost" to achieve the task (Wertsch 1984). These design mechanisms provide children with support when they need it and helped them to move through their gaps of knowledge, see Figure 3.3 for a graphical representation and summary.

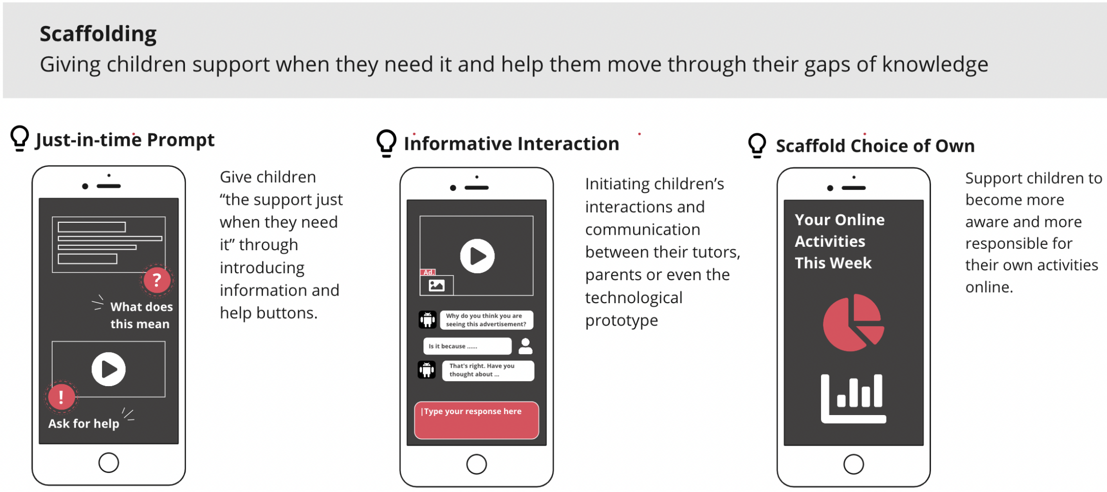


Figure 3.3: The autonomy mechanism *scaffolding* includes three design mechanisms: *just-in-time prompt, informative interaction* and *scaffold choice of own.*

Just-in-time Prompt. Just-in-time prompts or 'pop-ups' have been brought into the prototype designs as means to give children "the support just when they need it" (Hendrix, Eisenberg, et al. 2006). These designs provide an opportunity for children to conduct critical thinking and to make informed choices about the technology, typically through the use of 'help buttons' or 'information buttons'. For instance, *TalkBack* (Parker et al. 2013) developed a mechanism in which they present expert nutrition tips next to food ads online (e.g., displaying "Choose whole fruits rather than juice drinks with added sugar.") to promote children's critical thinking about food adverts. Similarly, in a prototype designed for supporting children navigating informational privacy online (*DOPA* (Yap and J.-J. Lee 2020)), a question is posed to children when they encounter targeted advertisements, prompting them to think about why this ads is generated for them, with the answer revealed at the following page, explaining how third-party tracking cookies work and thus help children interpret the targeted advert in front of them.

Informative Interaction. An important aspect of the ZPD theory is for the learner to interact with and learn from a more knowledgeable other (with knowledge and skills beyond that of the learner) (Wertsch 1984). Designs in *informative interaction* mechanisms work by initiating children's interactions and communication with these knowledgeable others, such as from their parents (M. Ko et al. 2015), teachers (Bekker, Bakker, et al. 2015), or the system itself (Badillo-Urquiola et al. 2019) when they need help, so as to give them the information to prompt their critical thinking and informed decision making. For instance, Badillo-Urquiola et al. (Badillo-Urquiola et al. 2019) designed a parent alert button and a "police popo" button, which children can click on and seek for advice when they encounter messages from strangers online; *We-Choose* (Hashish, Bunt, and Young 2014) explored mechanisms that enable children to communicate and collaborate with their parents to set content filtering rules and establish what is appropriate. Meanwhile, designs were also implemented in ways to encourage children to ask questions or have conversations with the prototype itself, thus to scaffold them to navigate through online information (Woodward, McFadden, Shiver, Ben-Hayon, et al. 2018).

Scaffold Choice of Own. The goal of ZPD is for children to move through their gaps of knowledge and be able to make their own decisions. Designs in this mechanism aim to support children to become more aware and more responsible for their own activities online. These designs are usually related to inspire children's intrinsic motivation and self-regulation. For instance, *Familync* (M. Ko et al. 2015) allowed children to self monitor their usage, thus to encourage their goals on limiting screen time and increasing study time. Similarly, another app called *Teen-alyse* (Sangal et al. 2021) developed for children's self-regulation through showing them their app usage, together with a comparison of the rules set for children and whether they have exceeded that rule. Another design example was implemented in *MediaKids* (Poblet et al. 2017), in which they explored the design idea of helping children to set up family media agreement with their parents, and encouraging them to follow through these rules by reminding them from time to time.

3.4.2 Decomposing

The second category of design mechanisms were synthesised based on their mutual standing on providing *fictional context* for children to develop digital autonomy.

Thinking in abstract concepts could be difficult for children (Siegler 1991). Decomposing provides a natural way of formalising tacit knowledge through breaking down and providing analogs — “a familiar scenario to think with” (Ackermann et al. 1996), which has a critical place to help children explicitly link their existing knowledge to the more abstract digital concepts. These autonomy mechanisms were designed to break down somewhat complicated concepts into entities that are more approachable for children, and provide analogies and metaphors that make concepts easier for children to process and to relate to themselves, advocating child-centred discovery learning, see Figure 3.4 for a graphical representation and summary.

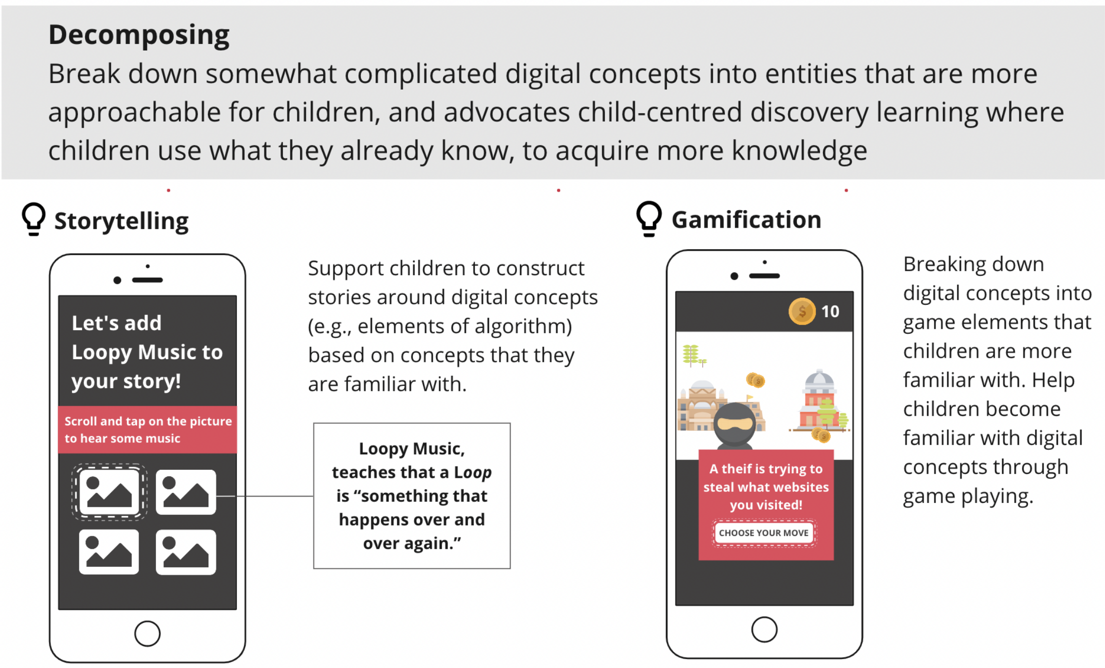


Figure 3.4: The autonomy mechanism *decomposing* includes two design mechanisms: *storytelling* and *gamification*.

Storytelling. Storytelling has long been known as effective ways for children to increase their comprehension skills (S. Miller and Pennycuff 2008). Designs have been made to break down elements of digital concepts into stories by including relatable elements such as familiar characters. By exposing children to a range of digital scenarios, children were encouraged to develop agency and critical thinking for informed decision making, through relating these story scenes to their own

experience. For instance, a series of comic-based stories were developed to support children to get familiar with internet safety scenarios (Baciu-Ureche et al. 2019). Other designs also worked on supporting children to construct their own stories using digital elements. For example, in *StoryCoder* (Dietz et al. 2021), children were introduced to the concept of loop by incorporating loppy music when creating their story. Kumar et al. (Kumar et al. 2018) supported children to create their own stories around privacy decisions (e.g., the reader receives a suspicious email purportedly from a former classmate), while encouraging them incorporate decision points (e.g., whether the character in their story should click on the link inside the email) in their stories. Similarly, Hou et al. (Hou et al. 2015) invited children to create their own stories, with the focus to help them explore their online identity.
Gamification. Similar to storytelling, designs in gamification mechanisms work by breaking down digital concepts into game elements, and help children construct their knowledge as they interact with the games. We have observed different styles of implementation of gamification mechanisms: using gamification as a mechanism to *increase children's engagement*, or to *help children develop critical thinking as well as computational thinking skills* through asking children to follow a specific set of game rules. In terms of the former, Garcia et al. explored ways to incorporate gamification to motivate children (e.g., gain more star by finishing more exercises). On the other hand, most other gamification designs we observed fell into the second category. For instance, Bowyer et al. (Bowyer et al. 2018) developed a card game that gamified data concepts, instructing children to play card games following rules which help them get familiar with how data were collected and processed (Bowyer et al. 2018). Similarly, another data card game was developed by instructing children to put data cards into a black box, thus mimic machine learning processes (Dowthwaite et al. 2020). Apart from card games, other games were developed to compare online privacy protection to protecting village against privacy thieves (M. Williams, Nurse, and Creese 2019), or to instruct players to follow main characters through a series of events in their daily digital life and help them make smart decisions online (Maqsood, Mekhail, and Chiasson 2018).

3.4.3 Peer Support

Design mechanisms in this category typically originates from their mutual standing on providing *social context* for children to develop digital autonomy. Design mechanisms in this category encourage social interaction between children and their peers in order to promote their digital autonomy. Unlike scaffolding mechanisms, peer support mechanisms encourage children to turn to their peers instead of adults. Comparing with the traditional parent-child support model, peer support contribute to higher levels of active engagement for children (Shukla, Kennedy, and Cushing 1998), as well as improvements in skills, self-confidence and relationships (Coleman, Sykes, Groom, et al. 2017), see Figure 3.5 for a graphical representation and summary.

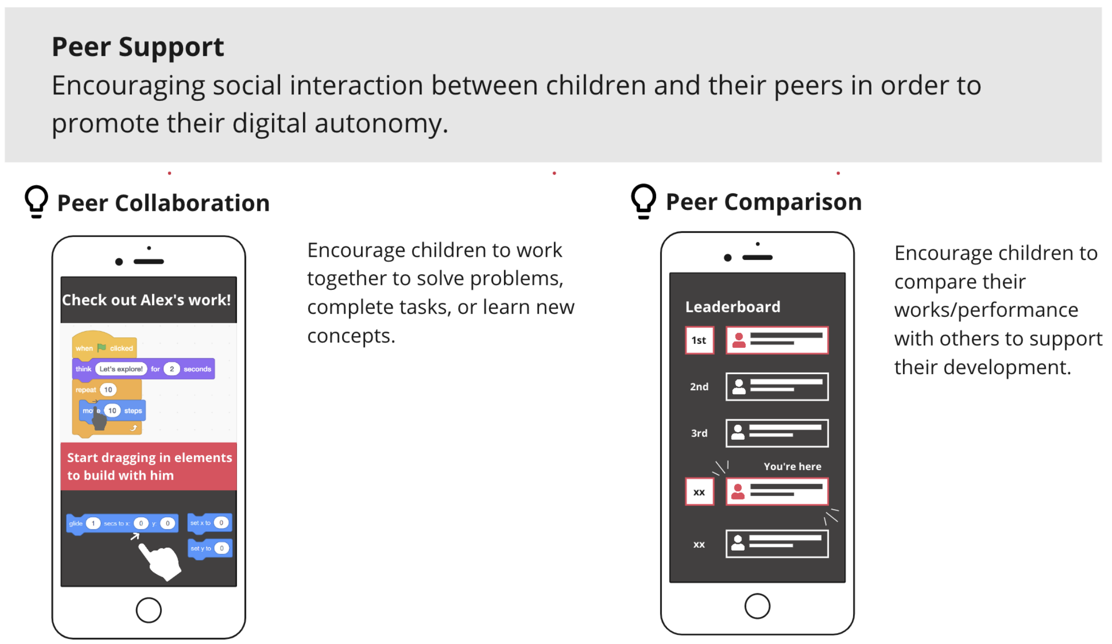


Figure 3.5: The autonomy mechanism *peer support* includes three design mechanisms: *peer collaboration* and *peer comparison*.

Peer Collaboration. Enabling children's collaboration has been a long established topic. Through collaborating on tasks, reframing ideas, listening to each other and articulating their points, children will gain a more complete understanding as peers than they could as individuals (Laal and Ghodsi 2012), as designs supporting them to develop their critical thinking as well as computational thinking skills. *Catriod* (Gritschacher and Slany 2012) is a Lego-style programming environment which supports children to build their work on each other — “standing on the shoulders of their peers”. Dasgupta (Dasgupta 2013) developed an online environment that encourages children to collaborate on coding projects. Other designs include setting up online forums that allow children to share their work (Dasgupta and Hill 2017), and online communities that allow children to work together, to collectively manage their privacy and security through posting and commenting as well as direct messaging one another (Akter et al. 2022).

Peer Comparison. The social cognitive theory suggests that human especially children learn new behaviour by modelling others’ behaviours through observational learning processes whereas children learn new behaviours and strengthen their behaviour by observing the effect of others’ behaviours and copy it through vicarious reinforcement processes (Bandura 2005). Design examples typically work through encouraging children to compete with each other. For instance, in *TalkBack* (Parker et al. 2013), an app designed for cultivating critical thinking when interpreting online adverts, “friendly competition” is encouraged such that whoever wrote the most critical comments is awarded the Top Talker position to encourage engagement. Similarly, another design worked by building communities that allow children to compare their work with others, observe what others do and go back to improve their own work (Gritschacher and Slany 2012).

3.4.4 Digital Playground

Design mechanisms from this category were synthesised based on their mutual standing on providing *embodied context* for children to develop digital autonomy. These designs encourage children to freely interact with a digital system which is connected with a physical artefact, thus extending the playground to the physical environment around them (see Figure 3.6 for a graphical representation and summary). The goal of these designs is typically around developing children’s computational thinking and digital literacy. For instance, Ofer et al. (Ofer et al.

Digital Playground
Encouraging children to freely interact with digital systems in more embodied ways, learning through playing.

Figure 3.6: The autonomy mechanism *digital playground*.

2019) explored the idea of a coding platform that controls a programmable hardware device for children's outdoor play, thus supporting children generate outdoor game ideas and implement them through coding activities. *DataMove* (Brazauskas et al. 2021) is an interactive physical computing artefacts developed that enabled children to explore number systems and data through embodied movement and dance. Hitron et al. (Hitron et al. 2019) worked on a platform on which children can train ML systems using a hand-held input device by performing different hand movements. Other designs include instructing children to make stuffed toys from pieces, connect them with electronics and then learn how to program it to be interactive (Rode et al. 2015); and supporting children to imagine, code and display their programs on 3D surfaces (Hsi and Eisenberg 2012); as well as systems that made children's programming edits immediately reflected in the behavior of a physical device (Cabrera, Maloney, and Weintrop 2019).

3.4.5 Nudging

A nudge is defined as "any aspect of the choice architecture that alters people's behavior in a predictable way without forbidding any options or significantly changing their economic incentives" (Leonard 2008). These autonomy mechanisms were designed to impose subtle design changes that could alter children's behaviors and reinforce positive ones, and is closely related to the dual process theories (Groves and R. F. Thompson 1970). See Figure 3.7 for a graphical representation and summary.

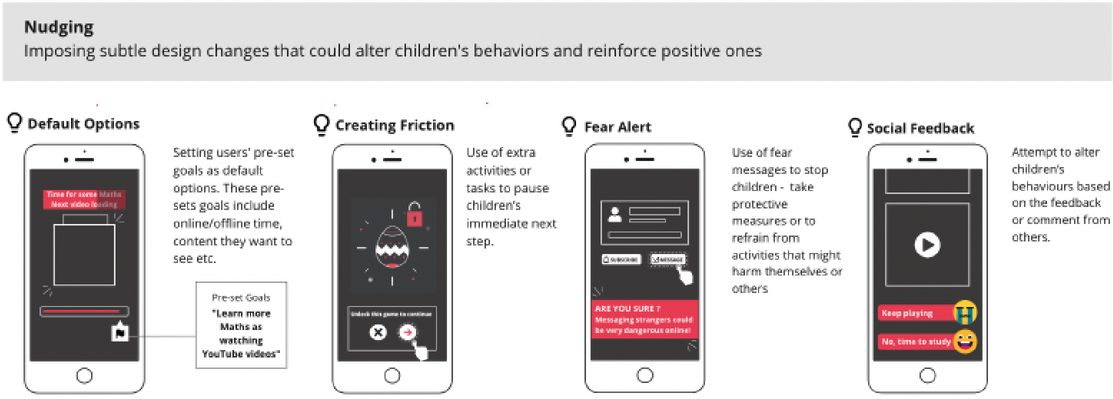


Figure 3.7: The autonomy mechanism *nudging* includes four design mechanisms: *default options, creating friction, fear alert* and *social feedback*.

Default Options. Perhaps the most widely known type of nudging is default options, which is simply what happens if you do nothing. The power of the default has long been acknowledged to have a significant impact on individuals' choices (Goldstein et al. 2008). In our review, we found that default options were mainly realised through making users' pre-set goals (e.g., online/offline time, content they want to see such as educational resources) as their default options so as to support their self-regulation. For instance, in *Coco's Videos* (Hiniker, Heung, et al. 2018), a video player designed for children to self-regulate their media use. It was found that lock-out mechanisms (no element of the screen was interactive, the user was locked out of the app once they reached this point) were more effective in terms of helping children to stick with their transition plan.

Creating Friction. Friction nudges aim to minimize intrusiveness and the sense of restriction while maintaining the capacity to change children's behaviours. The typical design examples we observe from our review includes the use of extra activities or tasks to pause children's next step so as to help them enforce self-regulation or enable stop-and-think for critical thinking. For instance, *MABLE* (Shin and

Gweon 2020) explored ways of introducing offline leisure activities (e.g., hide-and-seek games) for children to interrupt their continued screen activities. *Romi* (Yim, D. Ko, and W. Lee 2021) was introduced to children as a peripheral companion, such that children would need to interact with it before continuing their online activities. Gauthier et al. (Gauthier et al. 2022) explored a series of design features including the use of motion and colour, to initiate children's stop-and-think behaviours. Other designs include hiding some apps or content from children, making them harder to find (Akter et al. 2022).

Fear Alert. While friction nudges push children away from certain behaviours, remind nudges demand immediate attention and action from children. These designs make use of children's fear to stop them from doing harmful activities. For instance, Badillo-Urquiola et al. (Badillo-Urquiola et al. 2019) explored a series of alert messages such as "Stranger danger alert!", "Reminder! It's not safe to share your location with people you don't know!", with the goal of nudging children to take protective measures or to refrain from activities that might harm themselves or others. Similarly, Dempsey et al. (Dempsey et al. 2022) explored a series of designs on warning messages with children related to the disclosure of their private information online.

Social Feedback. Social feedback nudges attempt to alter children's behaviours, typically for self-regulation, based on the feedback or comment from others. For instance, *Plan&Play* (Hiniker, B. Lee, et al. 2017), a tool developed for supporting children's intentional media use, incorporated the design of a happy face of panda if children's current activity follows plan, and a sad panda face if their activity is off-plan. Other designs have also tried to display a privacy rating related to each options in order to alter children's choices online (Alemany Bordera, Del Val Noguera, and García-Fornes 2020).

3.5 Discussion

One of the key goals of this research is to create a landscape understanding of digital autonomy for children in the HCI community. Our analysis identified some critical factors to be considered in addition to the popular Spear et al's definition (Spear and P. Kulbok 2004; Hannum 2011; Lapidoth 1997) and some critical gaps in existing research.

While there is no consensus on what autonomy for children should look like (Silverberg and Gondoli 1996), Spear et al's definition on personal autonomy, based on a conceptual analysis of extensive literature, has been widely accepted by existing research around self-autonomy, and thus is considered as a great starting point for us to distill the more nuanced conceptualisation of digital autonomy in existing literature. We found that the majority of the research in HCI defined digital autonomy from the *self-regulation* perspective, emphasising the importance of *"supporting autonomy as triggering intrinsic motivation"* (Shin and Gweon 2020) and *"self-regulate through making responsible choices"* (M. Ko et al. 2015). This aligns closely with Spear et al's definition of 'behaviour autonomy' - the ability to make decisions independently and follow through on these decisions with actions. Meanwhile, our analysis identifies that more than half of the articles we reviewed have explored digital autonomy through supporting children's *critical thinking* and *computational thinking* abilities. Although many of these research have not provided an explicit definition of autonomy, they have identified supporting children's autonomy as their key goal and thus positioned their explorations within this context; in fact, their investigations align well with Spear et al's notion of 'cognitive autonomy' — the ability to acquire knowledge and understanding and to evaluate thought, voice opinions, and make decisions independently and to self-assess (Beckert 2007).

This review provides crucial insights regarding the conceptualisation of children's digital autonomy. We show that in contrast to the general expectation that digital autonomy is largely about self-governance or behaviour change (Rafael A. Calvo et al. 2014; Peters, Rafael A Calvo, and Ryan 2018), the ability of computational thinking and critically acting on these information is just as important. Indeed, research related to adolescence development has highlighted that the development of autonomy must be built upon a process of recognising, identity formation, to making independent choices (for example, independent of parental influences) (Neel, Jay, and Litt 1985; Paterson 2010). These previous research have described that the development of autonomy is influenced by multiple processes, including the cognitive processes of developing judgment and decision-making, negotiating social influences from peers or parent, as well as fostering healthy behaviour development (Chassin et al. 1995). Our conceptualisation analysis identified this focus on 'cognitive autonomy' support in the existing HCI research, although the conceptualisation of *cognitive autonomy* is probably less well-defined. By aligning existing HCI conceptualisations against the definitions from the other disciplines, we highlight the importance of considering the support of digital autonomy from a multi-dimensional aspect, which is essential to children's development process, and thus crucial for designing and building future digital autonomy support for children.

While our research provided a rich grounding for research around supporting children's digital autonomy in the behavioural and cognitive perspectives, our analysis also identified a critical gap in the existing HCI conceptualisation of digital autonomy - the support of children's development of *emotional* digital autonomy. Emotional autonomy refers to the ability to free oneself from emotional dependence (Lamborn and Steinberg 1993), and is a vital part of the social-individual relationship development of children. Various previous research in child development and psychology has suggested the important role emotions served to help people address or overcome problems and attain their goals (Keltner and Lerner 2010; Frijda and Mesquita 1994). Beyers et al. (Beyers and Goossens 1999) pointed out the importance of emotional autonomy for children through clarifying its key difference to behavioural and cognitive autonomy under a case study, suggesting that emotional autonomy of children would correlate more with their intrinsic motivation. Meanwhile, the lack of support of such autonomy in existing HCI literature is perhaps not surprising, as under the existing conceptualisation, digital autonomy is more positioned as skill development (e.g., one's self regulation skills, critical thinking skills as well as computational thinking skills), leaving the social-emotional aspect of digital autonomy under-explored. We argue that children's emotional autonomy development should be much more greatly emphasised, and be considered more jointly with their cognitive and behavioural autonomy development. It would be challenging for children to partake in digital citizenship or develop deeper digital literacy without social-emotional skills (Education 2017), as these are the essential process by which children understand and regulate their emotions and behaviours, and make responsible decisions (Children 2020).

3.6 Data Autonomy: A Working Definition

So far, we've explored the intricate concept of children's autonomy in the digital realm based on current research. What has become increasingly evident is the omnipresence of **data** in this digital age. Every interaction, every click, every choice a child makes online leaves behind a trail of data. This data doesn't merely act as a passive record; it actively shapes their digital experiences, influencing the content they see and how they engage online. If we genuinely aspire to empower children in the digital world, it's not enough to simply grant them freedom to navigate; we must equip them with the understanding and agency over their own digital footprint, thus making a compelling case for the transition from digital autonomy to data autonomy.

Shifting our attention back to our thesis's core subject, we highlight the pivotal role data plays in children's AI-based systems. Building on the insights from our discussions on children's digital autonomy, we aim to refine our grasp of supporting their autonomy, specifically in their data interactions. To achieve this, we developed a working definition on ***data autonomy*** rooted in the earlier established digital autonomy, viewing it through the distinct lens of data:

3.6.1 Unpacking Data in Data Autonomy.

To better define the scope of the term "data" within our notion of data autonomy, we draw on Solove's taxonomy of privacy (Solove 2002) and conceptualisations around datafication (Cukier and Mayer-Schoenberger 2013; Zuboff 2019; Mejias and Couldry 2019). Solove's taxonomy organizes data concerns into four categories: information collection (observation and recording of activities), information processing (storage, manipulation, and use of data), information dissemination (breaches of confidentiality, harmful disclosure), and information invasion (intrusions into physical, psychological, or digital spaces and decision-making). Meanwhile, scholars like Cukier and Zuboff have furthered the concept of datafication from social science and business perspectives (Cukier and Mayer-Schoenberger 2013; Zuboff 2019; Mejias and Couldry 2019). The concept of datafication involves converting phenomena into quantifiable data for analysis, and is anchored in two key elements: first, the external infrastructure that enables data collection, processing, and storage; and second, the value-generation mechanisms, which include aspects like monetisation, cultural production, and civic empowerment. This framework is intricate and operates on a global scale, encompassing dissemination, access, storage, analysis, and surveillance, largely under the control of large corporations and states. Building on these frameworks, we distill three critical elements essential to the term "data", specifically in the context of online platforms:

- The first, *Data Collection*, defined by Solove as "the watching, listening to, or recording of an individual's activities" (Solove 2002). In the context of online platforms, this involves the gathering and storing of user information. A significant dimension within this element is data sharing – how platforms distribute the collected information to other entities, such as internal departments, partner companies, third-party vendors, or advertisers.

- The second element, *Data processing*, refers to the process in which digital platforms process, analyze and make use of collected user's data. In the context of social media and other online platforms, this involves using users' data to generate services and content, often supported by algorithms.

- The third element, *Data inference*, involves the further processing and analysis of user data by online platforms to evaluate or predict personal aspects, such as work performance, economic situation, or health. This aligns with Solove's categories of data dissemination and invasion, as well as the value-generation aspect of datafication. What sets *Data Inference* apart from *Data Processing* is its capacity to *learn* about individuals or groups about their personal aspects, going beyond simply processing data for services like video recommendations.

3.6.2 Unpacking Autonomy in Data Autonomy.

In line with our discussion on data autonomy, we aim to provide clarity on our interpretation of "autonomy" within the digital realm. Here, we continue to use the previously mentioned widely acknowledged classification of autonomy for adolescence by Spear and P. Kulbok 2004 as:

- The first, *Cognitive Autonomy*, which refers to an individual's ability to think independently. It involves self-governance of the mental action or process of acquiring knowledge and understanding, to evaluate thought, to voice opinions, critically evaluate information, and to form personal beliefs.

- The second, *Behavioural Autonomy*, which refers to an individual's capacity to act independently, make their own decisions, and carry out actions based on their personal judgment and values. It involves the ability to self-regulate, take responsibility for one's actions, and behave according to one's own decisions and choices.

- The third, *Emotional Autonomy*, which refers to an individual's ability to identify, understand, and manage their own emotions independently. This involves the capacity to distinguish one's own feelings from those of others, handle emotional dependence, and maintain emotional stability without relying excessively on others.

3.6.3 Data Autonomy: A Working Definition

Building on the aforementioned concepts on "data" and "autonomy" across multiple disciplines. We now present an overview as a working definition of data autonomy in the digital realm:

> *Data Autonomy*, can be summarised as the empowerment and capability of individuals to comprehend, exercise control over, and reflect on the collection, processing, and inference of their data within the digital realm. This concept underscores the importance of informed understanding, active decision-making, and critical reflection in the way personal information is handled and utilized in online environments.

In this chapter, we delved into the intricate notion of digital autonomy for children. Acknowledging the pivotal role of **data** in children's digital experiences, we refined these concepts through the specific lens of data, culminating in a working definition for *data autonomy*. Our goal of laying down such a working definition is by no means seeking for an all-encompassing framework or formal definition for data autonomy. Instead, at this stage, our goal is to tease out the key themes that are relevant to the concept of data autonomy that will guide our subsequent investigation.

Moving forward, this thesis will embark on an insightful journey, examining the intricacies of data autonomy through empirical investigations and technical evaluations. In the chapters to follow, we aim to initially comprehend children's perceptions of datafication practices on AI-based platforms (Chapter 4), and then understand their expectations for support in navigating these practices and associated challenges (Chapter 5). Moreover, we plan to evaluate the influence of two technical prototypes on their autonomy regarding their data (Chapter 6 and Chapter 7).

Statement of Authorship for joint/multi-authored papers for PGR thesis

To appear at the end of each thesis chapter submitted as an article/paper

The statement shall describe the candidate's and co-authors' independent research contributions in the thesis publications. For each publication there should exist a complete statement that is to be filled out and signed by the candidate and supervisor **(only required where there isn't already a statement of contribution within the paper itself).**

Title of Paper	12 Ways to Empower: Designing for Children's Digital Autonomy.
Publication Status	Published
Publication Details	Ge Wang, Jun Zhao, Max Van Kleek, and Nigel Shadbolt. 2023. 12 Ways to Empower: Designing for Children's Digital Autonomy. In Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23). Association for Computing Machinery, New York, NY, USA, Article 91, 1–27. https://doi.org/10.1145/3544548.3580935.

Student Confirmation

Student Name:	Ge Wang		
Contribution to the Paper	Ge Wang is the principal author who contributed the ideation, data collection, data analysis, and paper write-up. Jun Zhao, Max Van Kleek, and Nigel Shadbolt critically reviewed the manuscript and offered valuable feedback.		
Signature		Date	22/09/2023

Supervisor Confirmation

By signing the Statement of Authorship, you are certifying that the candidate made a substantial contribution to the publication, and that the description described above is accurate.

Supervisor name and title: Dr Jun Zhao, Senior Research Fellow			
Supervisor comments			
Signature	Jun Zhao	Date	19 March 2024

This completed form should be included in the thesis, at the end of the relevant chapter.

Publications arising from this chapter:

- *Ge Wang, Jun Zhao, Max Van Kleek, and Nigel Shadbolt. 2022. 'Don't make assumptions about me!': Understanding Children's Perception of Datafication Online. Proc. ACM Hum.-Comput. Interact. 6, CSCW2, Article 419 (November 2022), 24 pages.* `https://doi.org/10.1145/3555144`. ***Best Paper Honourable Mention Award*** ★ ***Impact Recognition Award***

4

Understanding Children's Perception of Datafication Online

Contents

Here we address the thesis question: *In the current AI landscape, what insights can be gathered about children's understanding, perceptions, needs, and the obstacles they face regarding having autonomy over their data?* **(TQ1)**. Specifically, in this chapter, we delve into children's understanding and perceptions of datafication practices on the AI-based online platforms.

All studies in this chapter received ethical approval in accordance with university guidelines for human participant research, under reference CS_C1A_021_028.

4.1 Background and Motivation

The rapid adoption and increasing reliance of children on the online world and various types of algorithmic-based AI systems has raised corresponding concerns about the long-term effects of *datafication*, in which children's actions are pervasively recorded, tracked, aggregated, analysed, and exploited by online services in multiple ways that include behavioural engineering, and monetisation (Mejias and Couldry 2019; Zuboff 2019; Mascheroni 2020). Such activities take place invisibly behind the scenes of apps and services, and are less well understood or discussed as risks than other kinds of more easily characterised harms, such as the collection or disclosure of particular kinds of sensitive data. Given that most adults have little understanding of how their own data are being collected, processed, and used to shape their digital environments (Büchi, Fosch-Villaronga, Lutz, Tamò-Larrieux, and Velidi 2021), it is not particularly surprising that children too, lack a robust understanding or adequate mental models of how their data are processed or used (J. Zhao, G. Wang, et al. 2019).

On the other hand, there have been a variety of developments looking into how to support children developing their 'digital literacy' (Internet Safety 2020; Commission 2016; Institute 2021) as well as AI literacy (Druga, R. Williams, et al. 2018; Druga, Vu, et al. 2019; Long and Magerko 2020). However, such frameworks were often oriented around data privacy or online safety, with the algorithmic processing of data by online systems scarcely mentioned. Meanwhile, 'critical algorithmic literacy' (CAL) (Kafai, Proctor, and Lui 2020) puts particular emphasis on *understanding the implications of data processing*, by directing children's attention towards data and the algorithmic processes applied to them. The goal of CAL is not merely assisting the development of knowledge about algorithms but also an ability to engage in critique of algorithmic systems reflexively. The CAL framing proposed that computational thinking should include three key frames: the cognitive, the situated, and the critical thinking (Kafai, Proctor, and Lui 2020). *Cognitive thinking* focuses on the understanding of key computational concepts, practices, and perspectives and the associated skill building and competencies; *situated thinking* encourages learning to take place in contexts that the learner cares about so that they include their personal expression and social engagement in their pathway of learning; and finally *critical computational thinking* recognizes that computing is not an unequivocal social good, and emphasises the importance of supporting the questioning of larger structures and processes behind the computational phenomenon.

In this chapter, in order to address the topics of datafication and the thesis focus on empowering children with data autonomy. We feel that it is crucial to first understand how children currently interpret and perceive online datafication practices, pinpointing common barriers and knowledge gaps among them; especially since existing research has largely overlooked this specific child perspective.

For this purpose, we chose YouTube as our exemplary platform, given its widespread popularity among children worldwide (Ofcom.org n.d.; Foundation 2021) and its extensive datafication practices (Matamoros-Fernández 2017; Mitrou et al. 2014). We conducted one-to-one interviews with 48 children, aged 7-13, from UK schools undertaken between November and December 2021, identifying ***children's perceptions of the datafication practices*** on YouTube.

Our findings identified three key knowledge gaps in children's current awareness and perceptions of datafication practices online, including their lack of recognition of (i) who are involved in the data processing and how, (ii) data being transmitted across platforms, and (iii) their data ownership. Through situating our findings under a critical algorithmic literacy framework, our findings provided critical insights on how we could better support children in the datafied society through more transparency and autonomy-supportive designs.

4.2 Methods

4.2.1 Study Design

As explained above, we chose the YouTube platform to be used as an example, and conducted one-to-one semi-structured interviews with children.

As part of our child-centred approach, we drew inspiration from Druin's influential model for integrating children into the design process (Druin 2002). In her model, children can adopt various roles such as users, testers, informants, and full-fledged design partners. While the first two roles emphasize gathering feedback from children towards the end of the design cycle, the latter roles focus on deeper collaboration, viewing children as equal stakeholders in designing new technologies and valuing their contributions throughout the process.

For our study, we positioned children as *informants*, seeking to understand their experiences and perceptions concerning YouTube's datafication practices. We opted for semi-structured interviews because of their inherent flexibility. This method captures the richness of children's viewpoints, allows them to share unexpected insights, ensures their unique experiences aren't sidelined, and truly emphasizes their individual voices in an authentic context. By adopting this approach, we encourage a dialogue led by the child participants, granting them the agency to guide the conversation towards topics they find significant.

During the interviews, we elicit their responses to a collection of tasks that attempted to recreate their everyday experiences on YouTube, followed by a collection of scenarios that reflected different types of datafication practices in relation to how YouTube could process and make use of their data. Each interview session spanned approximately 60 minutes and was divided into two parts:

Part 1: YouTube Tasks

In part 1 of the interview process, we chose to walk through a series of tasks with participant children on their own devices (Figure 4.1) for the purpose of recreating and reminding them of their everyday experience on YouTube. In this process, children were encouraged to actively interact with YouTube while being instructed to solve a series of tasks.

Figure 4.1: Children were instructed to complete three tasks on YouTube.

Our problem-solving tasks were carefully designed based on 'critical interaction points' on YouTube, and were divided into three sessions: In *Task 1*, children were asked to show the researchers *how they would normally find their favorite videos on YouTube*. Entering the YouTube is a key interaction point and task 1 aimed to observe how children would go to their favorite videos (such as through a subscription list or just through searching function); in *Task 2*, children were asked to show the researchers *how they would find a video on a specific topic*. We consider the search function as a 'critical interaction point' as it's the most important functionality on the YouTube platform in order to find a video to watch. Here task 2 aimed to observe how children made use of the search function and subsequently, how children would choose and decide which video to click on from the search list; in *Task 3*, children were asked to wait for a current video to finish, and then show the researchers *how they would decide what to watch next*. A few 'critical interaction points' were involved in this task, including the 'next-up' videos that show up after a current video is finished, the 'videos on the right' which includes a list of videos recommended to children based on different categories of information (e.g. the genre of the current video watched, other videos from the same YouTuber, videos based on the watching history), as well as the personalised advertisements that show up at the beginning of the next video. Task 3 aimed to observe how children would normally react to such choices and examine their perceptions on such recommendation practices.

As children were going through the three tasks, they were also invited to explain their understanding about and experiences of certain technological terminologies (*homepage, autocomplete, autoplay*, and *personalised adverts*), if they appeared during their tasks, to evoke further discussions around their perceptions on the relevant data practices. For example, when adverts appeared during their completion of the task, we asked about their perceptions on personalised adverts and how they dealt with them. During the process, we tried to use the tasks to encourage children to recall their experiences as much as possible. Instead of focusing on discussing what were returned, we particularly focused on asking children questions about *how would you* or *what do you do then.*

Part 2: Video Scenarios

In the second part of the interview, we chose to present the children with two videos about a fictional character named Lola, a 10-year-old girl who likes to watch videos on YouTube, and how she learns about data collection and processing on YouTube (see Figure 4.2). In the videos, we used metaphors and compared the collection and processing of children's data against throwing bottles of elixir into a magic pot. Previous research showed that children as young as 5 can start to comprehend metaphors and could provide verbal explanations for metaphorical expressions (Rowe, Özçalışkan, and Goldin-Meadow 2008), and that metaphors and stories are effective way of building children's understanding of abstract concepts (Cameron 1996; Billow 1981).

Each video lasts about 1.5 minutes and was shared with participant children through a link to each video. After watching each video by themselves, children were invited to comment on specific plots presented in the video. A screen capture as well as the question were presented to the children through screen share by the researchers so that children can recall the content, and the questions were expressed in a language appropriate for the participant's age and development.

We have not intended to design or introduce the videos for education or learning purposes because the study was relatively short, lasting about 1 hour. Instead, the videos were mainly there for invoking discussions.

Video 1 - *General Perception of Datafication Online.* This video pictures YouTube as 'a magic pot' that requires access to a range of data in order to perform its magic, including video we watched, terms we searched for, websites we visited, our friend lists and our location. After having watched the video, children were prompted to articulate on how they perceive the general datafication practices on YouTube. Specifically, whether they were surprised by or happy about the data being collected by YouTube, and what do they think will happen to their data.

Video 2 - *Perception of Data Inference Online.* This video provides more details about *what* magic can be performed by YouTube and *how*. In this way, we talked about the data inferences performed by YouTube, using data collected by themselves as well as those from other platforms (such as web search history on other sites). Then, we invited children to think about and articulate on the data inference practices: specifically, whether they were surprised by or happy about how YouTube use different types of data about them to recommend new videos, learn more about their personal life, or send them adverts more personalised to their interests. In this way, we focused on the data inference part of the datafication process - online platforms could learn more subtle things about users (apart from interests) based on their data. This refers to the likelihood that YouTube could infer about a child on a more personal level, such as inferring their age, their socio-economic status, and more.

4.2.2 Study Method

Participants were recruited from local primary and secondary schools, and a public forum for recruiting family participants. Recruitment started in November 2021 after obtaining ethics approval, and 48 children were interviewed between November and December 2021. Participants were given the choice to take the study either online or in-person due to Covid restrictions. 46 children were interviewed online and 2

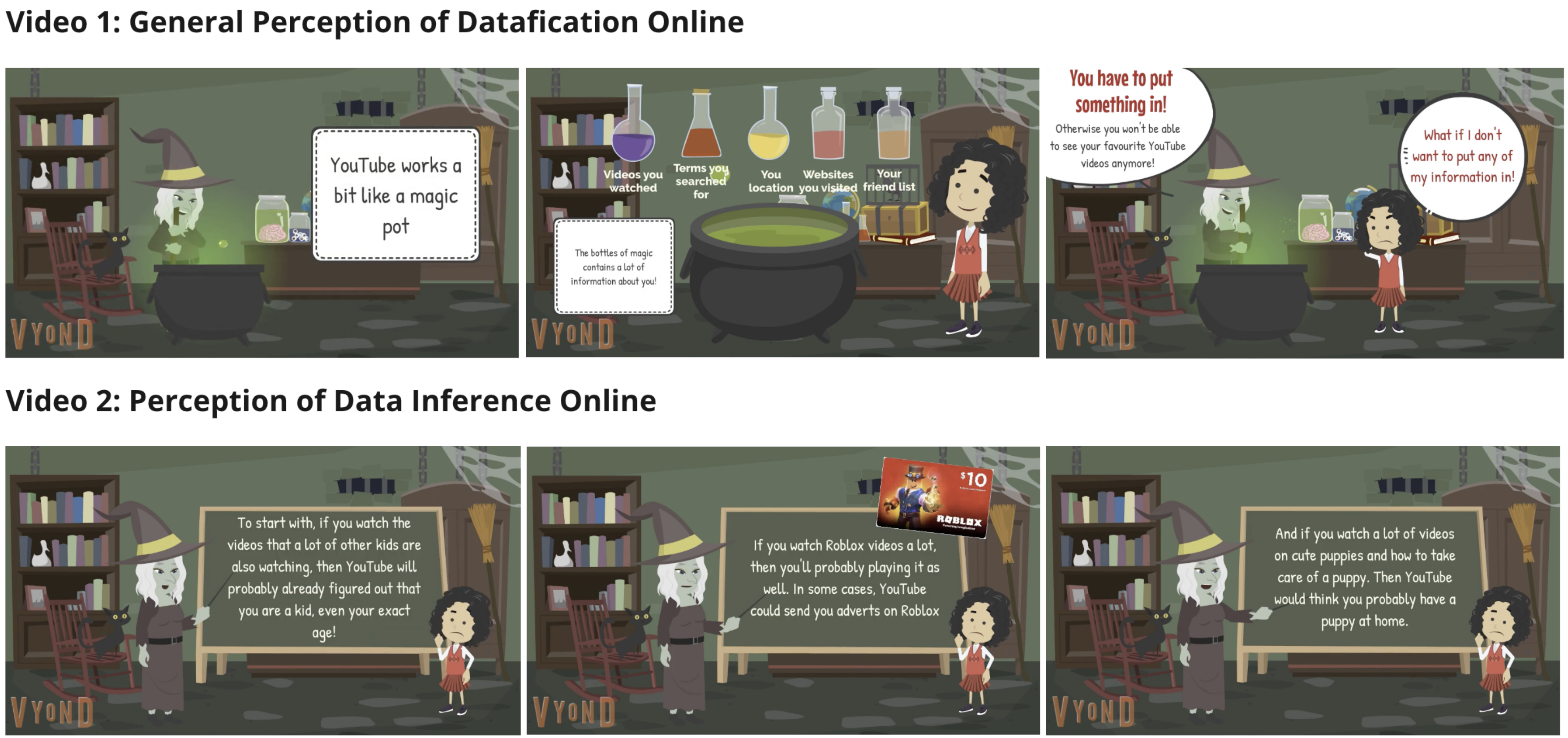


Figure 4.2: Screenshots of video 1 and 2. Video 1 is about children's general perceptions of datafication online. Video 2 is about children's perception of data inference online.

were interviewed in-person. At each of these sessions, parents helped children to set up their devices and then left the children alone until the interview was completed. The majority of the participants were recruited through school newsletters. Each study was facilitated by at least two researchers.

Participant Information

We had 48 participant children, including 23 boys and 25 girls. 24 were among the age range of 7 to 10, another 24 were among the age range of 11-13, with an average age of 10 (range = 7-13, s.d. = 1.76). Details about participants can be found in Table 4.1.

Age	#Boys	#Girls	#Total
7-yo	2	2	4
8-yo	3	1	4
9-yo	5	5	10
10-yo	2	4	6
11-yo	4	6	10
12-yo	5	6	11
13-yo	1	2	3

Table 4.1: Summary of participants' ages and genders

Study Process

Each interview contained four parts, including an introduction session, Part 1 - a walk-through of tasks on YouTube, Part 2 - a walk-through of videos on data practice scenarios, and finally an open-ended session about children's thoughts and needs as well as any issues not so far discussed. The whole study was planned to last around to approximate 1 hour.

Children were asked to bring their most used device when accessing YouTube to maximally recreate their everyday experience. In Part 1 of the study, children were given some time to read through the task sheet given. One of the researcher read along the tasks with them, and helped them throughout the process. As they were completing the tasks, researchers observed their activities and asked questions to invoke discussions. Children were told there are no “right answers” to whichever task they were on or question asked. In Part 2 of the study, children were shown with two videos on datafication practices scenarios. One of the researcher watched the videos along with them, and offered clarifications if children had trouble. After they finished watching each video, screenshots of the video were shown along with questions, to remind children of the content of the video. While children were first asked to respond to the predefined questions, like *Are you happy to share some information than others*, our researchers followed up any responses that required further clarification and we encouraged children to share their personal experience or values related to the scenarios, by asking questions like *has it surprised you* and *what would you do then*.

Screen and audio recordings were taken during the studies. Screen recordings on children completing tasks on YouTube were played back during the transcription and data analysis phases to highlight any notable patterns.

4.2.3 Data Analysis Method

Our analysis began with the transcription of interviews conducted with children. These transcripts served as the primary data source for our study, capturing the rich, detailed accounts of the participants' experiences and insights. To analyze these transcriptions, we adopted a grounded, thematic approach, as recommended by Braun and V. Clarke 2006. This methodological choice was informed by its robustness in identifying, analyzing, and reporting patterns (themes) within qualitative data, thus providing a flexible yet detailed tool for qualitative analysis. Below, we provide a formal description of the methodology:

Thematic Analysis Process. The thematic analysis was conducted in several stages to ensure a systematic and comprehensive examination of the data:

1. Familiarization with the Data: Initially, we immersed ourselves in the data by reading and rereading the interview transcriptions. This step was critical for gaining an intimate understanding of the data and for identifying preliminary insights and patterns that would guide the subsequent analysis.

2. Generating Initial Codes: Following the familiarization phase, we began the systematic coding of the data. This involved identifying and tagging relevant features of the data that appeared to be of interest concerning our research questions.

3. Searching for Themes: With the initial codes in place, we proceeded to organize these codes into potential themes and sub-themes. This process was not merely about categorizing data; it involved interpreting and making sense of the underlying ideas and patterns that emerged from the coded data. The search for themes was both an inductive and iterative process, allowing for the refinement of themes as we engaged more deeply with the data.

4. Reviewing Themes: Once potential themes were identified, we reviewed these themes in the context of the coded data and the entire dataset. This involved a critical evaluation of whether the themes accurately represented the data, leading to further refinement of the themes. This stage ensured that our themes were coherent, consistent, and meaningful.

5. Defining and Naming Themes: The final step in the thematic analysis involved defining and naming the themes clearly and concisely. This required a detailed analysis of each theme and its relationship to the overall data set, ensuring that the themes effectively captured the essence of the data related to the research questions.

More specifically, results from part 1 of the study contained both children's experiential descriptions about their own experiences of using YouTube, invoked by the tasks, and their perceptions of key tasks. So, we identified the experiential reflections mentioned by the children, and then carefully coded the specific synchronic elements mentioned (such as their emotions, sequence of actions, and information used to inform their action) by the children during their completion of each specific task. In this way, we could gain a more in-depth understanding about how participant children currently managed critical interactions points on YouTube and what elements of knowledge were used by them. This gave us a set of codes about children's usage patterns, their existing knowledge points, and their general expression of experiences about YouTube.

With respect to the data from the video-guided interviews, we tried to calibrate how children perceived YouTube's data processing practices, from the general datafication to the more specific data inference practices, by carefully examining their use of language, for example, how they described data, those who collected their data, what and how the data was processed etc.

4.3 Results

We present our results by first outlining children's overall experience and usage of YouTube. We then present children's general perceptions of datafication online, followed by an in-depth analysis regarding children's perceptions of the more specific data inference practices. Finally, we present our analysis on children's coping strategies and the design changes they want. While our participant children demonstrated different perceptions and varied level of understanding, we found no strong differences between children of different age or gender. We present individual children's quotes with their participant id.

4.3.1 Children's Overall Experience and Usage on YouTube

Most children (37/48) in our study owned their own devices (phone, tablet, computer), while the rest used their parents' devices or shared with their siblings. A fair amount of children (19/48) had and would sign in to their own YouTube accounts [1], 7 children used their parents' YouTube accounts, while the remaining 22 children did not have their own account and would not sign in to an account at all on YouTube. In terms of usage, around a third of children in our study spent over 5 hours per week on YouTube, 11 others spent 3 to 5 hours per week, and the remaining 21 children spent 1 to 3 hours per week on YouTube. Children in our study were generally attracted to game videos, animations and educational videos on YouTube.

[1]The minimum age to have an account is 13 years-old, as claimed by YouTube (YouTube 2021).

In general, most children (41/48) reported that they started their use of YouTube using the search button. They generally found 'autocomplete' quite useful for their search, and a few children (17/48) reported that they thought their search intentions could be influenced by the autocomplete list. However, the majority of them (41/48) reported that they were able to decide which search terms to use. The autocomplete could be helpful to remind them of something they searched and they generally thought themselves unlikely to be 'persuaded' by these terms.

Once the play of a video was completed, the majority of the children (35/48) reported that they knew what autoplay is, they commonly experienced this and they demonstrated different ways to respond to autoplay. The majority of the children (29/48) said they normally wouldn't pay any attention to the autoplay function, thus let the video to automatically go to the next one ('I think I won't do anything' - P20). Some children (23/48) said that they would only follow the videos in autoplay if they were from the same series that they were following ('Well if I'm watching a series then yes, cause the next video would be from that series' - P29); while others (25/48) said that they were more likely to choose what to watch next from the list on the right, instead of permitting autoplay ('I would just pick one from this list' - P4). However, only a few children (11/48) knew how to turn off autoplay.

A fair amount of children (19/48) had their own YouTube accounts and all of them reported they knew or have used subscriptions. Almost all children (46/48) reported that they knew what advertisements are and they generally disliked advertisements on YouTube, mainly because the advertisements being 'not interesting', 'not useful', and 'unskippable'.

In general, most children (38/48) reported having a positive experience on YouTube, for YouTube 'offering fun videos to watch', and 'I can learn stuff on YouTube'. The remaining 10 children said they had a neutral experience on YouTube, mainly because of 'boring advertisements', and some reported seeing 'weird stuffs' - inappropriate contents on YouTube before. None reported their experience on YouTube as negative.

4.3.2 Children's General Perceptions of Datafication Online

To start with, we present our findings on children's general perceptions of datafication practices on YouTube. Specifically, how children perceive what data is collected and what happens to these data.

Data Collected: Essential? Personal? And Owned by Whom?

All children knew that YouTube would collect data from them, however, they had different perceptions in terms of what data was being collected, and such perceptions were often related to whether they think such data was essential or personal.

To start with, all children we interviewed knew that data such as the videos they watched before, and terms they searched would be collected by YouTube. And again, all of them were perfectly fine with YouTube doing so, because they thought these data were *essential* for YouTube to generate videos for them:

> I'm not surprised by that (YouTube collecting videos watched and terms searched) at all, and I'm fine with it. Of course they would need that to provide videos for us. I don't see how else they could do that. (P10)

On the other hand, children were more reluctant for YouTube to collect data they think that were 'personal to them', and many did not think YouTube was already collecting such data as they could not see why their personal information such as their location information, their age and gender, would be needed for generating videos:

> Definitely not the location. I feel like the location is personal to you, and honestly I don't know why they would need that. (P3)

There was a strong theme of discussion around what data children considered as their 'personal data'. Interestingly, while almost all children strongly believed that data such as their location information and their age and gender were their own personal data, a majority of children regarded their behavioral data (e.g. videos watched, terms searched, channels subscribed) as 'not personal to me' and therefore could be accessed and quantified by platforms:

> Personal data is something that is not already on their platform, like my location. But I don't think the videos watched and terms are that personal to me, as they are on their platform already. (P16)

Furthermore, we found an interesting common thinking among children regarding who owns their behavioral data online (e.g. videos watched, terms searched, channels subscribed). A surprising amount of children (32/48) thought that it was YouTube (or the platform) that own these data, or have the rights to these data:

> YouTube should have the ownership. Because YouTube are providing these videos? They should have the right to collect these things I guess, I mean, these are like their thing. (P37)

> I don't know. It feels like I kind of technically own them. But it's kind of like YouTube's right to know and control them? Because we are on their platform? (P21)

In response to our question about "whether you are happy to share these information" or "Are you surprised", the majority of the children were happy to share data they regarded as essential for YouTube to function, but less happy about those they regarded as 'personal' or 'non-essential for YouTube':

> I'm happy to share videos I watched, terms I searched before, channels I subscribed. But definitely not my location, that is personal to me. (P14)

> The others are all fine cause I see why YouTube would need that. Location? No. I can't really think of how that would help them. (P31)

Data as Part of a Process: (Only) Used to Provide Better Services for Me

The majority of the children perceived that the data were used by YouTube to provide services for them, within YouTube:

> I knew they(recommended videos) are based on what I watched before and I think it's pretty smart of them to do so. And I always get videos I like. (P34)

Meanwhile, children also demonstrated this perception about the exchange of data for services on other online platforms. For example, some children mentioned that they thought Amazon would collect data such as their purchase history and search terms to recommend products to them; and some children also mentioned their experience on apps stores and games stores, that they thought platforms would collect what apps or games they previously downloaded to personalise the things recommended to them:

> It's just like Amazon right? They will know what stuffs I searched or bought, and then send me similar products. (P31)

> This is a suspicion, but I think Steam is doing that as well? I remember seeing somewhere that they would use stuffs like what I played before to recommend more games for me? (P48)

On the other hand though, while children had good understanding on how YouTube, and other online platforms would provide service for them using their data, most children thought that the *only* reason why online platforms would collect and process their data was *just for* offering them better services. Only very few children (7/48) mentioned data could be used by the platforms for monetisation. However, their understanding of data monetisation were still limited to - platforms will use their data to provide better service for them/users, so that websites will have more users/engagement, thus making more money:

> If they offer better service, they would attract more users, or longer watch time, then they will make money. (P6)

> The more people they have on their website, they could show that to their sponsors, and the companies would pay YouTube money to put up more advertisements. (P41)

And when children thought the platforms were just trying to offer them better services, they were generally fine with and perceived such monetisation as fair, because they were receiving better services:

> It's pretty fair. They are just trying to give you a better experience, and they are doing a pretty good job. I think they should earn some money. (P29)

> I don't really mind if they use my data to make money. It's a win-win situation actually. At least it's useful to me. (P11)

Data Silos: My Data will Only Stay on Where I Go

A key theme emerged was children's perception about where these data about them went: the majority of them believed that the data were collected by a certain platform and would be only used by this platform locally.

> They collect your data on YouTube, and used that on YouTube. I think it's pretty obvious? I don't think the others (other platforms) could use YouTube's data. (P29)

Some children showed some awareness about data being shared across platforms but only very few of them (5/48) were able to articular how such data sharing might be carried out:

> Maybe if you have YouTube tab open, and you change that to a different website. They might be able to track you to the next website. (P46)

Interestingly though, a fair proportion of children (22/48) thought data sharing/tracking was enabled through Google:

> Because YouTube is owned by Google, and YouTube and Google could see whatever you do on each. But if you have a different search engine then YouTube won't be able to track you. (P25)

Even for the few children who had some notion about the general data sharing practices, they struggled to picture how the data flow took place. And they commonly thought that their data would only be transferred among the websites they actually visited:

> I think it (my data) would only go into all the websites that I usually use. (P16)

> It's like footprints. I will only have footprints in places I go to, but not beyond them, like other places. (P28)

It was perhaps thus not surprising to see that a fair amount of children (19/48) thought that they could just delete the data they left on the Internet by simply uninstalling the apps they usually use:

> I guess I could take back all my information if I just delete the apps? (P11)

4.3.3 Children's Perceptions of Data Inference Online

Following children's perceptions of the general datafication practices online, we further investigated their understanding and perceptions of the data inference practices online.

What is Data Inference

Continuing the notion about data as being an element and part of a process, most children described data inference as the process in which YouTube is trying to makes guesses and assumptions about people:

> YouTube does not actually know any facts about you. So it has to guess about your age, what kind of school you go to, any news in your home. (P45)

Some children thought that data inference also included categorising people into certain groups based on what is inferred about them:

> It assumes about you, and put you together with the majority. (P48)

> They would categorise you in different groups, so that they could do things in each different groups. (P19)

While most children described data inference as 'making guesses' and 'broad assumptions' about them, there were a few children believed that data inference would enable YouTube (and other platforms) to know exact factual details about them:

> They will try to learn everything about you. And with all the information they have, they could even somehow know my name, and everything about me, what jobs my parents do etc. I don't want this but I think it's unlikely for them to get things wrong, with all that power they have. (P14)

How is Data Inference Conducted

Children showed different understanding in terms of how a data inference process was conducted. We summarised their understanding into four major themes: *operators* - who/what was conducting the data inference, *inputs* - what was the data inference based on, *algorithmic processing* - what was the (computational) process happen in between, and *outputs* - what were the outputs of data inference.

Starting with the *operators*, the majority of the children reported that they believed that it was not 'human' that 'manually' processed their data. They would describe the operators as some kind of 'machine' using terms such as 'a computer', 'a robot', 'an AI' or '010101'. On the other hand, only a minority of children (11/48) had some awareness that algorithms were used to process their data:

> I doubt it if just some person just sitting there doing it. So it's probably some kind of algorithm. (P7)

For these children, they showed varied perceptions about what an algorithm was. Some of those believed an algorithm was written by human, and followed a set of rules:

> A person wrote the algorithm, like they told the computer what to do. An algorithm just do what it's coded to do. (P24)

Interestingly, when trying to explain how algorithms follow a set of rules, many of these children did not recognised the 'data-driven' aspect of algorithm, but instead described algorithms as following very exact rules. For example, when asked about why they think algorithms could generate gender-stereotypical contents for people, they tend to think that algorithms were set to only give certain content to certain gender:

> I think it (YouTube algorithm) might be doing that intentionally, like it's set to only give certain contents to certain groups of people. (P3)

That said, many children described there was data *input* for this machine/algorithm, and the algorithms were closely based on what data they knew: they generally thought the data collected on YouTube (e.g. videos watched, terms searched) could be used to make inference about them:

> I mean if you only watch yr7 videos on YouTube, they would probably guess that you are a kid. (P27)

> Depending on what kind of videos you watched? Like if you watch videos in another language, it's pretty obvious to them that you speak that language. (P29)

On the other hand, many of these children reported that they thought only their data resulting from their interactions on a particular platform would be used to make inference about them. To be more precise, almost all children only mentioned the possibility of YouTube using data generated from their interactions on YouTube (e.g. videos watched); none, meanwhile, mentioned the possibility of YouTube using other data sources, such as what they did on other websites. This agrees with previous findings from studies with both adults and children, that individuals perceive data in silos on each platform separately and have little awareness of cross-platform tracking and information sharing:

> I don't really think that's (data inference based on cross-platform data) possible? Well I guess it's possible if they want to. Emm, I never really thought about this... (P34)

Finally, as for the *outcome* of data inference, many children mentioned the concept of a 'profile'. And they described such profiles as a list of things, or records about them:

> I guess they would have a profile about you, like what type of person they think you are. It's different from what they think you might be interested in. I don't know, it's hard to describe. (P18)

How Children Perceive Data Inference Practices - When Reflecting on Hypothetical Scenarios

During our interview, we used two hypothetical scenarios to establish a deeper conversation with children about their perceptions around how platforms could make inference about them. The first scenario illustrated how YouTube would send personalised ads on things of different prices to different children; and the second one showed how YouTube would send advertisements on Barbie Dolls and makeups to girls, and send advertisements on science books and robotic toys to boys. When reflecting on these scenarios, children showed strong oppositions against platforms conducting data inference on them, for a variety of reasons.

To start with, children had strong opinions and generally believed that platforms shouldn't make assumptions of people, arguing that such practices could be used against them or people like them:

> Don't make assumptions about me! They're just kind of saying that some people should spend more money than others. (P19)

> This is just stereotypical. They are trying categorise people by their gender and make guesses on them. But those guesses could be wrong. (P48)

Children further described such data inference practices spying on them without them knowing:

> It's rude. It's just rude. There are lines. In way, what YouTube does is technically spying on your personal life. (P19)

> It's like you go into someone's house without them knowing, take stuff away from them. YouTube is just doing that. (P18)

They then continued to argue there exists an unequal relationship between them and YouTube. When being asked about what they thought about the trade-off between YouTube trying to learn so much about them, while at the same time, YouTube was offering them videos to watch, they thought what YouTube took from them was more than what they gained on the platform:

> It's just unfair. The videos they give us might not be useful to us, but the things they learn from us is useful to them. (P40)

Some children continued to argue that YouTube's only job should be providing videos, and it didn't need to learn so much about them:

> I'm just here for the videos. Their only job is to generate videos. It's completely unnecessary for them to know anything about me. Well maybe just what videos I watch, but that's it. (P23)

How Children Perceive Data Inference Practices - When Reflecting on Their Own Experience

However, when being taken outside the scenarios and asked to reflect on their own experience. It was interesting to see that, the exact same children who demonstrated different feelings and values, as well as critical thinking before (under the hypothetical scenarios), were unable to relate such thinking to their own everyday experience on YouTube. For example, many of them did not think or weren't aware that data inference practices could happen in their own lives:

> I don't think they are actually trying to do that though, cause I don't really see any stereotypical content. (P28)

> I actually never thought about that. Now that I think about it, they could be doing that? It's pretty scary. (P46)

And when being further asked about why YouTube, or more generally, platforms would try to infer about them and learn things about them, the majority of the children thought this was dominantly used for "making the website more popular" or "offering me better videos to watch", even though they demonstrated different feelings when they were shown how different 'values' could be generated by the inference process:

> They try to learn things about users because they want to give them better services. And if you offer better services, your website would become more popular. (P2)

> It's because so that they will know what's new in your life and offer you better videos to watch. (P35)

And when reflecting on their own experiences, many children reported that they didn't care as much because they didn't feel such data inference, if any, would make any difference or affect them in their real lives:

> To be honest, I don't really care. It's not gonna influence me in any way even if they learn things about me. (P40)

On the other hand, for the children who had some real life experience in terms of how data inference could have impacts on them, they were more alert of such practices:

> I definitely don't want them to know about me. I remember my parents mentioned that there was this time they tried to refurnish the house, maybe they have searched for something on that? And then they started to get calls from agencies, like they know they own a house and asking whether they want to rent out the house. I can't remember exactly but it's really creepy. (P43)

4.3.4 Coping Strategies and Desire for Legibility and Control

When asked about how they managed their data, the children revealed several coping strategies, for dealing with the undesirable aspects of these datafication practices, which revealed a number of common concerns and perspectives among children.

Don't Know How to Cope

During our interviews, children mentioned several coping strategies that pertained to their data privacy, specifically around how to stop data collection from websites and how to better protect their personal information online. These strategies included turning off location tracking, and adjusting privacy settings.

However, there was little mentioned (or that children could do) to help them address how their data were used by the algorithms. They expressed confusion and frustration at their inability to stop, or defend themselves more directly against certain data processing practices, describing them as 'more subtle movements':

> So like at school, we've already definitely learnt not to just randomly give out your age, your personal details and stuff. But with these subtle movements from YouTube and probably other companies, I don't know what to do. (P34)

Some children mentioned they tried to read up more carefully on the consent agreements and terms of service, but found it frustrating that data processing practices weren't made clear in either, and that such practices were impossible to manage or control from their end:

> They don't really talk about these things you know, well, at least they weren't very explicit. (P9)

> It's just impossible for me to deal with it. What can I do? Maybe I can hide some data from them, like my location. But I can't control what happens after they got the data. (P14)

Desire for Legibility and Control

On the other hand, when discussing future mechanisms, almost all children (46/48) mentioned that they wanted to be able to stop their data from being processed and used in ways that were unknown to them, mainly through two types of mechanisms - transparency and control.

To start with transparency, almost all children (41/48), regardless of their age, talked about how they would like to have information regarding the data processing practices: not only about what data of them is being collected, but more importantly, how their data will be processed. For example, some children described a transparency mechanism in which they would be able to see how a certain video or advertisement was selected for them:

> Cause it just feels more comfortable knowing what they're really doing. Like how they generate this video, is it based on something I purchase on another website? Maybe they're doing something extra they don't tell you. (P32)

Some other children mentioned they wanted to be able to see if YouTube had a profile about them, and what YouTube put down about them in the profile:

> Yes I do want to see what they think about me. Probably be happier if they get things wrong though. I'd rather they didn't know everything about me. (P31)

Meanwhile, children also urged for more control mechanisms: more specifically around how they could more directly control what platforms could infer about them, and how platforms would conduct this inference process:

> I want to have like a settings or something like that, that I can adjust what they could learn about me. Maybe adjust what they could use to learn as well. (P5)

Some children just said they wanted all platforms to stop data inference from the beginning, and instead only focus on 'their job':

> To be honest, I just want them to do their job and nothing else. Just stop learning things about people cause that's not their job. (P28)

Pessimism About Platform Change

However, nearly all of the children (45/48) we interviewed expressed pessimism that YouTube, and other online service platforms in general, would change anything about their current practices. Several children discussed that they thought nothing would change, as such datafication and inference online had become such established and ingrained, and it would be hard to change unless there were regulations on it:

> If there's a law on it then yes. But if no they'll just carry on doing it. (P26)

And a surprising amount of children (37/48) thought no changes would be made, for the reason that companies would only care more about money:

> Well, I guess it impossible. Because they just want to make more money, to make their business bigger, and to get wealthier. They just care more about money. (P7)

4.4 Discussion

4.4.1 Key Findings and Contributions

Children are often regarded as not as capable or competent as adults for coping with the complexities of online life, including aspects of privacy, safety, and datafication (Sonia Livingstone 2018), viewed as 'diminished versions of adults' (Pangrazio and Cardozo-Gaibisso 2021). However, our findings showed that not only do children care significantly about various aspects of datafication, but they already possess rudimentary conceptual understandings of it, and a significant willingness to take action to shape it to their desires, possibly even more than adults. We found that children were well aware of the data collection practices on YouTube, especially on data such as the videos they watched before, and terms they searched. They largely knew that these collected data will be processed, and generally understood that such processing is not conducted by human, but through some kind of machine. They had some awareness in terms of online datafication practices would "make guesses on them", and the results of such guesses could be used on users such as themselves, although not necessarily understanding the full picture. On the other hand, our results reinforced existing findings that children do not always comprehend online datafication practices to a full extent (Stoilova, Sonia Livingstone, and Nandagiri 2020; Sonia Livingstone, Stoilova, and Nandagiri 2019; Pangrazio and Cardozo-Gaibisso 2021), and through a knowledge construction lens, we identified three key knowledge gaps in children's perception of data inference practices, including their lack of recognition of *who are involved in the data processing and how*, *data being transmitted across platforms* and *their data ownership*. These findings provide critical contributions regarding our understanding of how 7-13 years old children perceive datafication in their everyday platform; instead of looking broadly at children's understanding of 'what an algorithm is', our study provided deeper insights regarding children's ability to interpret the implications of data processing and their barriers. These findings also provided important future design directions for supporting the development of children's algorithmic literacy and self-autonomy.

4.4.2 Unpacking Our Findings Through the Lens of Critical Algorithmic Literacy

Although the focus of our study was to explore how children experience and perceive a data-driven algorithmic platform, our findings also provided crucial insights for the recent 'critical algorithmic literacy' (CAL) development, confirming the urgency of extending children's cognitive computational thinking. The CAL framework also provides a useful guidance for us to interpret the observed gaps of knowledge for devising key future design considerations.

There have been a variety of developments looking into how to support children in achieving 'digital literacy' (Internet Safety 2020; Commission 2016; Institute 2021; UNESCO 2019). However, such frameworks were often oriented around data privacy or online safety, with the algorithmic processing of data of online systems scarcely mentioned. There has been some work from researchers around AI literacy (Kahn et al. 2006; Heinze, Haase, and Higgins 2010; Druga, R. Williams, et al. 2018; Druga, Vu, et al. 2019; Long and Magerko 2020). However, such work generally pertains to AI systems (machine-based systems that could make predictions, recommendations, or decisions that influence real or virtual environments (Yeung 2020)), rather than the more specific data-driven algorithms. On the other hand, 'critical algorithmic literacy' (Dasgupta and Hill 2021; Aleman et al. 2021) puts particular emphasis on *understanding the implications of data processing*, by directing children's attention towards data and the algorithmic processes applied to them. The goal of CAL is not merely assisting the development of knowledge about algorithms but also an ability to engage in critique of algorithmic systems reflexively.

CAL is closely aligned with recent extension of computational thinking. Kafai et al. (Kafai, Proctor, and Lui 2020) proposed that computational thinking should include three key frames: the cognitive, the situated, and the critical thinking. *Cognitive thinking* focuses on the understanding of key computational concepts, practices, and perspectives and the associated skill building and competencies; *situated thinking* encourages learning to take place in contexts that the learner cares about so that they include their personal expression and social engagement in their pathway of learning; and finally *critical computational thinking* emphasises the importance of supporting the questioning of larger structures and processes behind the computational phenomenon. Our key findings provided the crucial empirical evidence for the need of situated and critical thinking in children's algorithmic literacy.

To start with, our findings resonates the emphasis that CAL should go beyond the basic *cognitive* understandings and *situated* thinking will complement children's understanding of the social aspects of algorithmic system. For example, although most of our study participants were able to recognise different types of data being collected by YouTube and how they were processed to provide better video recommendations, they struggled to situate these understandings in the complex and diverse social contexts under which data may be collected, processed, analysed or exploited. Supporting children to connect with different algorithmic situations is the key objective of *situated thinking*. Indeed, previous studies have shown how children's understanding of privacy and algorithmic implications could be boosted by participating in carefully designed computational tasks involving the sharing and processing of social media data (Dasgupta and Hill 2017; Hautea, Dasgupta, and Hill 2017). We can envisage that similar experiments could be designed to enhance children's situated thinking regarding the sharing of data across platforms and what their data ownership means under different contexts.

Other findings from our study demonstrated the importance of introducing *critical* thinking in algorithmic literacy. For example, many of our participants struggled to perceive their data were shared across different platforms and how these data from outside YouTube could affect their experiences on the platform. They were unable to perceive themselves and their data as part of one connected data-driven digital society, in which an extensive amount of personal data about each individual can be processed in various unanticipated ways. This ability of situating datafication in a broader digital society is rarely discussed in existing research of algorithmic literacy (Aleman et al. 2021; Pangrazio and Cardozo-Gaibisso 2021; Long and Magerko 2020). However, *critical thinking* promotes an understanding of the existing structure of power in order to increase children's awareness of ideologies, privilege, and opportunities, and several research have shown how the approach can push children to "conceptualise, create, and disseminate digital projects that break silences, expose important truths, and challenge unjust systems." (Kafai, Proctor, and Lui 2020)

Situating our findings under CAL also provided tremendous inputs for our design of future digital experiences for children. At the same of time of exploring a constructive application of behaviourally or psychologically grounded approaches to enhance children's self-autonomy and resilience, we must recognise the importance of supporting children's CAL. Existing approaches taken by the CAL community have shown some fruitful results. A good critical computational thinking ability would pave the crucial foundation for children to exercise informed choices in a transparent algorithmic system and chances of exercising their data autonomy.

4.4.3 Implications for Designing Future Digital Experiences for Children

Our understanding of children's current experience of datafication prompted an urgent need for rethinking what future data-driven digital experiences should be like for children, so that we can reduce the negative effects they have on children (Lords House of Commons; Joint Committee on the Draft Online Safety Bill 2021; Kidron, Evans, and Afia 2018). Our findings have provided some immediate indications regarding how we could better support children in the datafied society and a need for a fundamental shift of the current data governance structure.

To start with, children's inability to recognise data transitions across platforms and their data ownership indicated how the lack transparency of the current datafication approach is damaging children's development of identity (J. Zhao, G. Wang, et al. 2019; Mascheroni 2020) and their ability to effectively link such practices with, let alone effectively recognise and comprehend, data processing and inferences. As a result, we had the majority of the children believing that datafication was largely for generating better services and that datafication was a very localised phenomenon. Transparency mechanisms have been extensively studied to be brought in as 'hints' for children to remind them of and to help them better posit their decisions during their interaction with the online platforms (Resnick, Berg, and Eisenberg 2000; Westlund and Breazeal 2016). However, previous attempts on transparency mechanisms tend to focus on raising children's awareness on certain practices on the online platforms (e.g. their data is being collected). These existing transparency mechanisms were mainly oriented around empowering children with more *cognitive understanding* of key computational concepts through child-friendly ways, such as Lego's Caption Safety (Lego 2021) and Google's Be Internet Legends (Google 2021). On the other hand, as our findings as well as the CAL framework indicated, we must extend these existing transparency mechanisms to support children developing *situated recognition* in broader and more diverse contexts. A recent design example from the UK ICO has shown how the transparency of implicit data privacy risks could be conveyed to children in a much more effective way if these risks were situated in a variety of scenarios from children's lives that they are more familiar with (ICO 2021). Future transparency mechanisms should focus on seeking to promote children's critical thinking, that is to increase their awareness of potential outcomes, ideologies and values associated with datafication practices, instead of the factual hints on the surface.

In fact, the lack of transparency mechanisms could result in children not being able to make informed decisions online. Self-determination theory explains that children's ability to self-regulate requires intrinsic motivation to enact specific behaviors and internalization of norms (Grolnick, Deci, and Ryan 1997). Understanding social ideologies and values related to their decisions enables children to exercise executive function, control impulses, and make more informed decisions online. This has been reflected during our studies, that children particularly struggled at some of the critical interaction points on YouTube includes recognising autocomplete when they are conducting searches, personalised advertisements embedded in the videos, autoplay for the next-up videos, or a list of recommended videos personalised for them. Children's discussions showed that they were not always capable of recognising the datafication behind such interaction points: a lot of them would randomly select terms from autocomplete list without fully recognising how each term is generated for them, permit autoplay without thinking too much about how and why autoplay videos were generated for them, or click on advertisements links without realising that those advertisements were specifically targeting at them.

On the other hand, such behaviours of children resonate the importance of designing for children's digital experience at these critical points (Lukoff, Lyngs, Zade, et al. 2021), which could go beyond transparency, and more generally about enabling children to think a bit more and make more informed decisions online. Several studies have explored how to scaffold children at various critical interaction points. A recent study on YouTube, for example, introduced redesigned mechanisms to offer users with different level of control at points such as search, autoplay, recommendations, thus resulted in users feeling greater sense of agency (Lukoff, Lyngs, Zade, et al. 2021). Related research with children have shown that changing the designs of how a video finishes would greatly alter how children decide what to do next (Hiniker, Heung, et al. 2018). Meanwhile, there have been a variety of research on how psychologically-based design mechanisms could be brought in at these critical interactions. Studies have attempted to adopt the use of dual system theory (Cushman 2013) to support users' ability to regulate their digital experience (Lyngs et al. 2019; Sharma et al. 2021), by carefully framing messages to encourage more automated and reflective exercise of self-autonomy.

Our findings underscore that children have a pronounced aspiration to exert greater control over the usage and processing of their data. While existing literature has scarcely addressed the support children need for better autonomy in datafication, our study illuminates the pressing need to steer future design endeavors in this direction. It also emphasizes the potential interplay between children's pre-existing algorithmic literacy and the cultivation of their agency, confirming that bolstering children's media and digital literacy is pivotal for enhancing their resilience against detrimental content and experiences online.

4.5 Conclusion

In this chapter, we contributed a first understanding of the perceptions of children on the datafication and the more specific data inference practices that dominates their information consumption online. Children are often regarded as not capable of coping with the complexities of online life. However, our research in this chapter showed that not only do children care significantly about various aspects of datafication, but they already possess rudimentary conceptual understandings of it, and a significant willingness to take action to shape it to their desires.

Furthermore, our research affirmed the value of the critical algorithmic literacy framework as a potent tool for deciphering children's comprehension and views on datafication. This framework can offer a robust scaffolding structure to bolster their literacy development.

With this foundation in place, the next chapter will shift focus from children's perceptions to exploring the requisite support and challenges they encounter when navigating datafication practices.

Statement of Authorship for joint/multi-authored papers for PGR thesis

To appear at the end of each thesis chapter submitted as an article/paper

The statement shall describe the candidate's and co-authors' independent research contributions in the thesis publications. For each publication there should exist a complete statement that is to be filled out and signed by the candidate and supervisor **(only required where there isn't already a statement of contribution within the paper itself).**

Title of Paper	`Don't make assumptions about me!': Understanding Children's Perception of Datafication Online.
Publication Status	Published
Publication Details	Ge Wang, Jun Zhao, Max Van Kleek, and Nigel Shadbolt. 2022. `Don't make assumptions about me!': Understanding Children's Perception of Datafication Online. Proc. ACM Hum.-Comput. Interact. 6, CSCW2, Article 419 (November 2022), 24 pages. https://doi.org/10.1145/3555144.

Student Confirmation

Student Name:	Ge Wang		
Contribution to the Paper	Ge Wang is the principal author who contributed the ideation, data collection, data analysis, and paper write-up. Jun Zhao, Max Van Kleek, and Nigel Shadbolt critically reviewed the manuscript and offered valuable feedback.		
Signature		Date	22/09/2023

Supervisor Confirmation

By signing the Statement of Authorship, you are certifying that the candidate made a substantial contribution to the publication, and that the description described above is accurate.

Supervisor name and title:	Dr Jun Zhao, Senior Research Fellow		
Supervisor comments			
Signature		Date	19 March 2024

This completed form should be included in the thesis, at the end of the relevant chapter.

Publications arising from this chapter:

- *Ge Wang, Jun Zhao, Max Van Kleek, and Nigel Shadbolt. 2023. 'Treat me as your friend, not a number in your database': Co-designing with Children to Cope with Datafication Online. In Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23). Association for Computing Machinery, New York, NY, USA, Article 95, 1–21.* `https://doi.org/10.1145/3544548.3580933`.

  Best Paper Award

5

Designing to Support Children in Coping with Datafication

Contents

In the previous chapter, we delved into the ways in which children understand and perceive the prevailing datafication practices on their favorite everyday platforms. We also uncovered a noteworthy willingness among children to actively influence and shape these datafication practices.

In this chapter, our exploration persists as we tackle the overarching research question: *In the current AI landscape, what insights can be gathered about children's understanding, perceptions, needs, and the obstacles they face regarding having* *autonomy over their data?* **(TQ1)**. To narrow our focus further, we will delve into the support children require and the challenges they encounter while navigating datafication practices.

All studies in this chapter received ethical approval in accordance with university guidelines for human participant research, under reference R76800/RE002.

5.1 Background and Motivation

Our previous chapter on children's perceptions of the datafication practices online showed that while there are still key knowledge gaps in children's understanding of datafication practices, most children already possess rudimentary conceptual understandings of some aspects of datafication, and have a significant interest and willingness to take action to shape it to their desires. In this chapter, we aim to extend these existing understandings, and examine how children would like to be supported in order to take action on the datafication practices. More specifically, we aim to explore the following research question: **What kind of designs are desired and needed by children to help them navigate datafication, and what barriers might they encounter?**

To this end, we continued to use the YouTube platform as an example datafication platform (Matamoros-Fernández 2017; Mitrou et al. 2014) for our study. Through 10 co-design sessions with 53 children, aged 7–14, from UK schools undertaken between May and June 2022, we identified different types of envisioned support and design mechanisms desired by children of diverse age groups. Our findings provide crucial insights for creating age-appropriate support for children's algorithmic literacy development, highlighted and unpacked the importance of no one-size-fitting-all designs when supporting children's coping with datafication.

5.2 Methods

Given our focus on investigating how children want to manage datafication practices online, we chose the YouTube platform to be used as an example, and conducted a series of co-design activities with children, including fictional inquiry and feature redesign, to elicit their requirements. Same as the previous chapter, we continued to select YouTube as the exemplar datafication platform to ensure consistency.

As part of our child-centred approach, we drew inspiration from Druin's influential model for integrating children into the design process (Druin 2002). In her model, children can adopt various roles such as users, testers, informants, and full-fledged design partners. While the first two roles emphasize gathering feedback from children towards the end of the design cycle, the latter roles focus on deeper collaboration, viewing children as equal stakeholders in designing new technologies and valuing their contributions throughout the process.

In this study, we treated children as *equal design partners* alongside adult researchers, embracing the co-design methodology. Originating from approaches like user-centered design (Eason 1995) and participatory design (Muller and Kuhn 1993), co-design has evolved to foster mutual creativity and understanding between designers and users. In the realm of Child-Computer Interaction, co-design underscores the significance of children as equal contributors in the design journey (Woodward, McFadden, Shiver, Ben-Hayon, et al. 2018; K. J. Lee et al. 2021; Druga, Yip, et al. 2021). Through this methodology, we highlighted the distinctive perspectives of children. This approach has been praised for its effectiveness in both capturing children's viewpoints on digital trends and in identifying their design requirements (Woodward, McFadden, Shiver, Ben-hayon, et al. 2018; McNally et al. 2018).

5.2.1 Study Design

To encourage children's involvement and their voices in the co-design process, we planned each co-design session to be composed of 3 activities: 1). Pre-design activity, 2). Co-design activity #1: Fictional Inquiry, and 3). Co-design activity #2: "Big Paper" Feature Redesign . The fictional inquiry session was designed to be more open-ended and to collect children's perceptions and how they envision to cope with the datafication practices; while the feature redesign session was more scaffolded by drawing on the CAL framing, in order to allow us to identify the actual support/design mechanisms needed by children to manage datafication. Each session was designed to last about 1.5 hour, consisting of 5–6 children and 2 adult researchers as co-design partners. The co-design groups were then broken into two design teams for Co-design activity #1 (*Fictional Inquiry*) and #2 (*"Big Paper" Feature Redesign*), with each design team containing 2–3 children and one adult design partner. In each design activity, the adult design researchers acted as partners by designing with children and facilitating discussions. After each design activity, the two design teams came back together for discussion. Throughout the study, adult partners co-designed with children and facilitated discussion in a way that avoids influencing the direction of the design and carried out conversations by encouraging children to clarify their design intents, rather that trying to guide the direction of discussion.

Pre-design Activity.

The warm-up session included a game of "throwing a ball" (Morgan et al. 2002) and invited everyone in the room to share their favourite YouTube video with others. This session was designed as a break-the-ice session to help the children to relax and get familiar with each other and the researchers. Then children were asked about two questions: *How do you think your video recommendations are generated? How do you think your personalised advertisements are generated?* We followed up children's responses by asking them to explain any terms they mentioned, such as "cookies", "trackers", or "profile". We have not intended to introduce this activity for learning purposes. The adult researchers did not try to provide children with a "right" answer, or provide guidance to elicit responses to those questions in any way. Instead, children were encouraged to express their own perceptions and opinions about datafication and related issues without being judged as right or wrong. These questions were designed to give us an initial insight about children's understanding of the datafication practices online. Particularly, we invited children to talk about their perceptions of datafication practices taking place, and their understanding of data inferences and profiling.

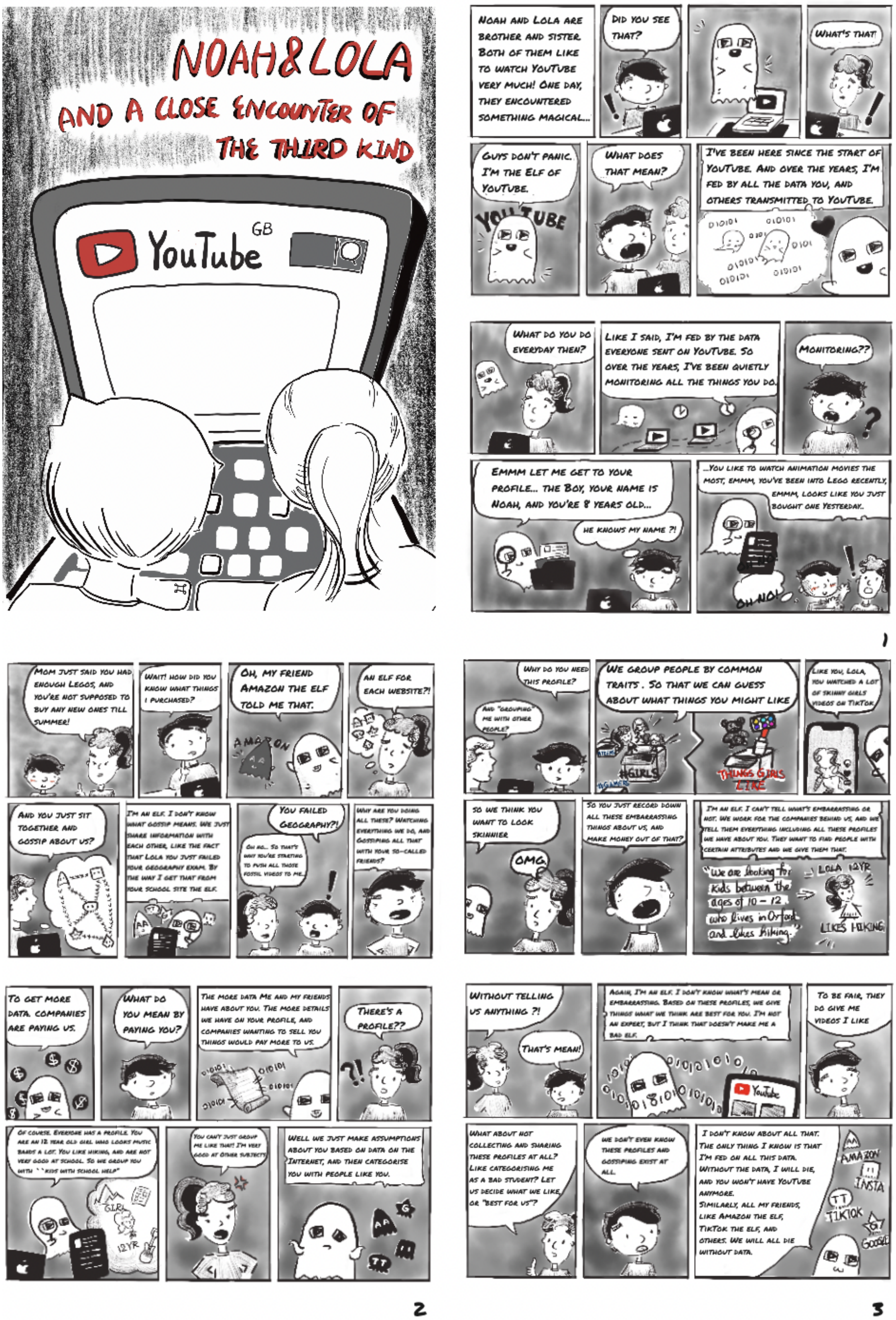


Figure 5.1: Fictional storyline used in co-design activity #1 for children to read through and complete as a group.

Co-design Activity #1: Fictional Inquiry.

In co-design activity 1, we conducted a fictional inquiry session. Fictional inquiry is a participatory design technique that entails creating an immersive fictional storyline and prompting participants to brainstorm within the context of this imagined reality (Hiniker, Sobel, and B. Lee 2017). By creating a fictional context for individuals to develop ideas, this method attempts to reduce the constraints of reality and free participants to be more generative. Prior work has showed this technique to be effective with children as young as 5 for eliciting their requirements (Hiniker, Sobel, and B. Lee 2017; Dindler et al. 2005).

In our work, fictional inquiry was used for facilitating children to think about their experiences with datafication practices, as well as how they may want these experiences to be different/better. For this, we created an original story titled "Noah and Lola: and a close encounter of the third kind" (Figure 5.1). The story describes Noah and Lola, who are brother and sister, meeting the YouTube elf (representing the YouTube platform) one day and having to decide what to do with the elf. In the story, the datafication practices were described to children through conversations between the three main characters, and the story was left open-ended. During the design, we made a careful choice of our character (use of "elf" instead of characters that were typically perceived more negatively such as "monsters" or "witches" (Hémar-Nicolas et al. 2021; Arrúa et al. 2017)) to avoid introducing any negative affections to the practice of datafication and tried to remain as neutral as possible by explaining both the good side and down side of datafication in the story.

At the start of this session, all children and adult design partners were invited to role-play when they read aloud the story as a group: each design partner took the part of a character from the story, and read out its lines. We noticed that this helped children to become more engaged and pay more attention to the story. After finishing reading the story, the adult design partners encouraged the group to discuss on "*What happened in the story?*" This helped the adult design partners to confirm what children understood about the story, and facilitate discussions to clarify the story if needed. The group was then split into two design teams (each with 2–3 children and one adult) to draw their suggestions about what Noah and Lola should do with the YouTube elf. After working on this task for 30 minutes, the design teams came together and each team presented their stories in front of the whole group, while an adult design partner took notes. The adult design partners then synthesized story ideas across the group and facilitated group discussions based on these ideas. During the design process, design partners were explicitly reminded not to focus on how well they can draw, but what goes into their story. We have not intended to design or introduce the fictional storyline for education or learning purposes due to the relatively short duration of the study; instead, this activity mainly aimed to invoke children's discussions about datafication and their envisioning on how to cope with these practices.

Co-design Activity #2: "Big Paper" Feature Redesign.

Previously in Chapter 4, we highlighted the effectiveness of the Critical Algorithmic Literacy (CAL) framework in fostering children's algorithmic literacy. While no single framework is universally accepted as the best for advancing children's digital literacy, and we don't posit CAL as the definitive approach for datafication knowledge, we've chosen to consistently employ the CAL framework in this study. It serves as a foundational bedrock for our design mockups. By integrating elements of CAL into our co-design activities, we aim to stimulate insightful discussions among children and pinpoint their nuanced design requirements.

According to the CAL perspective, computational thinking encompasses three pivotal facets: cognitive, situated, and critical thinking (Kafai, Proctor, and Lui 2020). *Cognitive thinking* focuses on the understanding of key computational concepts, practices, and perspectives and the associated skill building and competencies; *situated thinking* encourages learning to take place in contexts that the learner cares about so that they include their personal expression and social engagement in their pathway of learning; and finally *critical computational thinking* recognizes that computing is not an unequivocal social good, and emphasises the importance of supporting the questioning of larger structures and processes behind the computational phenomenon.

We introduced a series of design mockups, grounded in CAL elements, to our child participants (refer to Figure 5.2). The mockups offer variations on two key mechanisms on the YouTube platform—video recommendations on the homepage, and personalised advertisements that show up at the beginning of a video. These two mechanisms were chosen as they were considered to be the most representative for the datafication practices on YouTube. For each of the two mechanisms, we created 3 mockups for transparency, and 3 mockups for control, varied with the 3 design thinkings from CAL:

- The **cognitive-thinking** version of *transparency* mockups (the left card in Figure 5.2) provides the basic information, such as category of data being collected and used to generate video recommendations/personalised ads, without going into the details and the implications; and the *control* mockups were designed to offer a basic control on these different categories of data. (e.g., *We choose this ads for you based on: the time of day or your general location, your age and gender, your interaction with similar ads, and our estimation of your interests.*)

- The **situated-thinking** version of *transparency* (the middle card in Figure 5.2) and *control* not only display the data being collected and used for video recommendations/personalised ads, but also provide a contextualised explanation and control option according to children's personal experience on YouTube and multiple other online platforms. For example, in addition to the YouTube videos children watched, the design mockup also shows children the websites they visited and the products they purchased on third-party platforms. (e.g., *We choose this ads [Worms Rumble - Launch Trailer | PS5, PS4] for you because: you searched for "Worms Rumble" 8 times last week on Google, you purchased a PS4 console this week on Amazon...[other online activities].*)

- Finally, in the **critical-thinking** version of *transparency* (the right card in Figure 5.2) and *control*, we tried to reveal the bigger picture behind video recommendation and personalised ads, explaining the process of profiling and offering controls on this profiling process. (e.g., *We collect all your activities across all websites you visited into a profile, that means all your digital footprints on the Internet. Your profile is as follows: Love gaming, particularly into Worms Rumble (information from YouTube and Steams)... [other interests assumptions].*)

Again, in each session, 5-6 children and 2 adult researchers participated as co-design partners. The adult partner first presented the group with the CAL-inspired mockups on a big screen. Each child partner spent five minutes reviewing the CAL-inspired mockups, which were showcased on the screen in a random order. Each child partner was then invited to go through what they like and don't like about each of the mockups in front of the group. The co-design group was then broken into two design teams (each with 2-3 children and one adult), to think about how they would like to redesign the given mockups. Each team received a packet of printouts of the mockups, pens and pencils, scissors, markers, and tape. We used the "Big Paper" paper-prototyping technique (Guha, Druin, Chipman, et al. 2004), where design partners directly iterate upon previous designs by cutting out, drawing upon, and marking up printouts with their suggestions, additions, and changes (Guha, Druin, and Fails 2013). After working on this task for 40 minutes, the design teams came together and each team presented their ideas in front of the whole group while an adult design partner took notes. The adult design partners then synthesised design ideas across the group and clarified and elaborated these ideas through discussions with the children. In our study, the CAL-inspired mockups were used for facilitating children's discussion and brainstorming on their needs when dealing with datafication practices; and whether such needs would have any age-appropriate implications. We do not claim that CAL is the best approach for scaffolding children's knowledge about datafication; however, we

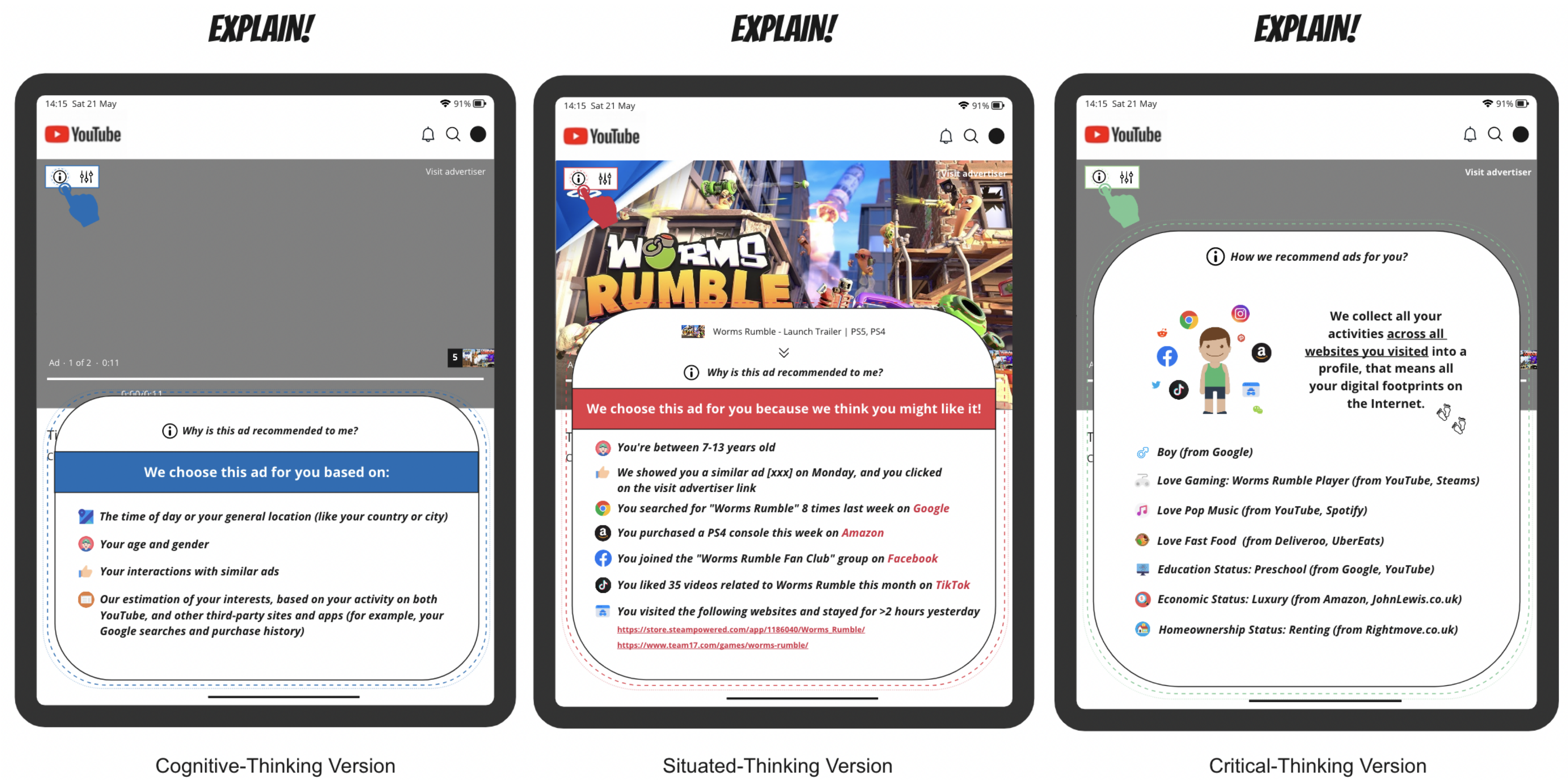


Figure 5.2: Transparency mockups redesigning the personalised ads mechanism on YouTube (EXPLAIN! cards). We created three versions of the mockups drawing on the CAL framing: cognitive, situated, critical. Each mockup was evaluated and later redesigned by the participants in the co-design activity.

introduced elements of it into our mockup designs as a way to invoke discussion and the more nuanced needs of children.

5.2.2 Participants

Participants were recruited from local schools, and a public forum for recruiting family participants. Recruitment started in May 2022 after obtaining institutional research ethics approval, and we conducted 10 co-design sessions (with an average group size of 5) with 53 children between May and June 2022, contributing towards a total of 19 co-design teams made up of children and adult design partners. In alignment with Chapter 4, we chose to continue working with children aged 7 to 14. Among the 53 participant children, 25 were between 7 to 11 [1], another 28 were among the age range of 12 to 14 [2], with an average age of 11 (range = 7–14, s.d. = 2.05).

Apart from the age of participant children, we also made a careful selection ensuring the diversity of the demographic background of our participants. Children were recruited from five local schools: two private schools [3], one grammar school [4], and two state schools [5]. For the participants recruited from public forums, we also noted down the type of schools they attend, participants' demographic information including their ethnicity, YouTube usage behaviours, and some basic information on their schools' and family education on topics related to datafication.

While participants were recruited from schools and public forums, the co-design sessions were not conducted in a school setting to mitigate the potential influence brought by the typical power structure in a school setting in which adults such as teachers exercise authority over children (Van Mechelen 2016). Instead, children signed up to our study were invited to our lab. This also allowed us to carefully organise children into groups, so as to ensure that single-session participants were of diverse ages, genders, ethnicity groups, types of schools attended, and related

[1] Primary school age in the UK.

[2] Secondary school age in the UK.

[3] Private schools (also known as 'independent schools') in the UK: charge fees to attend instead of being funded by the government.

[4] Grammar schools in the UK: government-funded schools that are allowed to select their pupils by means of an examination taken.

[5] State schools in the UK: government-funded schools that provide education free of charge.

experiences on datafication. Meanwhile, we tried to maintain an equal partnership between children and adults throughout all session. For example, sessions often begin with a snack and casual conversation; participants wear informal clothing; and children do not need to raise their hands to speak nor refer to adults by their titles or last name. Extra attention was paid to children who were younger, or knew less about the subject, or simply being shy.

5.2.3 Data Analysis

After completing the ten design sessions, we transcribed all the video recordings. There was a total of 927 minutes of video data (not including snack time at the beginning of each session), which resulted in a total of 1853 utterances used for analysis. Out of the 1853 utterances, 1396 were made by children (75.3%), the rest were made by adult design partners (researchers). We analysed the data using a grounded, thematic approach (Braun and V. Clarke 2006) to develop codes and themes related to each of the three parts of the study. For a comprehensive explanation of thematic analysis as a data analysis methodology, please refer to section 4.2.3. Photographs of children's drawing were also consulted to complement our analysis.

Results from part 1 of the study (Pre-design Activity) contained children's understanding of the datafication practices online. During the coding process, we tried to calibrate how children understand the YouTube's data processing practices, and this led towards a codebook on children's understanding of the datafication practices. Results from part 2 of the study (Co-design Activity #1: Fictional Inquiry) contained children's brainstorming on their designed story endings for the fictional storyline. Our analysis identified how children currently perceive datafication and their envisioning on how datafication should be dealt with. This led towards to a codebook on children's conceptual model on how the current datafication could be dealt with, what went wrong and what could be done better. With respect to the data collected from part 3 of the study (Co-design Activity #2: Feature Redesign). We first categorised children's rankings and comments around the three sets of CAL-inspired mockups (cognitive, situated and critical) for both transparency and control (EXPLAIN! cards and CONTROL! cards). For children's proposed redesigns of the mockups, we clustered their proposed design mechanisms. This gave us a codebook on the specific design mechanisms/support children want to have for knowing about and controlling the datafication practices, which is used to code the rest of the transcriptions.

5.3 Results

We present our results by first outlining children's overall understanding of the datafication practices online. We then present children's perceptions of datafication and their envisioning on how to cope with datafication. Finally, we present our analysis on children's preferences of the CAL-inspired features, followed by an in-depth analysis regarding children's desired design mechanisms for coping with datafication, with specific examples from our design sessions (e.g., [DSx]). We present individual children's quotes with their participant id and age.

5.3.1 Children's General Understanding on Datafication Online

In general, we noticed that children's understanding of datafication aligns with their age group, and can be broadly categorised into three phases (7-9Yr, 10-11Yr, 12-14Yr). This categorisation is also found to be aligned with UK ICO's Guidance on Age and developmental Stages (*Annex B: Age and developmental stages* 2020). In the UK, age 11 is also an important transition year for children to enter secondary schools. Here we report their understanding of datafication clustered into three major themes: *data collection, data sharing*, and *data inference* (see Figure 5.3).

Starting with children between 7-9Yr, children in this age group largely demonstrated an awareness of the *data collection* practices on YouTube, especially how their recommended content could be generated based on their activities on YouTube (e.g., videos watched, terms searched, location information). Meanwhile, most of them only considered location information as personal information and think the platforms own this data; instead, they do not think their online activities data such

Age Group	Datafication as *Data Collection*	Datafication as *Data Sharing*	Datafication as *Data Inference*
7 to 9Yr	Data on my interactions with YouTube (videos watched, terms searched, likes and comments) is collected to generate videos for me **These data are not owned by me, and they are not my personal information**	Collected data will be stored by YouTube, may be shared across departments within YouTube **Data won't be shared (traded) across platforms**	Some basic conceptual knowledge on collected data could be used to *do things,* but can't articulate more **Not aware that data could be used to infer personal aspects of them**
10 to 11Yr	Not only what I do on YouTube, but things I do on other platforms would be collected to generate recommendations **Not sure data on which platforms would be collected (tend to think within Google)**	Collected data will be shared across platforms **Data will only be shared across platforms that are related (e.g., under the same company, visited through the same browser)**	Data is processed through algorithm/some kind of automatic process, through which guesses about me (e.g., what I might like) could be inferred **Find it hard to understand how "guesses about me" could mean for other aspects in their life apart from "giving me better videos"**
12 to 14Yr	Data collection is enabled through *Cookies* There are terms and conditions on data collection in user agreement **Can't really describe how exactly is data collection enabled (apart from knowing the term *Cookies*)**	*Trackers* are used to track my activities across platforms **Can't really describe how exactly is data sharing enabled (apart from knowing the term *Trackers*), only vague ideas on data might be traded, but don't know how**	Data inference is making *profiles* about people **Can't describe how exactly is inference conducted** **Think they have rights, but don't know how to implement those in a datafied society**

Figure 5.3: Summary of children's understanding of datafication, organised into three themes.

as videos watched and terms searched are personal information, and many (7/11) regarded these data as "*owned by YouTube, not by me*" (P5, age 7). Meanwhile, children in this age group generally did not have much idea on the *data sharing* of their data, especially how it could be shared across platforms. None of them thought data would be shared, or traded by YouTube to another platform/company, as *"Why would they do that, I thought they [YouTube, TikTok] are enemies."* (P11, age 7). As a result, they tend to think data would only flow within *"different departments of YouTube"* (P12, age 8). In terms of *data inference*, most children in this group were not aware that once their data is collected, it could be used to infer more sensitive things about them. This point was also reflected in the later on co-design sessions, in which many of them demonstrated confusion on how profiles about users could be set up.

Children between 10-11Yr all demonstrated an awareness of the *data collection* practices on YouTube, and how their recommended content could be generated based on not only what they do on YouTube, but also their activities across multiple platforms *"If you visit a coffee website, YouTube will give you ads on coffee."* (P19, age 10). While some children (8/13) were able to discuss how data could be transmitted across platforms, they demonstrated different understandings in terms of how such across-platform *data sharing* is enabled. Many of them (10/13) thought data sharing was only performed between YouTube and websites that they visited through Google Chrome, as *"YouTube and Google is the same company."* (P40, age 11). As a result, they generally thought they won't be tracked as long as they don't sign in to a Google account, or just use other browsers such as Safari and Firefox. In terms of *data inference*, most children (9/13) in this group had some awareness or had heard about this topic, describing data inference as *"make guesses about me"* (P46, age 10). Some of them also mentioned the concept of algorithms, such as how their recommendations is *"automated by algorithms"* (P29, age 11) or *"some kind of machine"* (P24, age 10). Meanwhile, they were less certain about the specific things that could be inferred about them. For instance, while they described data inference as *"categorise people and influence the content they get"* (P17, age 11), they generally believed such inference was only for better videos per se, and found it hard to connect to other aspects in their life.

For children aged 12Yr onward, they tend to have more understanding on topics around datafication. Unlike the younger age groups who sometimes have some understanding on a certain topic but cannot describe fully, many children in this age group can accurately name specific terms in datafication, such as *data collection* is enabled through *cookies*, and cross-platform *data sharing* could be achieved through *trackers*. Meanwhile, some directly used the term *profiling* when referring to *data inference*. Children in this age group were more aware of the monetisation process behind scenes. For example, in conversations during design activities, many (19/28) mentioned that the companies are trying to make money – *"They collect our data because they want to sell it."* (P43, age 14); *"They all have some kind of partnership, for trading users' profiles."* (P45, age 14). On the other hand though, children were less certain when talked about how the datafication practices actually work *"I guess it's [enabled through] some kind of algorithm, but honestly don't know anything about it."* (P32, age 13). Furthermore, some of them (13/28) demonstrated confusion in terms of why platforms nowadays would try to infer/learn things about people, and were generally unsure about what rights do they (or can they) have in a datafied society.

5.3.2 Children's Envisioning on How to Cope with Datafication

In the part 2 of the study, children were guided to draw their own endings to the fictional storyline provided to them, making suggestions to the fictional siblings regarding what they should do with the YouTube elf. This activity encouraged children to think about how they would perceive and cope with datafication. Although the drawings from the children provided some valuable indications of their requirements, we found the discussions and think-aloud during the process presented additional and richer insights.

Demand for more respect from platforms

One of the most reflected points from our children participants was how they generally felt that they could be more respected by the platforms, through more interactive communications and dialogue. As such, children proposed three possible directions:

Increased Transparency. Many children expressed that more transparency about how platforms function and what they do with their data would be critical for them. Particularly, they often referred to the lack of transparency in the current terms and conditions provided by the websites, as a direct example of how their trusts of platforms are affected: *"Literally anything could be in the terms and conditions, if you're willing to sell your soul or something, but nobody would pay attention."* (P35, age 11). Some children also brought up the more design-level lack of transparency, such as how they identified the *"sneaky"* (P17, age 11) design patterns could hinder their autonomy online, and nudge them towards unwise decisions: *"If you click no (to the cookies) they say the app won't function, like a bunch of negative things."* (P16, age 10). Many children argued that they think it is important for platforms to make more child-specific considerations. Some children mentioned how they would want a child-friendly version of the terms and conditions, which can *"actually help me know what's going on"* (P41, age 12); while children suggested that they would like to be especially reminded of the *"important matters"*, such as *"profiling and selling of info - that's what most people actually care about"* (P39, age 9).

Increased Control. Children also brought up how they think platforms could be more respectful by providing users with more power of control: *"Yes being able to see things is nice, but what's important is that we would actually be able to do something about it."* (P27, age 14) Starting with data collection, children talked about how they would like to control the *"types of data being tracked online"* (P18, age 9) as well as the *"types of data being traded online"* (P20, age 11). Some children also mentioned how they want platforms to minimise the datafication practices: *"limit the data they harvest to only what's needed"* (P36, age 8), and *"delete our profile after an amount of time"* (P24, age 10). Furthermore, many children talked about how they want to be able to control their profiles. In particular, some children brought up the idea of having *"privacy-preserving profiling"* options, and they brainstormed on ways such as anonymising or randomising all profiles so that "platforms won't know as much about my life" (P28, age 11). Meanwhile, children also mentioned how they want to be supported when making decisions, and how they expect the platforms to guide them throughout this process: *"That's something the elf should be helping us with — making good decisions"* (P15, age 13).

Increased Sympathy. Related to develop communication and interactions with platforms, a strong theme is that children want to build a relationship with the platforms. Many children, especially the younger ones, described how they want to have *"actual bonding with the elves"* (P11, age 7), and building *"friendship with the platform"* (P23, age 9). In particular, they used words such as *"thinking for me"* (P22, age 14), *"be considerate"* (P46, age 10). A group of children [DS5] envisioned a scenario in which they can tell their secret (which is their data and profiles) to the elf, and the elf would help to protect that secret — *"the elf wouldn't tell anyone"* (P27, age 14). This idea echoed with the story from another group of children [DS3], in which they described a concept of a *"value-sensitive elf"* — instead of just being an algorithm which is programmed to do its job, the children want the elves to be actually taking caring of them and protecting them online: *"The elf can actually tell that, oh this information might be too sensitive or embarrassing for this child, so I won't hurt them."* (P7, age 12).

Demand for unbiased digital experiences

How datafication may influence their experience online was extensively discussed by the children during the fictional inquiry session, ranging from bias and discrimination, to targeted promotions, and to filter bubbles. To start with, some children discussed how they think the current datafication practices could create bias and discrimination due to them trying to *"group people online"* (P19, age 10), and how bias could arise due to *"gender steoreotypes"* (P33, age 14) or *"how rich they think you are"* (P17, age 11). In terms of how to cope with this, children argued that it is the platforms' responsibility to avoid bias and discrimination and they proposed a *"scrutinising algorithm"* that platforms could develop to assess whether bias and discrimination exist, and that algorithms used for profiling should not be based on sensitive categories: *"when they are profiling on people, they should know there are things people don't want you to profile."* (P36, age 8). Meanwhile, children argued that advertisers shouldn't be marketing on people based on sensitive categories from the beginning: *"It's 2022, and you're still targeting boys v.s. girls, you will get cancelled."* (P41, age 12); *"Ban advertisers from using some parameters for their ads, so now they can't ask YouTube to target certain groups."* (P18, age 9) Apart from bias and discrimination, some children also argued about how datafication could lead to echo chambers online: *"The profile would restrict the things you see, pushing you to whichever group they think you're in."* (P9, age 11); *"so people don't have a full view"* (P33, age 14) Children brainstormed on new mechanisms to increase their content diversity, a group of children [DS6] introduced an *"explorer mode"* in which users would be given more freedom to see what's outside their world — *"Like I'm a boy living in the UK and speaks English. I would want to see what a girl, say living in India would see."* (P29, age 11). Many children also mentioned how they want to more directly see whether they are in an echo chamber — how their content were limited by the profiling of them, for example, *"we recommend this to you because you are tagged with this"* (P12, age 8).

Demand for fundamental changes made to the datafied society

An interesting theme emerged from children's discussion is that, many of them believed that datafication is becoming a social issue, and they talked about how stakeholders - platforms, users, regulators should take an active role in the increasingly datafied society. Such belief has been even transformed into some kind of *data activism* — *"fundamental changes need to be made"* (P42, age 13).

Increased Public Awareness. Many children talked about how the datafication phenomena and associated consequences should be made aware by the public. They felt currently such practices were largely unknown by the general public, and they talked about how *"social movement"* (P5, age 7) and *"campaigns on social media"* (P4, age 8) should be brought in thus to *"spread the world"*. A group of children [DS2] even brainstormed on organising public protest on the datafication phenomena and its related issues: *"Noah and Lola would make a website, and ask people to join their protest."* (P6, age 12). Apart from relying on the public efforts, some children also talked about how they want new regulations to be made for protecting them against these datafication practices online, such as *"an upgraded version of GDPR"* (P42, age 13).

New Business Models for the Datafied Society. A large proportion of the children demonstrated a strong awareness that data is online platforms' main source of money. In fact, many of them have already accepted it as a norm that companies would be making money based on their data, in exchange of the services they offer: *"It's all about data selling, and it's very difficult to shut down something that is their core business."* (P8, age 14) On the other hand, some children envisioned there should be some kind of revolution on this data-centred business model. Children in our study generated several versions of *"new business models"*. For example, a 'Weight Loss Scheme" [DS6]: *"The Internet could go on a diet, like eating less but healthier - taking less but more effective data. And whoever signed up to it would be promoted more."* (P30, age 12) Another group of children [DS8] generated a similar idea to this — a "Fair Trade Union". Companies joining would be checked on the fair use of users' data, and thus get promoted more. A group of children [DS10] envisioned a *"Data Pass Scheme"*. The idea was that users would pay companies for them to stop taking their data (through purchasing "data blocks"). On the other hand, some children began to question why platforms have the rights to make money from their data in the first place: *"We should be the ones getting paid as it's our data."* (P41, age 12) and described such data-centred business model as *"the ultimate scam"* (P42, age 13). Finally, some children [DS5] talked about some initial ideas towards a data-decentralised structure - future platforms would only be passing requests from their local device: *"Every phone or laptop will have a creature living inside. But it only does things locally, like providing you service based on what's in your phone, but not giving your information away to YouTube or TikTok."* (P25, age 11) and platforms would *"only be parsing whatever is requested by these creatures."* (P24, age 10).

5.3.3 Children's Desired Designs Mechanism for Coping with Datafication

In the part 3 of the study, children were asked to comment and redesign the provided CAL-inspired mockups. We observed three themes about how children would like to be supported when coping with the datafication practices on YouTube. First, children demonstrated *age-related* design needs for them to make more informed choices. Second, children envisioned more *humane* designs that treat them in more respectful ways. Finally, children desired for more *autonomy-supportive* designs for more active engagement. Importantly, these proposed design mechanisms aligned closely with the themes emerged from their envisioning on how to cope with datafication during the fictional inquiry, demonstrating children's desire of transforming their conceptualisation to concrete design practices.

Age-Related Needs for More Informed Choices

Although the fictional inquiry sessions indicated some differences of responses from different age groups, our feature redesign activities led to more specific age-related observations. Children often showed different preferences and expressed different needs for how they want to be supported, often depending on how familiar they are with datafication concepts, which was typically related to their age.

Starting with children who had possibly less datafication knowledge (typically between 7–9Yr), almost all of them (10/11) preferred the *cognitive-thinking* designs the most, which present children with the basic information about what data was being collected and how. They found these designs *"easy to understand and configure"* (P36, age 8). Meanwhile, they largely disliked the *situated-thinking designs*, which listed their activities and showed how their recommendations were made based on these activities; and they reported feeling *"being judged"* (P49, age 9) and *"unsafe"* (P11, age 7). They also disliked the critical-thinking designs, considering the designs mentioning profiling as *"random"* (P4, age 8) and *"feels unreal"* (P5, age 7). Interestingly, even though these children were previously able to describe some basic datafication concepts such as how platforms take and make use of users' data; they seemed to have difficulty in relating such abstract concepts to their own datafication scenarios, thus showed understanding barriers and felt intimidated when actually being showed what datafication can learn about them. Children also exhibited desire for more straightforward and more direct support for coping with datafication. For example, P11 (age 7) suggested removing the "complex sentences", explaining: *"I feel like 'we build your profile' is a bit too deep"*, and just using words like *"We could guess what you like."* Children also wanted simpler control configurations and they especially liked the idea of having a *"one button for all"* (P12, age 8), proposing designs that can *"stop profiling in just one click."* (P5, age 7). Meanwhile, they preferred more direct and obvious support when making decisions. For example, they wanted for designs that *"just tell me what to do"*, giving them direct instructions on *"if clicking on this button is good or bad"* (P36, age 8), or direct parental help *"Mom will tell me what to do."* (P10, age 8).

We found that children (usually between 10-11Yr) in our study started to demonstrate a different set of preferences, going through a transition phase. They showed more positive perceptions towards the situated-thinking designs for them being *"more related to me"* (P46, age 10), while some disliked the cognitive-thinking designs, for them being *"too vague"* (P16, age 10) in the explanation, and being *"too broad and general"* (P20, age 11) for exercising controls. These children were able to connect abstract datafication concepts to *me*, and care more about how *me* would be affected by the datafication practices. Such a contrast to the younger participants was also reflected in their design proposals, which included significantly more designs on managing things about *me*. For example, P16 (age 10) envisioned for designs to tell them more specifically how profiles were formed around *me*: *"How they made that guess. Maybe like put up a search history, saying that I have searched this on this day so they think I'm a child or a boy or something."* Children also showed greater interest in controlling things about *me*, such as *"removing these guesses about me on my profile"* (P25, age 11) and *"choosing which websites can receive my profile"* (P52, age 10). In general, children want to be able to decide what goes onto their profile, how it's generated and could be used. In terms of how they want to be supported, it is interesting to see that, unlike younger children who preferred direct help, children from this age group wanted for support that help them make their own choices: *"Tell me the consequences of my choices, but let me decide."* (P35, age 11), and they expected parental involvement in more communicative ways instead of just telling them what to do: P17 (age 11) explaining *"We added a button here to invite parents to do these settings with us."*

Finally, from around 12Yr onward, children in our study largely found the cognitive-thinking insufficient and preferred the situated and critical-thinking support. Meanwhile, they demonstrated greater interest in the critical-thinking designs. Apart from *things about me*, children from this age group started to also become interested in *things about people other than me*, and more broadly around the datafication phenomena and its implications. This is also reflected in their designs, such that they started to want designs that explain *"the full picture of datafication"* (P38, age 14) to them, including topics ranging from algorithms used in the datafication practices — *"the kind of formula, weights of factors used in the algorithm"* (P34, age 13), to what are the *"data partnership between the websites I visit the most"* (P31, age 12). Children from this group not only care about *what* is being done by the datafication practices, but also *why* the datafication practices were performed and how it would have greater impact. They expect designs to be delivered in ways with both facts and reasoning behind the facts. When comparing to younger children, children in this age group demonstrated interest beyond *things about me*, and extend such interest to "people around me" and the greater society. For example, P45 (age 14) talked about wanting to learn *"how my data, my friends data, my parents data, and everyone's data is combined and merged by them [platforms] and how it [datafication] would have effect on every single one of us."*. P31 (age 12) also described: *"Why they [platforms] are doing all these, how it benefit them and how it may have impact on us as users, and maybe even how it would impact the society."*

Demand for More Autonomy-Supportive Designs

Children in our study envisioned several key designs to assist them to have greater autonomy, i.e., to take more *active roles* when coping with datafication practices, where the children felt like they should be the ones to initiate the action. To start with, some children proposed that it is important for them to receive alerts and notifications in a more visible way so that they could take an active action. For example, P11 (age 7) described the design of *"a huge question mark that you can click on, at every place they brought up this profiling thing. To me, that's more important than other things"*. Furthermore, many children expressed that it is crucial to have more simplified designs that encourage *active* actions, instead of having to navigate complicated user interfaces. As a step in this direction, P18 (age

9) sketched a home screen that present users with all the settings they can configure on one page, instead of *"hiding all settings under sixteen layers of pages"*. Some more active designs include the ones that enable children to inspect on things to ask for clarifications and modifications [DS6], as in P28's (age 11) sketch with their teammate: *"We designed for this e-highlighter function, so that you can highlight the bit where you find confusing or don't like. Like them guessing that you like fast food, the website would then go back and review that assumption they made about you."* Some still more active designs were when the (typically older) child expressed a desire to control and personalise how data profiling is computed about them: instead being treated as passive recipients that can only configure things once all the datafication is done, children proposed ways to actively engage in the whole datafication process. Such mechanisms included designs that support them to *"choose which pages I visited can be used to generate my profile"* (P38, age 14) [DS7]; *"deciding the models used to build my profile, like I can assign a value on how much this thing I did online matters, or if that's just a random thing I did."* (P51, age 10) [DS4]. Similarly, instead of just being told how their profiles might be shared with other platforms, a group of children [DS6] designed for mechanisms that they can create a list of platforms themselves, deciding on who can have their profile or not.

Call for More Humane Designs

Another theme that emerged from children's design activities was their desire to have positive experience and willingness to build a positive relationship with the platforms. Children expressed their expectations to be treated more equally and more humanely by the platforms: *"If the platforms were humane, which I think they are, which I hope they are. They would know they are dealing with actual people, we are not just statistics in their database or whatever."* (P52, age 10). To start with, children described how they want to have *"more positive experience"* (P12, age 8) on the digital platforms, such as through the use of smiley icons and more friendly tones [DS2]: *"We changed this sentence to 'Will you allow us to recommend you videos based on your profile?'. Because with that, it feels like a nice lady trying to* *talk to you, unlike a machine just trying to generate info about you."* (P6, age 12). Some children also talked about the use of *"more humane way"* (P43, age 14) of delivering the designs: *"Don't just show me the numbers, it feels really cold."* (P50, age 13). Children expressed great emotions and angers when they felt they are not *"being treated as a human"*: *" I'm being dehumanised. The way they deliver these things [e.g., data policy] ... give me information, but don't care about my reactions at all."* (P51, age 14); *"It feels like all that matters to them is 'me' as my data, like not 'me' as a human-being."* (P2, age 13). One interesting observation we had on children's redesigns is their tendency to personify a platform, almost in an unconscious way. A direct example of this is how often some kind of bot - personified version of the platform would show up in different children's designs (e.g., [DS3], [DS4], [DS9]). Children talked about how they want to be able to have *"actual conversations"* with it, because *"That's what human do, they talk with each other, not just showing each other with numbers and statistics."* (P12, age 8) Other children [DS4] also designed for mechanisms that can *"take care of me, know what I want"* (P18, age 9) — designs that can tell what's personal or secretive to a child and help children to hide these information from the datafication practices, or even mechanisms that can *"protect me when I need it"* - P39 (age 9) redesigned for a mechanisms in which platforms are now able to identify if one's profile contains "sensitive information" and whether such information would cause harmful effect on them. When performing these redesigns, themes including *"friendship"* and *"relationship"* were frequently brought up (e.g., [DS2], [DS5], [DS10]) : *"Treat me as your friend, not just a number in your database"* (P6, age 12).

5.4 Discussion

Through our co-design activities, we identified a strong need for providing age-appropriate support for children of different ages. Our results showed that the depth of children's understanding of datafication varied significantly between age groups. We found that children in the younger age group preferred simpler designs (cognitive-thinking inspired ones) that offered them more simplified information, helping them to grasp the basic ideas on datafication and its implications, whereas almost all children in the older age groups preferred designs that are more situated to their actual digital experiences and to provoke their critical-thinking, and showed great willingness towards having more information "related to me" as well as learning about the more in-depth problems behind the datafication phenomena. Our findings provide critical insights into design implications for future age-appropriate solutions that support children in coping with datafication. We therefore propose that *there is no one-fits-all design solution* when it comes to designing for children. How shall future designers address the various needs in children from different age groups, and how shall they unpack this no one-fits-all design then?

Rethink what 'transparency' means for children of different age groups. To start with, we have observed that, in contrast to how some of the child-specific technologies have been carefully considered according to children's age and developmental needs (Finegan and Austin 2002; *Age appropriate design code* 2020; Daugherty et al. 2014), today's digital platforms have given much less consideration regarding how children should be informed of the ubiquitous datafication behaviours online. This poses a strong need for future designers to *rethink what 'transparency' means for children of different age groups.*. Our co-design activities have particularly focused on exploring the type of data transparency that children would care for and be able to make sense of. The general assumption is that children have less awareness about the datafication practice and data-based exploitation in their digital worlds (J. Zhao, G. Wang, et al. 2019; Stoilova, Sonia Livingstone, and Nandagiri 2020). Our research has shed new lights on this presumption. Children may have less ability to develop the nuanced mental models exhibited in the previous research with adults, however, our observations of how different age groups perceived datafication differently provide important indicators for future design developments: keeping languages simple is rudimentary for supporting younger children (7-9Yr), who also are more likely to need more parental involvement and support; whereas connecting datafication with a child's individual interests or context may provide a more convincing perspectives for designing algorithmic transparency for older children (10Yr+). For 11Yr+ children, we recommend designers to consider scaffolding children for more in-depth thinking including what are their roles and rights in a datafied society. Future design investigations must be cautious of the age-specific needs from children and a child-centred design approach is crucial to the process.

Reposition children as active participants than passive consumers online. A key insight from our findings is children's strong desire on self-autonomy over their own digital experience and significant willingness to configure the datafication practices on them. This indicates a strong need to *reposition children as 'active' participants than passive consumers in the process of designing empowerment tools for them.* Designers should consider options to facilitate children to actively engage in the various datafication practices (e.g., how a specific video has been chosen for them, how data about them is shared and used); instead of treating them as passive recipients of content. This support should also be considered according to age-specific needs of children. For younger children (7-9Yr), designers should consider allowing children to conduct *direct* control, providing protection for children against the datafication practices not wanted (e.g., a button to turn off all data inference), and avoid hiding such option under layers of menus. We also suggest designers to offer children *direct* guidance on what would happen is a certain choice is made and send out alerts and notifications in a more visible way. For the slightly older children (10-11Yr), many children expressed their desire for real-time support as they make choices online, which could be better supported with mechanisms such as just-in-time visualisations of how choices could effect their online experience (e.g., a visualisation on the changes in recommendations). Older children (12-14Yr) demonstrated a strong need for more fine-grained control, as they seek ways to actively shape their own datafication experience. These children also demonstrated tendency to seek help from their peers instead of parents, and thus setting up mechanisms for peer support or for informing children of their data rights may be more meaningful.

Demonstrate care and respect. Another interesting theme observed from our findings is children's great willingness in building a *"positive relationship" or "friendship"* with the online platforms, and their desire in being treated in more humane ways - they want to be treated as humans, not *"a random number in the database"*. Such an expectation echoes with the recent line of work in UX literature around designing for dignity/humane by design, which promotes designing systems so that users experience dignity throughout the system, and its core holding is that designers must view users not just as a means to making capital, but treat them with respect and dignity themselves (*Humane by Design* 2022; *7 Heuristics For Humane Design* 2020). Children expected themselves to be respected and taken care of by the platforms. For the younger children (7-9Yr), they require "positive elements" to reinforce this reassurance, such as through smiley icons and nicer tones. An interesting finding is that the children consistently conceptualise ways for platforms to have conversations with them, and respond in a human-like manner (e.g., human characteristics). Prior studies have showed that personified voice assistants with traits such as different accents and personalities could be more favoured by young children (Garg and Sengupta 2020). Mechanisms such as using basic conversational elements to present information or respond to choices made (e.g., "You did a great job!") could be an effective direction to explore for younger children. However, such approaches must ensure children's innocence is protected but not exploited and guided by careful, ethical considerations. Older children (10Yr+) are keen to receive explanations that show them how things have been done in their best interests, and demand themselves not to be treated as numbers. Future designs should consider providing more than factual details, and complement such with contextualised explanations on how and why things are done. Designers should also consider allowing children to respond to information given and decision made for them (e.g., "I don't agree with this" button), and accommodate their requests in an equal and respectful way.

5.5 Conclusion

In this chapter, we contributed a first understanding of how children would like to be supported when coping with datafication practices and the challenges they faced. We found that children demonstrate age-related needs for more informed choices, call for more humane designs, and autonomy-supportive designs for more active engagement.

Specifically, we pinpointed key design opportunities that bolster children's autonomy development: introducing developmentally-appropriate transparency mechanisms, repositioning children as active participants rather than passive consumers, and emphasizing care and respect. These insights lay the groundwork for our subsequent exploration: designing and developing technological prototypes tailored to the multifaceted needs of children, fostering their data autonomy in AI-driven platforms.

Statement of Authorship for joint/multi-authored papers for PGR thesis

To appear at the end of each thesis chapter submitted as an article/paper

The statement shall describe the candidate's and co-authors' independent research contributions in the thesis publications. For each publication there should exist a complete statement that is to be filled out and signed by the candidate and supervisor **(only required where there isn't already a statement of contribution within the paper itself).**

Title of Paper	'Treat me as your friend, not a number in your database': Co-designing with Children to Cope with Datafication Online.
Publication Status	Published
Publication Details	Ge Wang, Jun Zhao, Max Van Kleek, and Nigel Shadbolt. 2023. 'Treat me as your friend, not a number in your database': Co-designing with Children to Cope with Datafication Online. In Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23). Association for Computing Machinery, New York, NY, USA, Article 95, 1–21. https://doi.org/10.1145/3544548.3580933

Student Confirmation

Student Name:	Ge Wang		
Contribution to the Paper	Ge Wang is the principal author who contributed the ideation, data collection, data analysis, and paper write-up. Jun Zhao, Max Van Kleek, and Nigel Shadbolt critically reviewed the manuscript and offered valuable feedback.		
Signature		Date	22/09/2023

Supervisor Confirmation

By signing the Statement of Authorship, you are certifying that the candidate made a substantial contribution to the publication, and that the description described above is accurate.

Supervisor name and title: Dr Jun Zhao, Senior Research Fellow			
Supervisor comments			
Signature		Date	19 March 2024

This completed form should be included in the thesis, at the end of the relevant chapter.

Publications arising from this chapter:

- *Ge Wang, Jun Zhao, Konrad Kollnig, Adrien Zier, Blanche Duron, Zhilin Zhang, Max Van Kleek, and Nigel Shadbolt. 2024. KOALA Hero Toolkit: A New Approach to Inform Families of Mobile Datafication Risks. In Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (CHI '24). Association for Computing Machinery, New York, NY, USA, Article 536, 1–29.* `https://doi.org/10.1145/3613904.3642283`

6

KOALA Hero Toolkit: A Hybrid Toolkit to Inform Families of Mobile Datafication Risks

Contents

Based on our theoretical underpinning of data autonomy in Chapter 3, and our empirical findings of children's understanding, perceptions, needs and obstacles when facing datafication practices in Chapter 4 and Chapter 5. In this chapter and next, we explore the thesis question: *How can we take a child-centred design approach to develop technical interventions for children; and how might these technical interventions influence, support, and enhance children's sense of autonomy over their data within AI-based platforms?* **(TQ2)**.

All studies in this chapter received ethical approval in accordance with university guidelines for human participant research, under reference R83638/RE001.

6.1 Background and Motivation

In this chapter, we look into how to design technical interventions that enhance the critical awareness of children regarding the risks of datafication around them. As previously discussed in Chapter 2, the majority of existing online safety tools for children adopt a parental-oriented approach, focusing on restrictions or monitoring. These tools often provide minimal guidance for children to understand the subtle risks associated with datafication. Such approaches not only diminish mutual trust between parents and children (Ghosh, Badillo-Urquiola, et al. 2018; G. Wang, J. Zhao, et al. 2021; P. Wisniewski et al. 2017), but also place a significant burden on parents to possess a deep understanding of relevant issues (Sonia Livingstone and Byrne 2018; Schiano and Burg 2017), given that most adults remain unaware of how their own data is collected, processed, and exploited to shape their digital experiences (Büchi, Fosch-Villaronga, Lutz, Tamò-Larrieux, and Velidi 2021).

To try to overcome these problems, we developed *The Koala Hero Toolkit*, a hybrid (digital and physical) toolkit that comprises a mobile tracker app, a set of data cards, and a tasksheet accompanied by worksheets. This toolkit is designed to help children and parents collaborate to better understand implicit datafication practices online, especially those associated with the use of mobile apps; while informing and fostering trust and communication between parents and children. More specifically, we aim to explore three research questions:

- **RQ1**: How do families with children perceive and navigate risks associated with datafication on mobile devices?
- **RQ2**: How, if at all, does the Koala Hero Toolkit change families' perceptions of such risks, and the thought processes undertaken in risk evaluation?
- **RQ3**: Does the Koala Hero Toolkit support families in making more informed decisions about datafication? What additional support might they require?

Through 17 user studies with 17 families (17 parents and 23 children aged 10 – 14), we found that the Koala Hero Toolkit strongly influenced both parents and children's perceptions of datafication risks, and prompted families to critically reflect and introspect about potential datafication risks associated with mobile apps. Families also demonstrated changes in how they made data-related decisions as supported by the toolkit, often demonstrating more democratic and interactive family-joint decision-making processes as a result.

6.2 Designing the Koala Hero Toolkit

As discussed before in Chapter 2.3, existing support for family privacy and datafication hardly considers fostering family communications and co-development, which misses the opportunities of facilitating children's risk coping development. Furthermore, research also shows that a critical understanding of the implications associated with datafication is crucial for users, particularly children, to take actions (G. Wang, J. Zhao, et al. 2022; G. Wang, J. Zhao, et al. 2023a). To address these vital needs, we believe it is essential to create practical support mechanisms to: 1) raise families' critical awareness about the implications of datafication by drawing on theoretical frameworks; and 2) promote family joint engagement by encouraging open family discussions on the subject and supporting collaborative family approaches to privacy and datafication issues. With these objectives, we designed and developed the KOALA Hero Toolkit. This toolkit encompasses three key components:

- A mobile tracker app, designed as a practical tool for families to navigate and control mobile datafication risks.
- A set of data cards, designed to facilitate discussion and support situated understanding.
- A task sheet accompanied by worksheets, designed to facilitate interactive family engagement activities.

In this section, we first introduce the design considerations behind the toolkit, followed by the design specifics of each component. Please note that we do not position this toolkit as a tool for educational purposes, as such a claim may necessitate structured assessments and clear benchmarks (Scalise, Douskey, and Stacy 2018; McGrath et al. 2015), which are beyond the scope of this study. Instead, its design aimed to raise and support the families' awareness of potential datafication risks, as well as stimulate and guide their thoughts and discussions concerning data-related issues around them.

6.2.1 Design Considerations

We drew inspiration from Kafai et al.'s computational thinking framework (Kafai, Proctor, and Lui 2020), which advocates a multiple-perspective approach for supporting children's development of computational thinking, and contains three key frames: i) *Cognitive thinking* focuses on the understanding of key computational concepts, practices, and perspectives and the associated skill building and competencies; ii) *Situated thinking* encourages learning to take place in contexts that the learner cares about so that they include their personal expression and social engagement in their pathway of learning; and finally iii) *critical thinking* emphasises the importance of supporting the questioning of larger structures and processes behind the computational phenomenon. While existing online safety and privacy measures often prioritise children's cognitive understanding, they often overlook contextualization in personal situations or the promotion of critical examination of observable phenomena (Dasgupta and Hill 2021). While we do not claim the approach from Kafai et al. to be definitive, we used it as inspiration to enhance our family support design.

Our design goal is to create a toolkit with an app that enables families to discuss and collaboratively understand the datafication risks on mobile devices. We began by outlining the app's key components, such as the ability to detect and disable trackers, or gain an overview of trackers associated with all the apps on the device. We ensured that these components comprehensively address all aspects of Kafai et al's theory, which are detailed in Figure 6.1.

Design Justification	DataTalk Tracker App						DataTalk Data Cards	DataTalk Tasksheet & Worksheets
	Intro Dashboard & Help Message	Apps Dashboard	Trackers Dashboard	Trackers View	Data Destinations View	Trackers Control		
Cognitive thinking	X			X	X		X	X
Situated thinking		X	X	X	X	X	X	X
Critical thinking						X		X

Figure 6.1: Mapping the Koala Hero Toolkit features to Kafai et al.'s computational thinking framework. The toolkit is aimed to be age appropriate and encourage open-ended play in its design choices.

6.2.2 Koala Hero Tracker App

As a result of these design considerations, the Koala Hero Toolkit contains the following key components, summarised in Figure 6.1. Figure 6.2 illustrates the six main screens of Koala Hero tracker app, which are described in detail below:

Intro Dashboard & Help Button, shown in Figure 6.2(a), supports children and parents with *cognitive* understanding of datafication risks. It provides a short story video that presents introductory information about data trackers, portrayed as different types of "elves" (e.g., functional/essential, social, or advertising), providing children and parents with basic information about trackers and their functions. We used animated and colourful elves to engage children. Users can access this introductory dashboard at the app's launch or by clicking the help button anytime.

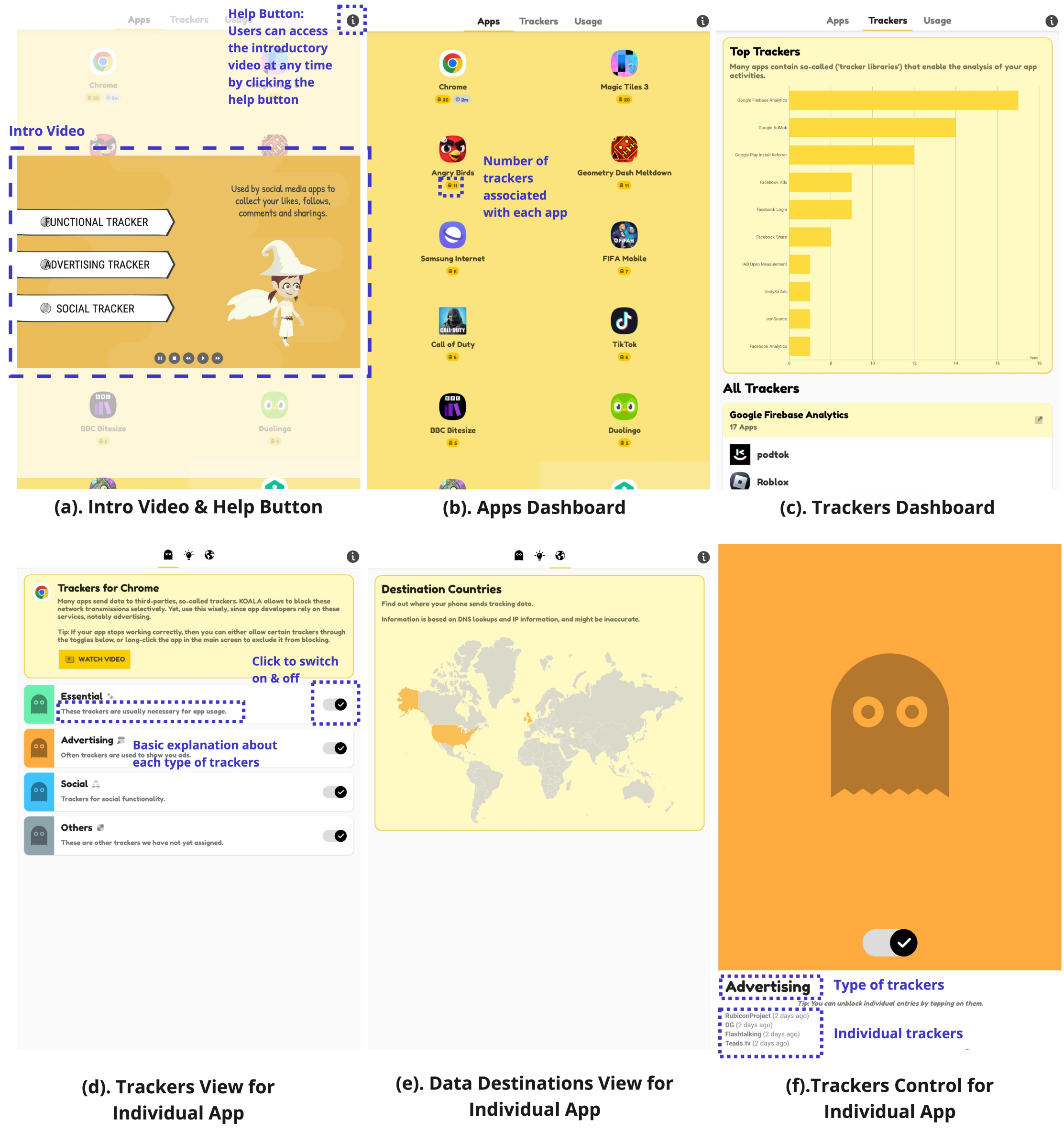


Figure 6.2: Koala Hero tracker app. (a). Intro Dashboard & Help Message (b). Apps Dashboard (c). Trackers Dashboard (d). Trackers View for Individual App (e). Data Destinations View for Individual App (f).Trackers Control for Individual App

Apps Dashboard, shown in Figure 6.2(b), provides a summary of all the apps on children's device and the number of trackers associated with each app. By designing the Koala Hero tracker app as an app that children can install on their own devices, children can see the trackers associated with their own apps. This encourages children's *situated* understanding of the datafication risks they are currently experiencing during their use of devices, creating a situation that is more likely to be relevant for the families, and allowing children to be more in control with their app choices.

Trackers Dashboard shown in Figure 6.2(c), provides a summary and ranking of all the trackers found on children's device. Similar to the apps dashboard, this feature also fosters *situated* thinking and encourages families to explore apps and trackers associated with higher risk factors. By configuring which trackers to block in Tracker View, children can also feel a sense of achievement by observing how the overall ranking of trackers may change.

Trackers View, shown in Figure 6.2(d), provides a summary of all the trackers of each individual app, grouped by types of trackers. Next to each tracker category, a help button provides users with a reminder of the basic explanation about each type of trackers, enhancing their *cognitive* understanding and allowing children to access help information whenever they need. As users click into each of the tracker categories, we provide a list of the exact trackers that have been collecting their data (Figure 6.2(f)), fostering a more *situated* reflection on the datafication risks in their daily lives.

Data Destinations View, shown in Figure 6.2(e), provides a map view of where in the world their data is being sent to for each app, supporting both the *cognitive* and *situated* thinking of children and parents around their datafication risks. Using a more intuitive map presentation, children can feel that they can explore and engage at their own pace.

Trackers Control for individual app is shown in Figure 6.2(f). Children and parents can use the "block" buttons to block data collection by certain companies. By enabling children to explore this function alongside parents and exercise agency over their own data choices, we aim to support families' *critical thinking* and enhance their sense of control.

Implementation. The development of the Koala Hero app represents a significant original contribution to the field of digital privacy and tracking analysis on the Android platform. This Android application is ingeniously constructed on the foundation provided by TrackerControl (TC) (Kollnig and Shadbolt 2022), which in turn is an extension of the well-regarded Android firewall, NetGuard (Bokhorst 2021). Our work with Koala Hero app enhances and extends these existing tools by incorporating both dynamic and static analysis techniques specifically tailored to scrutinize app tracking behaviors more comprehensively.

For dynamic analysis, Koala Hero leverages Android's VPN functionality to meticulously examine and, when necessary, block the network communications of other applications installed on the device. This process involves the real-time monitoring of data packets being sent and received by apps, allowing for the immediate identification of potentially malicious or privacy-invasive tracking activities. To facilitate this, network traffic is methodically cross-referenced with an extensive database of known tracker domains. This database includes, but is not limited to, lists maintained by Disconnect.me (Disconnect.me and Mozilla n.d.) and App X-Ray (Kollnig, Binns, et al. 2021; Binns et al. 2018), ensuring a comprehensive coverage of known tracking entities. Given that Android version 7 and higher do not allow the installation of custom certificates into the System Trust Store and that most apps secure their network traffic with HTTPS/TLS, it cannot gain access to the *contents* of other apps' traffic, only to the contacted *domains*, thereby protecting the integrity of the traffic. On the static analysis front, Koala Hero identifies embedded tracking libraries within the apps themselves. This is achieved by analyzing the apps' codebase against the Exodus Privacy tracker database (Exodus n.d.), a renowned repository of digital tracking signatures. This method allows for the detection of tracking components without the need to observe the app's network behavior, offering a valuable complementary perspective to the dynamic analysis.

6.2.3 Koala Hero Data Cards

To better facilitate hands-on exploration for families and foster deeper, more situated reflections, we incorporated physical elements into our toolkit design. The second component of our toolkit is a set of 18 data cards (see Figure 6.3), each illustrating the different types of data might be collected from children's apps, ranging from

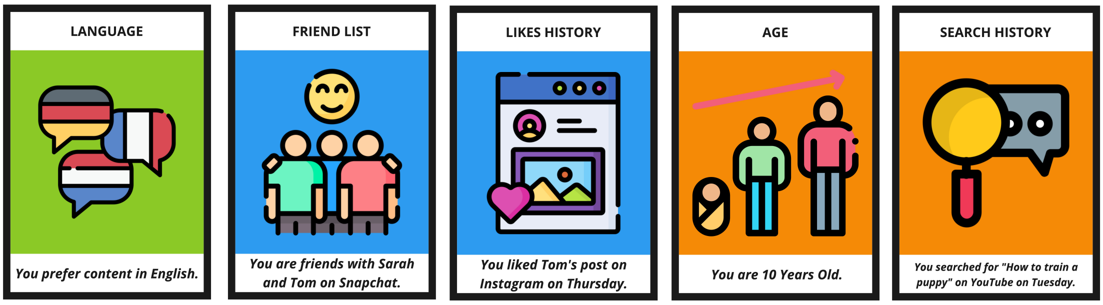


Figure 6.3: Examples of the Koala Hero Data Cards (18 in total).

account username, personal information such as age, gender and location, to online behavioural data such as browsing and search history. While these 18 data types don't cover all data online platforms collect, they represent the most significant categories from our survey of major platform privacy policies, like Google (*Google Privacy&Terms* 2023), Amazon (*Amazon Privacy Notice* 2023), and Meta (*Meta Privacy Policy* 2023). The data cards, designed to be physically printed and played by children and parents (used in tasksheet activities), were created to support the *cognitive* comprehension of data associated with their mobile apps and online activities. Furthermore, we curated the data card descriptions to help children and parents contextualise their own activities, such as "You are friends with Sarah and Tom on Snapchat", facilitating their *situated* understanding of how the data integrates into their daily lives.

6.2.4 Koala Hero Tasksheet & Worksheets

The third component of the hybrid Koala Hero Toolkit is a tasksheet and accompanying worksheets for families to complete together (see Figure 6.4). The tasksheet is designed to weave together and offer directions for families to effectively navigate through the Koala Hero tracker app and utilize the Koala Hero data cards. The tasks included selecting 3 most used apps by the children (one social media app, one gaming app, and one educational app) on their device, creating a context for children and parents to have more *situated* reflection on their own experience. For each of the selected app, the participants were then instructed to first go through

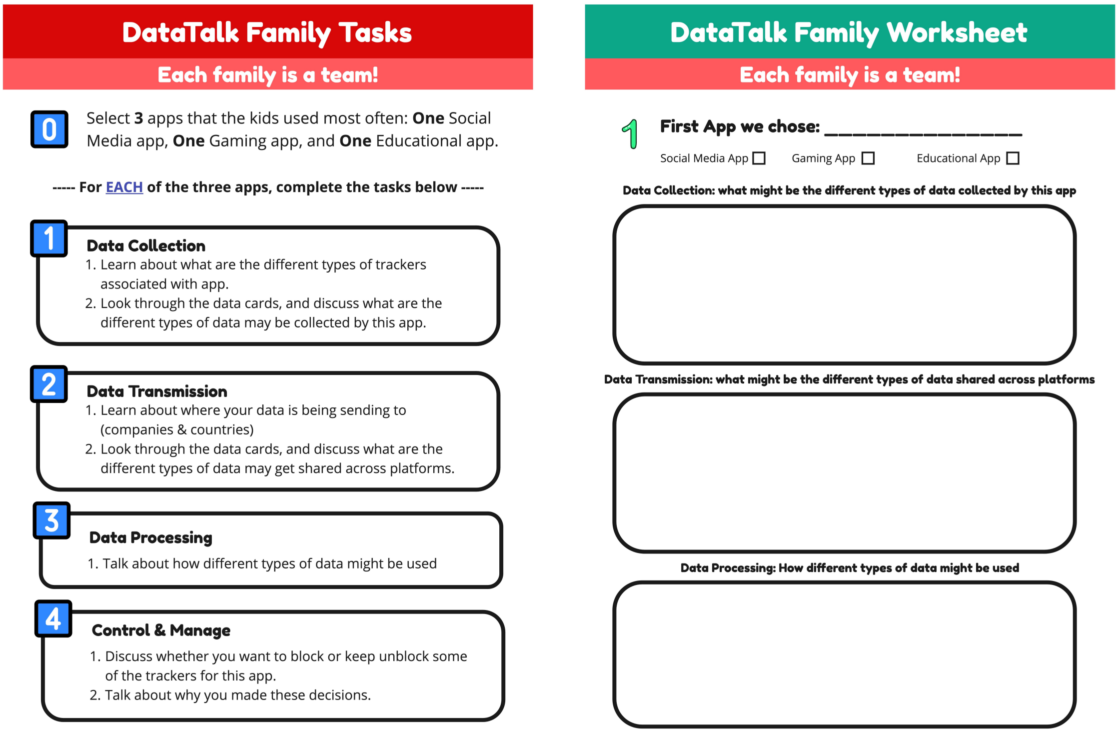


Figure 6.4: Tasksheet and worksheets for families to complete.

the app in Koala Hero tracker app to learn about the different trackers associated with this app, and where their data is being sent to, supporting their *cognitive* understanding of the datafication risks around them; they were then instructed to play around with the data cards, and to discuss topics around the collection, transmission and the processing of their data. Finally, participants were instructed to control and manage the trackers associated with this app through the Koala Hero tracker app and discuss their decisions. Through combining digital (Koala Hero tracker app) and physical (Koala Hero data cards) resources, this component fosters meaningful discussions between children and parents, encouraging them to *critically* reflect on the data collected and its potential implications for data transmission and processing. It also encourages exploration of the larger structures ad processes behind the mobile data, uncovering deeper datafication risk considerations.

6.3 Methods

6.3.1 Study Overview

In our study, we aim to address our three research questions: firstly, to explore how families currently perceive and navigate datafication risks; secondly, to assess how our toolkit might influence families' perceptions of datafication risks; and thirdly, to explore how our toolkit can help families become better informed about these risks, while also identifying any additional support needed. Participants were invited to our lab for the study, and were asked to bring the children's most used Android device. Our user study consisted of three parts: 1). an onboarding session, 2). activities with the KOALA Hero Toolkit session, and 3). an open-ended family reflection session. To complement our data, we also used pre-study surveys (for parents and children to complete separately) to establish a baseline of their individual perceptions and experiences on mobile privacy in the onboarding session; and post-study surveys (for families to complete together) to reflect on their joint experiences and thoughts in the reflection session. Each study was designed to last about 1.5 hours, with one parent and 1 – 3 children from the same household. Each study was facilitated by 1 – 2 researchers, who took observation notes and offer assistance without disrupting family activities, intervening only to provide technical support when needed, aiming to minimise their involvement in the family's activities unless assistance was requested.

Onboarding Session (20 minutes).

We began with a "whose fave is it" game (Dombro, Jablon, and Stetson 2011), where parents and children guessed each other's favorite mobile app. This aimed to ease participants and create a balanced power dynamic for the study. Parents and children were then asked to separately complete a pre-study survey, available in versions tailored for both groups (see supplementary materials). The surveys contained questions about basic mobile privacy knowledge, including data collection, sharing, processing, and their current understanding and practices on these topics.

The questions in both versions of the surveys were nearly identical, with the child-version employing more child-friendly language to enhance accessibility. We didn't use the surveys to measure their knowledge but to help us establish a baseline of their existing perceptions and experiences on mobile privacy. Participants were encouraged to think aloud and explain their choices as they filled out the surveys. The 1-min intro video (see Figure 6.2a) was then played to familiarise the participants with basic concepts around data trackers and datafication (e.g., What are trackers, their different types, and what can they do with your data?). After the video, families were encouraged to discuss what happened in the video. This helped us to confirm whether the participants understood the content in the video, and facilitate discussions to clarify the video if needed. The video wasn't meant for educational purposes but to familiarise participants with concepts for their upcoming activities.

Activities with the KOALA Hero Toolkit (40 minutes).

In this session, participants began to engage with the KOALA Hero toolkit (as introduced in Section 6.2). They were firstly asked to install the KOALA Hero tracker app on the children's most used Android device. The KOALA Hero data cards were also presented to them. The participants were then given 3 minutes to navigate the app and acquaint themselves with its functionality as well as read through each of the data cards. After this initial setup, each parent-child pair was asked to go through the tasksheet and complete the tasks on the accompanying worksheets. Participants were told there are no right or wrong answers, and the goal of these tasks is not to evaluate the usability of the KOALA Hero tracker app, but to facilitate a better use of the hybrid toolkit. Participants were provided with pens and pencils to jot down their observations and ideas on the worksheets. While writing was optional, they were encouraged to verbalise their thoughts aloud for the audio recording. Researchers primarily took an observational role and only intervened for technical support when needed.

Family Reflection Session (30 minutes).

This session is designed as a wrap-up session for parents and children to conclude their thoughts and observations from the day, and to bring up any topics that have not yet been addressed. Participants were encouraged to share any surprising or interesting discoveries they had, and reflect on how the toolkit had influenced their perceptions and thought processes concerning datafication risks. An exit survey was presented to the parent-child pairs, which was completed collectively as a family unit, fostering a shared dialogue on their experiences and conclusions of the study. The survey contains summary questions about their thoughts on the day's experiences and their sentiments towards datafication (see supplementary materials). The survey's purpose was to help participants articulate their thoughts rather than quantify responses. The participants were encouraged to think aloud as they navigate through the survey. As they did so, researchers paid close attention to noteworthy comments and followed up with further questions or discussion as needed. Finally, families were encouraged to maintain the app on their devices and to take the toolkit, including data cards and tasksheets/worksheets, back home for continued use.

6.3.2 Participants

Participants were sourced from local schools and a public family recruitment forum starting April 2023, after obtaining institutional research ethics approval. We conducted 17 user study sessions with 23 children and 17 parents between May and June 2023. Each session consisted of a single family, typically comprising one parent and one to three children. We recognise that "family" could be a broad term, encompassing relationships like grandparents, aunts, and cousins. Yet, this study focuses on what is typically the nucleus of a family: parents, children, and siblings. This aims for an understanding of primary dynamics before potentially expanding to extended family in future studies. Each participant received a £15 e-gift card as a thank you gift for their participation.

To participate in the study, each child was required to have access to an Android device, and have a parent who was at least 18 years old, to participate alongside them. We carefully selected the age range of child participants to be between 10 and 14, for several reasons: previous research has shown that from 10 onward, children gradually transitioned away from mainly parent-guided online activities (Ofcom 2021); evidence has also shown that children under 13 are active users on many social media platforms despite of the age restrictions claimed by these platforms in their terms and conditions (O'Keeffe et al. 2011; Richards, Caldwell, and Go 2015), exposing them to a wide range of risks online (Shklovski et al. 2014; Shipman and Marshall 2020; Acquisti, Brandimarte, and Loewenstein 2015; Zuboff 2019). Of the 23 children, 12 were aged 10–12 and 11 were 13–14, averaging 12 years (range 10–14, s.d. = 2.06). 14 identified as boys, 9 as girls. Of the 17 parents, most were 35–44 (9), followed by 45–54 (6), and 25–34 (2). 12 were moms and 5 were dads.

Beyond considering participants' ages, we ensured a diverse demographic background. Participants were sourced from three local schools: one private school (which charges fees to attend instead of being publicly funded), one grammar school (government-funded schools that select their pupils by means of academic performances), and one state school (government-funded schools that provide inclusive educational free of charge). For those recruited from public forums, we recorded their ethnicity, children's school type, and parents' education and employment status (see Table 6.1). Out of the 17 families, 7 identified themselves as Asian, 5 as White, 2 as Black, and 3 as mixed ethnicity families. 10 children were in private schools, 9 in state, and 4 in grammar. Parents held master's (9), PhD (3), bachelor's (3), or high school degrees (2). 12 parents worked full-time, while others were part-time-employed (2), self-employed (2), or full-time parents (1).

6.3.3 Data Collection

Data collection took place throughout May and June 2023, during which we conducted 17 user study sessions with 23 children and 17 parents. All sessions were audio-recorded with the participants' consent, which was obtained through signed physical consent forms. For the children's consent forms, we ensured each child had a clear understanding of the study, including the fact that their interactions would

Demographic Info	Children (n=23)	Parents (n=17)
Age (Children)	10–14 (Avg: 12, SD: 2.06)	-
Age (Parents)	-	25–34 (2), 35–44 (9), 45–54 (6)
Gender (Children)	Boys: 14, Girls: 9	-
Gender (Parents)	-	Moms: 12, Dads: 5
School Type	Private: 10, State: 9, Grammar: 4	-
Ethnicity	Asian: 7, White: 5, Black: 2, Mixed: 3	-
Education (Parents)	-	Master's: 9, PhD: 3, Bachelor's: 3, High School: 2
Employment (Parents)	-	Full-time: 12, Part-time: 2, Self-employed: 2, Full-time parent: 1

Table 6.1: Overview of Participant Demographics

be recorded and anonymised. The first and second authors transcribed the audio recordings, systematically removing all personally identifiable information pertaining to the participants or any individuals they mentioned, from session recordings, notes and transcripts. There was a total of 1586 minutes of audio data, which resulted in a total of 3506 utterances used for analysis. Out of the 3506 utterances, 3025 were made by the participants (86.3%), and the rest were made by researchers. The survey data has been employed as descriptive information, offering an overview on families' perceptions and practices on mobile datafication risks, and used to enrich our analysis; however, it has not been used to identify any direct correlations or measurable effects, nor employed as a source of direct quantitative measurements.

6.3.4 Data Analysis

We analysed the data using a grounded, thematic approach to develop codes and themes (Braun and V. Clarke 2006). For a comprehensive explanation of thematic analysis as a data analysis methodology, please refer to section 4.2.3. Photographs of families' activity sheets and their use of data cards were also consulted to complement our analysis. Our final codebook included themes related to families' existing perceptions of risks, families' thought process development, noticeable family changes of perceptions of datafication and decision-making, and their desire for additional support. More specifically, results from part 1 of the study (*Onboarding Session*) mostly contained families' existing perceptions and practices of navigating mobile datafication risks (RQ1). Results from part 2 of the study (*Activities with the KOALA Hero Toolkit Session*) mostly contained findings on how the toolkit influenced families' perceptions and thought processes on handling mobile datafication risks (RQ2), as well as our observations on their decision-making processes and the kind of support they required (RQ3). Results from part 3 of the study (*Family Reflection Session*) further enriched our understanding of the support families needed (RQ3). However, it's important to note that the findings for each of the three research questions are not strictly confined to any single part of the study. For example, families might discuss their existing perceptions and practices at any point. The study is best viewed as an integrated process where themes can emerge and be explored at any stage.

6.4 Results

We present our results by first outlining children and parents' existing overall perceptions and practices of datafication on mobile devices (RQ1). Next, we present families' thought processes regarding datafication risks and subsequent change of perceptions (RQ2). Finally, we explore the impact the Koala Hero toolkit had on the families' joint decision-making and family dynamics (RQ3). We provide quotes from individual children and parents, identified by their participant ID, along with the age of each child. A child participant is represented by (C#, age x) and a parent participant by P#. Child participants who are siblings from the same household are denoted as C#a, C#b, and so forth.

6.4.1 Families' Existing Perceptions and Practices on Datafication Risks

Here we present families' existing perceptions and practices on datafication risks prior to using the Koala Hero Toolkit (RQ1), as reflected in their pre-study surveys and as observed in their conversations at beginning of the studies.

Most families reported that children had their own devices, with only one using a parent's device and another sharing among siblings. Of the children, 9 out of 23 used their devices for 1–3 hours weekly, 5 used theirs for 4–6 or 6–8 hours, and 4 used theirs for over 9 hours weekly. Most children (17/23) responded *not very concerned* about apps and companies collecting their data. 13 children reported *neither agree nor disagree* and *disagree* when answering *I would like to discuss privacy issues with my parents.* Parents exhibited a higher awareness of online risks compared to their children, but their primary focus was on online safety rather than data and privacy. For instance, 11 out of 17 parents used parental controls, mainly to limit access to age-specific apps. When selecting apps for their children, 12 parents prioritized its functionality and educational value, while only one considered privacy aspects. However, a majority (14/17) expressed concern over data collection practices by apps and their parent companies. While 15 parents wished to discuss this with their children, only three felt confident doing so. Even though nearly all parents had prior online safety talks with their children, the discussions rarely touched on data and privacy. This aligns with survey data, where 16 parents admitted they had rarely or never addressed these issues at home.

Furthermore, as illustrated in Figure 6.5, families demonstrated a diverse range of understanding when it comes to the more specific data concerns, such as the collection, transmission, and processing of their data by their mobile apps:

- In terms of data collection, almost all children and parents mentioned how they "kind of knew" this before. 17/23 children and 14/17 parents selected *agree* and *strongly agree* to *I know that some of my data, such as my usage of the apps, and some of my personal data (such as name and age) will be collected by mobile apps.* 11/23 children also reported having heard about terms such as "trackers/cookies", and some parents (12/17) mentioned them knowing apps can get data from their phones but do not know how.

- In terms of data transmission, fewer children and parents knew data could be shared between platforms. Only 8/23 children and 9/17 parents selected knowing *I know that my data may be used by other companies, who may have an agreement with the app.* This was also reflected later on in their conversations, such that a fair number of children (11/23) thought *"My data will only stay within the app."* (C10, age 10).

- As for data processing, almost all families had the initial idea as *"Data processing is solely for offering us better services"* (C7, age 12). Most parents (12/17) and children (19/23) selected *disagree* and *strongly disagree* to question *I know how data can be used to learn about personal aspects about me (e.g., whether I'm a boy or girl, the type of school I go to)* and *I know how data can be used to make inference on personal aspects about my family (e.g., relationship status, parent or not, favourite family holiday destinations).*

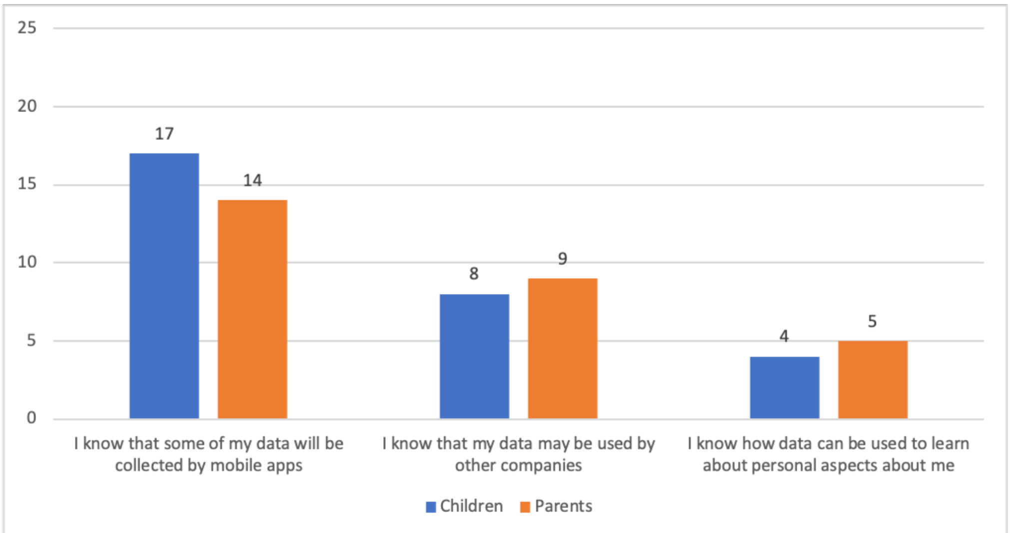


Figure 6.5: An illustration of existing perceptions about datafication risks among 23 children and 17 parents, with numbers above each bar indicating instances where children or parents share that perception.

6.4.2 Families' Thought Process: From Cognitive Understanding, Situated Reflection to Critical Thinking

Throughout our study, families demonstrated a variety of ways to utilise the Koala Hero Toolkit, exhibiting many playful ways to explore and make sense of datafication assisted by our toolkit.

Cognitive discussions

Families mostly started by trying to establish a cognitive understanding about trackers and associated datafication risks. For example, they often took some time to review the tracking information linked to each app via the Koala Hero tracker app, and they used this data as a springboard for their discussions. Families also tried to make sense of what possible data could be collected through navigating the Koala Hero data cards, *"Okay, what might be the things collected... Device info, 'You use an IPad in iPadOS16'. Oh that make sense, of course they gonna collect that."* (C6, age 10). Some families spontaneously used the Koala Hero data cards for some kind of role playing. Parents acted as the app, telling the children, *"I'm going to take this and this from you."* (P9) Following this, the entire family would begin deeper discussions on what this meant in the context of their daily experiences.

Various way to situate the discussions

While interacting with the Koala Hero Toolkit, families frequently engaged in *contextual reflections*. Most families related the study's topics to their personal experiences to understand them better. Siblings often recalled shared experiences, with one commenting: *"Remember when we liked socks on Instagram and then saw them on Amazon? That's them sharing our info."* (C15b, age 12). Additionally, children related the toolkit's information to what they already knew, noting, *"Friend list? We discussed this in school. They always say, 'Don't expose yourself on Facebook."'* (C7, age 12). Apart from families reflecting on past experience of their own, they also tried to make sense of the topics through linking with everyday scenarios, *"You wouldn't share everything about yourself with a stranger* *at a party, but yet we're giving out all our stuff out online."* (C9, age 13). We also noticed a trend among parents using scary examples to alert their children, *"If someone really wants to know it's very easy. [Children's name], where does he go to school? When is he online? If some bad person use this information to do some bad things about you!"* (P1)

Critical thinking from both parents and children

Along with their cognitive understanding and situated reflections, families also showcased various ways of critical thinking. For instance, families would combine the Koala Hero data cards to self-reflect the impact of losing control of various data points: *"Let's look at this card, location; and then this card here, '10 Years Old.' What they might find out about you? Maybe what school you go to?"* (P3). Some families related and applied the data cards to real-world issues they may have encountered, *"To me, each data card is like a box they've created for people; they make assumptions and put people into distinct boxes."* (C12c, age 13), *"Looking at these data cards, I see supermarket products. They're like items on shelves that big companies can freely pick and buy about people."* (C14a, age 13). We observed numerous instances where children critically reflected these observations: *"Wait, if they merge your friend list, language, and browsing history, they could probably deduce your ethnicity."* (C12a, age 11).

Apart from critically reflecting on the datafication risks and relevant concerns around them, families sometimes went on to reflect on what this meant for the whole society and the future of technology. In particular, many parents brought up the concept of "tradeoff", such that they believed *"Nothing is truly free. If a game is offered at no cost, they're likely selling your information."* (P9), *"It's the future of technology, everything is digitalised and monetised, and we rely on these services. Of course there are both good side and bad side about it, but it's crucial to be aware of them."* (P15)

6.4.3 Families' Change in Perceptions on Datafication

The various thought processes of families also affected their perceptions about mobile datafication risks (RQ2).

Families becoming more surprised/concerned/confused

To start with, families demonstrated great surprise as they explored their chosen apps using the Koala Hero tracker app, especially when viewing the associated trackers and destination countries of their data, *" Oh wow, I thought it's illegal for them to share the data, especially to a different country."* (P12). Meanwhile, almost all families were shocked at the amount of trackers associated with their apps:

> Let's look at Magic Tiles. On my God! 21?! (C12a, age 11)
> I did kind of know they are collecting from us. But I just didn't realise it's this many! I thought like, maybe everyday I run into 2 trackers, but it's like 20 just in this app. (C12c, age 13)
> At night you lock your door and keep yourself secure, but you never think about this opening you up. (P12)

They also demonstrated great confusion and concern around the mismatch between their perceived purpose of the apps and the types of the associated trackers: *"Tracker 'Amazon.com'. Amazon is very commercial, isn't it? So why would an educational application have an essential need to send your data to someone who wants to sell you stuff? That's like going to a school, and before you go into the front gate of the school, you are given a brochure of stuff to buy."* (P15)

Families becoming more aware of data inferences

As families continue on the activities supported by the Koala Hero Toolkit, we noticed that families started to become more aware of the power of data inference, and that more personal things about them can be learnt: *"Wait, now they're using Google tracker, and what you do on Duolingo will get sent to Google. But then I also use Google Chrome, so that means they also know what I search for and all that stuff. They might even able to pinpoint where I go to school, and like what they teach in school. This is overwhelming."* (C3, age 10).

Meanwhile, some families started to realise that data can be used not only for deducing the personal details of individuals, but also for collective profiling, *"I* *think, because they have so much data from such a big range of people, it makes the data more useful. The things would have more impact on a particular group of people."* (C13b, age 13).

Families also demonstrated a deeper level of concern about how their data could be used to push and nudge them into certain things and opinions: *"If they wanted to influence you, they could easily do it. They can get you hooked on something or got you looking at something or someone favourably, because they know so much about you."* (P9).

The role of stakeholders in datafication becoming more prominent

Both parents and children demonstrated increased concerns around what companies and platforms could do with their data: *"It's an uneasy feeling. They know all these thing about you and it's completely up to them now."* (C11, age 14). Almost all families expressed surprise and concerns on the dominance of big companies such as Google, Amazon and Meta. The families became aware of this issue as they look through the trackers of their apps, and they soon found out that the major companies were present in every single one: *"Oh Google again. I swear this is the tenth time I've seen their trackers today."* (C8a, age 10), *"If we look at the trackers, literally everything is sent to either Google or Amazon."* (P11). Families talked about their concerns on how powerful the data will be in the "hands of giants":

> Google knows everything, and Facebook also knows everything. It's crazy! (C14a, age 10)
> Our data is now in the hands of giants. Is there anybody monitoring what they do? (C14b, age 14)
> Yes it does reminds me that in essence, Google, Amazon, Meta, they are data companies. They make money off people's data and make huge profits. (P14)

Meanwhile, some families talked about the "unbalanced power strucutre" between big companies of them: *"In the future, the internet will record all our actions, and big companies might know us better than we know ourselves. As more people use it, their analysis will sharpen. There's going to be some point where you will just need to listen to the computer to decide your next step."* (C1, age 13). Families emphasised how they wanted to have a "stronger mind" when dealing with the datafication risks: *"I don't want to be distracted and manipulated by their little games, it's my data!"* (C4, age 11).

The impact of datafication on children becoming more tangible

Families started to contemplate the actual harm such practices could pose to children. Parents in particular, expressed concerns on data being taken from children from such a young age, and what that would mean for their future: *"They'll track you until you grow up. How do we know if they'll use that data against your future?"* (P8). Some parents also expressed concerns on children could become normalised and indifferent to these risks over the years as they grow up: *"What you just said sounds like normalisation to me. 'Data tracking, it's gonna happen. I'm normalised to it'."* (P10).

We also noticed a consensus among all families around the importance for children to be aware and kept informed of such datafication risks from a young age: "*We want to protect them, not bubble-wrap them. Awareness education, like this toolkit, is essential. They need to understand the risks and that data collection is often for others' benefit. Especially at this age, as they start exploring social media and online platforms, they need this awareness, and it takes time."* (P5). *"It's now in the back of our mind, and it's funny how before we were like, not concerned at all, and now we become concern of these things cause you realise things."* (C3, age 10).

6.4.4 Families' Change in Decision-making and Desire for Additional Support

In this section, we outline the ways in which the Koala Hero Toolkit enabled families to conduct joint decision-making, as well as the additional support they desired (RQ3).

Families felt more equipped.

In the meantime, families expressed how they now feel better equipped to make decisions when they return home and moving forward into the future. Both parents and children reported them liking how the Koala Hero tracker app **allowed them to see things in front of themselves**, *"In general, I think it's just good to see things in front of your face. I was suspicious of all this data tracking and stuff before, but now I can actually see that 21 trackers in front me, and I can better decide next time."* (C8b, age 13).

Parents also reported on how the toolkit **provided a structure** for them to start talking about these issues with their children, and engaging the children more effectively, *"It really helps me to have a structure to talk to my son. It's a great opportunity to get them interested, so they want to learn more about it when they see these practices again."* (P3).

Families further talked about they felt **better equipped to contemplate today's discussion in future scenarios**, particularly if they encounter similar situations and need to make decisions, *"Next time you get a recommendation. Think about they might think of you, and what they might know about you, not just, 'oh, okay, I'm just gonna look at it."'* (P9)

We observed instances in which families made a series of **weighted decisions** based on information exchange and family discussions enabled by the toolkit. Some families chose to only allow the essential trackers and block all others for all their apps. Others weighed the significance of a specific app to them and its function (e.g., educational apps), and chose to only allow all trackers for this, but not for others. A few families made more nuanced decisions by considering the companies behind. They opted to block all trackers from "dominant" companies such as Google and Amazon, while thoughtfully assessing the purpose of the tracker to ensure it did not interfere with their apps' essential functions.

Families having more equal family dynamics for joint decision-making

Through interactions with Koala Hero, families exhibited more balanced dynamics during activities. The toolkit design, especially the tasksheet's emphasis on children's favorite apps, often made **children took the lead in discussions** since they were the expert of their apps. Even in instances where parents initially assumed the lead role, they often found themselves asking questions like, "Is it true that this app does this?", and "What do you think?", indicating a shift towards a more equal family dynamics. At the same time, we observed a recurring pattern across all our studies where families began to **explore unfamiliar subjects together**, sharing their insights and helping each other throughout the process:

> Okay, I think this is a little bit difficult even for me. If we see these examples here [Koala Hero data cards], hmmm, they might look into your browsing history? (P8)
> Oh but you don't really browse anything with Angry Bird. (C8a, age 10)
> I think it collects game play content though. (C8b, age 13)

We noticed that families continuously engage in a process of discussion, negotiation, and conflict resolution as they **carry out joint decision-making**. Different family members might have varied viewpoints on aspects like which trackers to block or which apps needs inspection. As they exchange thoughts, new ideas and perspectives emerge:

> I would absolutely turn off all the trackers. (C14a, age 13)
> Why? They send you stuff you like and it's up to you. (P14)
> It's concerning that everyone thinks they're unaffected when they are. I've long felt that this data collection isn't right, but others around me just don't feel that. It should be a scary thing. (C14a, age 13)
> What mom said sounds like normalisation to me. (C14c, age 14)

Additional supports desired by the families.

Families, particularly parents, expressed a need for more data-centric controls. These included being able to govern the specific data being collected, managing data for real-time activities, and gaining a better understanding of precisely how the data will be utilised by the companies. Families also talked about having more contextualised information on "What can go wrong". Many families in our study expressed a desire for real-world examples and news to educate their children and other siblings on data misuse. *"Have examples here, like news on how our data are not being used legally. What can go wrong. As a way to educate."* (P12), *"It's frustrating being the only concerned one while others are indifferent. They should see the actual negative impacts on lives."* (C14a, age 13).

Besides the support from Koala Hero Toolkit, families sought external assistance, notably from schools and potential legislation. Both parents and children reported how they want the schools to take up more responsibilities, such as checking on the privacy practices of the educational systems they are using; and to extend the current education beyond solely online safety such as stranger danger, embedding data awareness education, *"The school needs to become more aware and hold hands of parents. We can then talk to kids at home about it."* (P10). Meanwhile, families also expressed the need for future legislation that actually works, *"Honestly, I don't think GDPR is effectively protecting our data, especially with all these trackers and I'm under 13! The internet is constantly evolving, so maybe the laws should too."* (C7, age 12).

6.5 Discussion

6.5.1 Impact on Families' Thought Process

Throughout our observations we noticed a consistent progression in family participants' thought processes. It's noteworthy that this pattern remained consistent across all the families, despite the varying ages (10–14) of the child participants. We also observed that regardless of families' initial understanding about datafication risks, this progression mirrored Kafai et al.'s three forms of computational thinking (Section 6.2): from *cognitive thinking* (i.e., understanding of basic computational concepts) – families figuring out basic concepts such as what are trackers, what data are being collected; to *situated thinking* (i.e., situate the abstract computational concepts in context children know and care about) – children connecting with their real life experiences, parents reflecting on their data practices for the children; and finally to *critical thinking* (i.e., supporting the questioning of larger structures and processes behind the computational phenomenon) – families questioning the dominance of tech giants, reflecting on how users are inferred and monetised in the datafied future, and making informed decisions based on their consolidating of information. While this is not a linear progression, we noticed that the cognitive and situated abilities are critical for enabling users' critical thinking.

How was such progression achieved then? While we do not claim that the KOALA Hero toolkit are the sole factors driving this change, we notice a strong relationship between some specific design features we included and the *triggering moments* experienced by families, which appeared to have fueled the progression in their thought process. We identified four critical *triggering moments* during families' interactions with the toolkit. The first *triggering moment* is when parents and children **see things right before their eyes on the tracker app**, supported by the information disclosure features on the KOALA Hero tracker app (e.g., trackers view and data destination view). This was the moment when almost all participants expressed surprise for the first time, expressing this either contradicted their previous beliefs or confirmed unverified suspicions. The second is when families started to **read through examples on the data cards**. The sentences are formulated in such a manner (e.g., "You are xxx / You did xxx last Friday") that all families started to connect with their own experience and reflect on what they did before spontaneously. The third *moment* is when families **documented their responses on the worksheet**. This formal requirement to write down answers encourages families to integrate all their observations and speculations derived from a range of activities, such as experimenting with data cards, identifying top trackers in the tracker dashboard, observing the total number of trackers on their phone, and speculating potential data collection based on their activities. And finally, the fourth *moment* is when families **exhibited controls on the tracker app** for each of their favourite mobile apps. This transition from passive observation to active control prompts an array of discussions and critical reflections.

These four *triggering moments* correspond closely to the four stages of Kolb's learning cycle (McLeod 2017): *Concrete Experience*, learners encounter a new experience (e.g., see trackers' existence for the first time); *Reflective Observation*, after an experience, learners reflect, question, and discuss (e.g., connect examples to real life data experience); *Abstract Conceptualisation*, learners classify concepts and draw conclusions from events (e.g, families systematically jot down their observations and refine their thoughts); and *Active Experimentation*, learners test out their new ideas and lessons gathered from the experience (e.g., families experimented with controlling trackers based on their reasoning).

While we do not position KOALA Hero as an education tool, our experience has shown that the use of what we call *triggering moments* can be valuable in promoting moments of *reflection, conceptualisation and experimentation.* These *triggering moments* can take various forms: they can be embedded as "pause" moments within systems to motivate user engagement and raise awareness; set up "game rules" that nudge users to explore with existing information; introduce "consolidation phases" for users to conceptualise their experiences; and provide opportunities for users to test their hypothesis. These ideas also resonate with several recent design principles for children, such as the Four Lenses of Play by Bekker et al (Bekker, De Valk, and Eggen 2014), and Project Zero's Agency by Design (*Agency By Design* 2020), which encourage the prioritization of playful elements in digital designs to enhance children's engagement and exploration of new knowledge. We encourage future research to explore further the nuances of these design choices for offering effective engagement for families.

6.5.2 Implications on Family Joint Engagement

One of our research questions was to investigate how KOALA Hero could provide improved support to families in making more informed and joint decisions. Our aim was to explore whether we could provide an alternative to the existing parent-led approaches, which can undermine children's autonomy, hinder the development of their risk coping abilities, and potentially damage trust and communication within families. Hence, we integrated design elements that fostered families' situated reflection and enriched their experience with playful physical components and activity sheets.

Our observations indicate that KOALA Hero enhanced family engagement, with several instances of active negotiation and collaboration for collective decision-making, suggesting the toolkit's positive influence on family joint engagement. Previously, it was commonly assumed that parents, due to their greater expertise, would take the lead in guiding their children's navigation of the digital space (Shifflet-Chila et al. 2016). However, our observations indicate that, particularly in the context of datafication risks, this assumption may not always apply. We noted a shift in expertise dynamics, with parents not always leading and children sometimes initiating discussions. While we observed some tensions in families making joint decisions—for instance, when parents noticed the large number of trackers related to an app and wanted to disable it completely, despite it being the child's favorite and most-used app—the friction tended to be minor, and families generally found a way out. This was mainly because our toolkit offered a more moderate approach by allowing families to turn off only certain types of trackers. Moreover, our toolkit provided a joint learning experience for families, enabling them to explore and collaborate with each other during the process. In general, all families in our study showed a trend towards more balanced engagement over time, with initially dominant members like parents or older siblings increasingly valuing others' inputs, and quieter members becoming more vocal.

In our study, we noted instances that could be described as 'bonding moments,' which seemed to have encouraged joint decision-making among family members. Some of these points include moments to **complete tasks that children might have superior expertise or a greater personal interest instances**, such as clarifying the function of their apps or navigating the data destination view of trackers. Another set of *bonding moments* happened when family members **encountered points of disagreements and had to engage in negotiation** with each other. For example, when families used the Data Cards to discuss datafication risks, it often involved bringing up real-life examples and putting forward critical arguments. Fisher (Fisher 1993) and Wegerif (Wegerif and Mercer 1996) proposed that there are three distinct types of conversations when guiding children's involvement in collaborative activities, namely disputational talk (i.e., disagreements and counter-assertions), cumulative talk (i.e., speakers build positively but uncritically on what the other has said), and exploratory talk (partners critically yet constructively engage with each other's ideas, providing justifications and alternatives). All three types of talks were observed within our family participants. In particular, we noticed that when families engaged with our toolkit, such as using Data Cards and relating app information to real-life scenarios, it appeared to influence a shift from initial disputational talk (simply expressing disagreement without deep engagement or reasoning) (Wegerif and Mercer 1996), to more explorative discussions. In these discussions, statements and suggestions were not just exchanged but also critically examined, with challenges being justified and alternative ideas presented, aligning with the principles of exploratory talk as described by Mercer and Barnes (Mercer and Wegerif 2002; Barnes 2008).

Though exploratory talk is often seen as an ideal outcome for constructive interactions between parents and children (Barnes 2008; Willard et al. 2019), it doesn't naturally occur just because they are using the same device (Brito et al. 2017; Zaman, Nouwen, et al. 2016). Factors like parental dominance (Kutrovátz et al. 2022; G. Wang, J. Zhao, et al. 2021) and challenges in recognizing shared goals (Tomasello and Hamann 2012; Woodward, Esmaeili, et al. 2018) can impede this process. Guidelines such as Playful by Design (S. Livingstone and Pothong 2021) suggest considering age-appropriateness and open-ended play to promote a balanced and explorative learning environment in families. Our observations of 'bonding moments' offer insights that complement these guidelines, highlighting the potential of designing shared experiences to enhance family joint engagement. However, these are preliminary findings and require further empirical investigations, and we recommend that future designs consider incorporating features that might support these dynamics for further investigation.

6.5.3 Implications on Legislative and Policy Development

Our study unveiled grave concerns among families regarding the impact of datafication on children. Families expressed their apprehensions about the influence of data inference and collective profiling on society at large. They voiced discomfort about the apparent data monopoly held by a select few tech giants. Children were particularly anxious that every aspect of their lives was being datafied and used in hidden ways to shape their thoughts and actions (*"have to have a strong mind"* C4, age 11). Parents echoed these concerns and extended them to worries about their children's future in a heavily data-driven society, with inadequate safeguards in place (*"legislation that actually works"* P7). Both parents and children yearned for greater transparency regarding how major platforms (e.g., Google, Meta, Amazon) utilise their data. At the same time, we observed how families transitioned from initially knowing little about the implications of datafication risks to developing heightened cognitive awareness of these issues and a strong desire for change. While our toolkit shows promising impact on users' perceptions, addressing all these complex issues likely exceeds the scope of a single toolkit.

The robust demand from families necessitates a fundamental reassessment of current legislative and policy development regarding children. However, current regulations across the globe mainly address traditional online safety issues for children, such as harmful content and stranger dangers (Parliament et al. 2022; Veale and Zuiderveen Borgesius 2021). Examples include the US's Kids Online Safety Act (*Kids Online Safety Act: Legislation to impose responsibility on online platforms and equip children and parents with tools* 2023), mandating account setting safeguards for minors, the UK's Online Safety Bill (Parliament et al. 2022), focusing on illegal content and age verification, and China's Cyberspace Protection Regulations (*China releases regulations to protect minors in cyberspace* 2023), enforcing protection against inappropriate content with strict penalties. While there is an increasing number of legislative efforts globally specifically targeting children's online safety, subtler risks associated with datafication largely remain under the radar. Recent initiatives such as 'Child Rights by Design' by the Digital Futures Commission (*Child Rights by Design* 2023), and 'Responsible AI for Social Empowerment and Education' at MIT (*Responsible AI for Social Empowerment and Education* 2023) provided useful starting points by addressing child-centric AI technologies, though with less emphasis on data-centric perspectives. Meanwhile, projects such as 'Agile-EDU' (ERSTAD et al. 2023) explore data in educational systems, laying a solid foundation for understanding data in everyday contexts. Our research highlights a clear demand from children and their families for enhanced ability to access and control, particularly from a data perspective, and an immediate impact on their perception of datafication risks through raised awareness. This calls for a comprehensive revision of the current data governance framework related to technologies accessible to children and more targeted legislation addressing families' specific datafication concerns. We advocate for sustained exploration into the creation of comprehensive ethical data governance systems, replacing the current data-driven approach to innovation and re-balancing the power between users and platforms, allowing families to assert their data rights and setting the groundwork for a more ethical data landscape in our society.

6.6 Conclusion

In this chapter, we introduce Koala Hero, a multi-component hybrid toolkit made up of a tracker app for mobile, a set of data cards, and a tasksheet supplemented with worksheets that informs families of mobile datafication risks around them, and encourage enriched discussion on relevant issues. Our goal is to examine whether we could change families' perceptions of datafication risks and families' approach to the discussions of these risks. Through user studies with 17 parents and 23 children aged 10 – 14, we identified more significant awareness of datafication risks from families, their progression of thought process, and a more collaborative family decision-making approach. We identified critical *triggering moments* and *bonding moments* that can nurture family data literacy development and cultivate collaborations and negotiations for family joint informed decision-making.

Our findings in this chapter underscore the potential effectiveness of a child-centred approach compared to the conventional parent-guided method, especially concerning its impact on children's awareness and the subsequent shift in their perception of datafication. Moving forward, our next chapter will move beyond critical awareness of datafication. Specifically, we aim to delve into how, building on this critical awareness, we can cultivate a child-centred approach that empowers children to assert their autonomy over their data.

Statement of Authorship for joint/multi-authored papers for PGR thesis

To appear at the end of each thesis chapter submitted as an article/paper

The statement shall describe the candidate's and co-authors' independent research contributions in the thesis publications. For each publication there should exist a complete statement that is to be filled out and signed by the candidate and supervisor **(only required where there isn't already a statement of contribution within the paper itself).**

Title of Paper	KOALA Hero Toolkit: A New Approach to Inform Families of Mobile Datafication Risks.
Publication Status	Published
Publication Details	Ge Wang, Jun Zhao, Konrad Kollnig, Adrien Zier, Blanche Duron, Zhilin Zhang, Max Van Kleek, and Nigel Shadbolt. 2024. KOALA Hero Toolkit: A New Approach to Inform Families of Mobile Datafication Risks. In Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (CHI '24). Association for Computing Machinery, New York, NY, USA, Article 536, 1–29. https://doi.org/10.1145/3613904.3642283

Student Confirmation

Student Name:	Ge Wang		
Contribution to the Paper	Ge Wang is the principal author who contributed the ideation, design and development of the KOALA Hero Toolkit, conducted all the user studies and all data analysis, and wrote the paper. Jun Zhao initialised the ideation and design of the toolkit, and provided critical input throughout the analysis and write-up of the manuscript. Konrad Kollnig, Adrien Zier (summer intern) and Blanche Duron (summer intern), helped with the testing of the tracker app (which is a part of the toolkit). Zhilin Zhang, Max Van Kleek, and Nigel Shadbolt critically reviewed the manuscript and offered valuable feedback.		
Signature		Date	22/09/2023

Supervisor Confirmation

By signing the Statement of Authorship, you are certifying that the candidate made a substantial contribution to the publication, and that the description described above is accurate.

Supervisor name and title: Dr Jun Zhao, Senior Research Fellow			
Supervisor comments			
Signature		Date	19 March 2024

This completed form should be included in the thesis, at the end of the relevant chapter.

Publications arising from this chapter:

- *Ge Wang, Jun Zhao, Samantha-Kaye Johnston, Zhilin Zhang, Max Van Kleek, and Nigel Shadbolt. 2024. CHAITok: A Proof-of-Concept System Supporting Children's Sense of Data Autonomy on Social Media. In Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (CHI '24). Association for Computing Machinery, New York, NY, USA, Article 536, 1–29.* `https://doi.org/10.1145/3613904.3642294`. ***Best Paper Honourable Mention Award***

7

CHAITok: An Android Mobile Application to Support Children's Data Autonomy on Social Media

Contents

So far in this thesis, we've established that children have a basic understanding of online datafication and a strong desire to act on it. In the previous chapter, we explored how technological interventions can heighten the critical awareness of both children and their families about datafication practices.

Now, in this chapter, we further address the core question: *How can we take a child-centred design approach to develop technical interventions for children; and how might these technical interventions influence, support, and enhance children's autonomy over their data within AI-based platforms?* **(TQ2)**. Specifically, we'll explore and examine technological innovations that empower children to review, manage, and control their datafication practices on AI-based platforms.

All studies in this chapter received ethical approval in accordance with university guidelines for human participant research, under reference R86939/RE001.

7.1 Background and Motivation

In this chapter, we present *CHAITok*, an Android mobile application designed to support children's autonomy over their data, specifically in the context of social media platforms. We chose to narrow our focus to social media platforms in this chapter because they represent the most prevalent and significant type of AI-based platforms in children's daily lives. Through mixed-method user studies involving 109 children aged 10 to 13 across multiple schools, we seek to address three research questions:

- **RQ1:** How, if at all, do children currently experience and perceive the handling of their data on social media platforms?
- **RQ2:** How does CHAITok influence children's user experience and autonomy over their data?

- **RQ3:** What are children's expectation towards having *data autonomy* on social media platforms?

In this chapter, we highlight the unique role of data in shaping children's autonomy on social media. We emphasize *data autonomy* as a fundamental right for children and call for further research, design innovation, and policy changes focused on this critical issue. The contribution of this work is as follows: (1) we investigate how children's current experiences with social media affect their data autonomy perceptions; (2) we design, develop, and assess a proof-of-concept system that demonstrates how popular social media platforms like TikTok can be integrated with autonomy-supportive features that promote children's autonomy over their data; (3) we provide crucial insights into the expectations of children aged 10 to 13 regarding data autonomy online. Our research offers vital insights into how children presently perceive data autonomy online, and how we can empower children's data autonomy through a socio-technical journey. Our findings inspire design recommendations to respect children's values, support children's evolving autonomy, and design for children's digital rights. We emphasize *data autonomy* as a fundamental right for children and call for further research, design innovation, and policy changes focused on this critical issue.

7.2 Data Autonomy: Scope and Definition

To establish the foundation for our design and development, we revisit our working definition of *data autonomy* as outlined in Chapter 3. To recap, the central themes of our working definition are:

7.2.1 Unpacking Data in Data Autonomy.

- The first element, *Data Collection*, defined by Solove as "the watching, listening to, or recording of an individual's activities" (Solove 2002). In the context of online platforms, this involves the gathering and storing of user information. A significant dimension within this element is data sharing –

how platforms distribute the collected information to other entities, such as internal departments, partner companies, third-party vendors, or advertisers.

- The second element, *Data processing*, refers to the process in which digital platforms process, analyze and make use of collected user's data. In the context of social media and other online platforms, this involves using users' data to generate services and content, often supported by algorithms.

- The third element, *Data inference*, involves the further processing and analysis of user data by online platforms to evaluate or predict personal aspects, such as work performance, economic situation, or health. This aligns with Solove's categories of data dissemination and invasion, as well as the value-generation aspect of datafication. What sets *Data Inference* apart from *Data Processing* is its capacity to *learn* about individuals or groups about their personal aspects, going beyond simply processing data for services like video recommendations.

7.2.2 Unpacking Autonomy in Data Autonomy.

- The first, *Cognitive Autonomy*, which refers to an individual's ability to think independently. It involves self-governance of the mental action or process of acquiring knowledge and understanding, to evaluate thought, to voice opinions, critically evaluate information, and to form personal beliefs.

- The second, *Behavioural Autonomy*, which refers to an individual's capacity to act independently, make their own decisions, and carry out actions based on their personal judgment and values. It involves the ability to self-regulate, take responsibility for one's actions, and behave according to one's own decisions and choices.

- The third, *Emotional Autonomy*, which refers to an individual's ability to identify, understand, and manage their own emotions independently. This involves the capacity to distinguish one's own feelings from those of others, handle emotional dependence, and maintain emotional stability without relying excessively on others.

7.2.3 Data Autonomy: A Working Definition

Building on the aforementioned concepts on “data” and “autonomy” across multiple disciplines. We now present an overview as a working definition of data autonomy in the digital realm:

> *Data Autonomy*, can be summarised as the empowerment and capability of individuals to comprehend, exercise control over, and reflect on the collection, processing, and inference of their data within the digital realm. This concept underscores the importance of informed understanding, active decision-making, and critical reflection in the way personal information is handled and utilized in online environments.

7.3 System Design and Development

How can we design for children’s data autonomy on their daily-used social media platforms? We selected TikTok as our primary platform for implementation due to its global popularity among children (Media 2022; Ofcom 2023). In this section, we present the design and development of *CHAITok*, an Android mobile application that is constructed upon the foundation of TikTok and rooted in our working definition on data autonomy. CHAITok is not intended to be a complete replacement for TikTok; instead, it functions as a *proof-of-concept* system built on TikTok’s existing framework. Our aim is to explore how can we foster children’s data autonomy within a social media environment that is already familiar to them.

7.3.1 Ideating

Our design objective is to introduce design elements that support children’s development of different forms of autonomy, as defined in our working definition of *data autonomy*. This support of autonomy is expected to occur as children engage with TikTok’s datafication practices, including data collection, processing and inference. We initiate our design process by creating a separate design canvas for each of the three phases of datafication. Within each design canvas, co-authors suggested a set of design features aimed at supporting one of the three forms of autonomy. These feature sets for each design canvas were then collectively evaluated by co-authors based on criteria like expected impact on users, novelty, and technical feasibility. The highest-scoring features for each design canvas were implemented.

7.3.2 CHAITok

In the CHAITok app, we implemented the three design canvases as sequential pop-up panels that appear after a set of 15 video recommendations. This means that while using CHAITok, rather than endless scrolling through videos, users will be shown a pop-up data panel after every 15 videos viewed. This staged approach guides users through the core aspects of datafication: from data collection, processing, to data inference. While users may experience diverse datafication practices on social media, we used the pop-up panels to direct their attention to specific sets of practices.

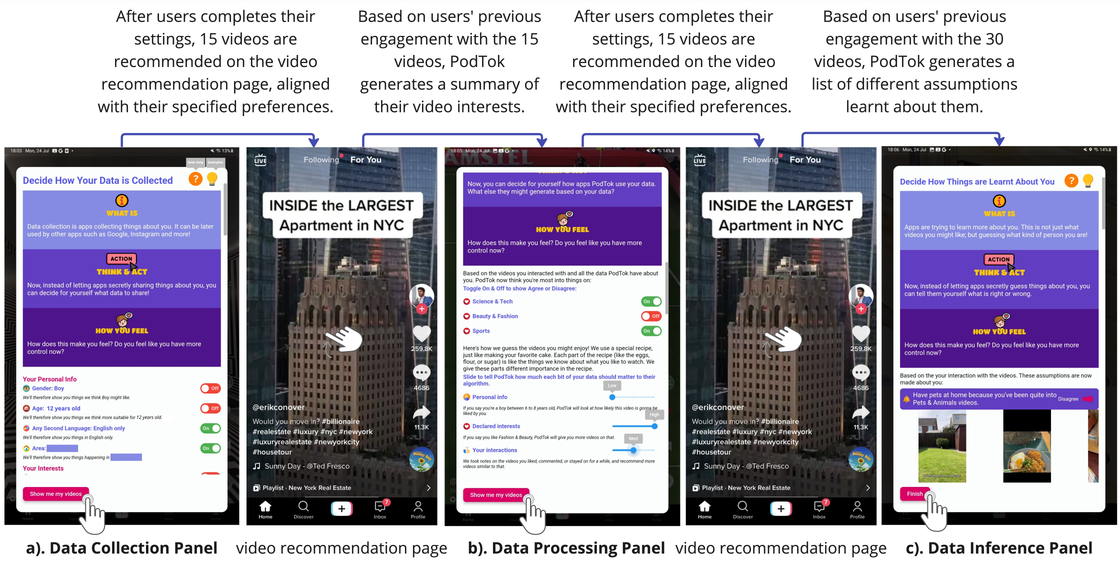


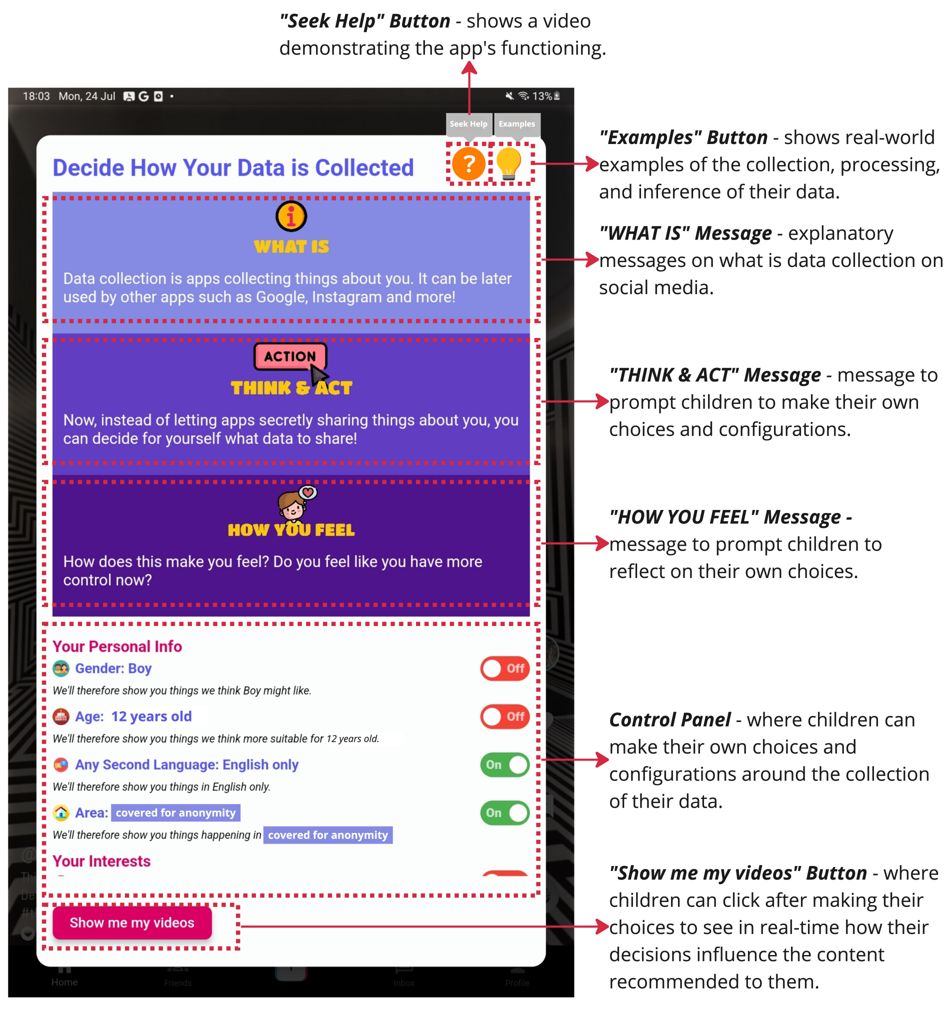
Figure 7.1: A visual representation of the user interaction flow in the CHAITok application from panel to panel.

Upon launching CHAITok, users first see the data collection panel. Users could configure the panel by setting what personal data they allow CHAITok to use for creating video recommendations. They can then click 'Show me my videos' and receive a set of 15 recommended videos. Our video recommendation page in CHAITok resembles TikTok's 'For You' page, where users can swipe up to view videos, as well as like and comment, mirroring typical TikTok behaviors. All user interactions are recorded. As users swiped through these 15 videos, the data processing panel is triggered, where users can adjust a new set of configurations, which can influence the next set of 15 video recommendations. Similarly, after swiping through the second set of 15 videos, the data inference panel appears. Figure 7.1 shows a visual overview of the user flow in CHAITok and we introduce the detailed features of each panel in the sections below.

Data Collection Panel (Figure 7.2)

This panel appears upon launching the CHAITok application, and is designed to support children's autonomy on the collection and sharing of their data. To support children's ***cognitive autonomy*** we include the following features: a **"Seek Help"** button that displays a demo video on navigating CHAITok; an **"Examples"** button that presents real-world examples and news related to datafication on social media; and a **"WHAT IS"** explanatory message which offers a brief explanation of what data collection entails on social media platforms.

In the section dedicated to children's ***behavioral autonomy***, users are first nudged to review and manage the various data types they wish to share with CHAITok through the **"THINK & ACT"** message. After reading through these messages, users can then use the **toggle buttons** in the **"Control Panel"** to indicate their preferences in two key areas: 1) personal details to be used by CHAITok for video recommendations, including age, language, areas, as well as user-selected video interest categories (e.g., Pets & Animals, Gaming); and 2) sharing or withholding online activity data from platforms like Google (e.g., "your browsing history on Google"), or Amazon (e.g., "your purchase history on Amazon") with CHAITok. The selection of data types stems from our analysis of TikTok's privacy policies (TikTok 2023). We do not claim this offers the most comprehensive representation of data that might be collected by TikTok, but serves to stimulate children's reflection and consideration of the data collection process.. For each study, we pre-filled some personal details based on the year groups we visited, including age (e.g., "12 years old"), area (e.g., [school's location]), and second

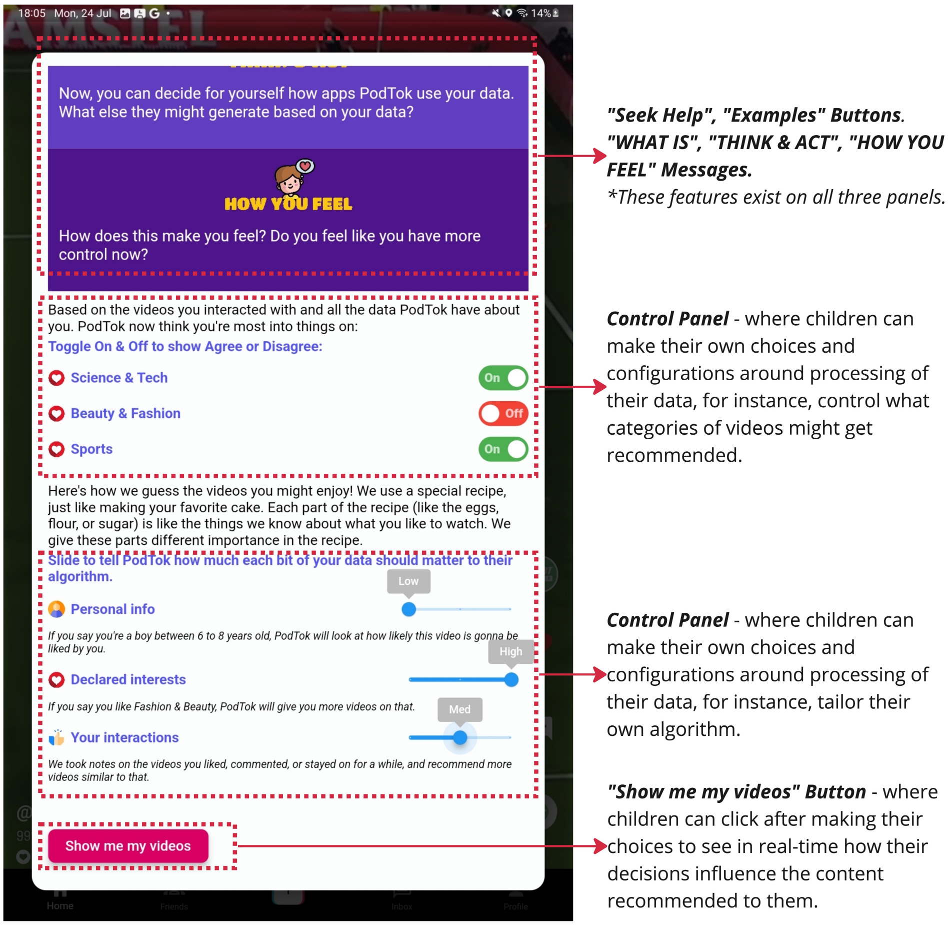


a). Data Collection Panel

Figure 7.2: Screenshot of Data Collection Panel. All three panels follows the same interface design, and contains a "Seek Help" button (shows a demonstration video on how to navigate CHAITok) , an "Examples" button (shows real-world examples of the collection, processing, and inference of their data). A "WHAT IS" message (explanatory message on data collection/processing/inference). A "THINK AND ACT" message (prompting children to make their own choices). A "HOW YOU FEEL" message (prompting children to reflect on choices). A control panel (where children can make their own choices on the collection/processing/inference of data). A "Show me my videos" button (children can click to instantly see how their choices affect the recommended content).

language (e.g., "French" if teachers informed us they were studying it). While not aiming for perfect accuracy, these steps were taken to help children more contextually make their decisions.

Finally, to support children's development of ***emotional autonomy***, they are encouraged to reflect on their sense of autonomy by the **"HOW YOU FEEL"** message in the panel. Once completing all the configurations, users can click **"Show me my videos"** to see 15 recommended videos on their video recommendation page. For instance, if users activate the "Second Language: French" or "Area: [City X]" toggles, they will receive recommendations for French-language or locale-specific videos, respectively.



b). Data Processing Panel

Figure 7.3: Screenshot of Data Processing Panel.

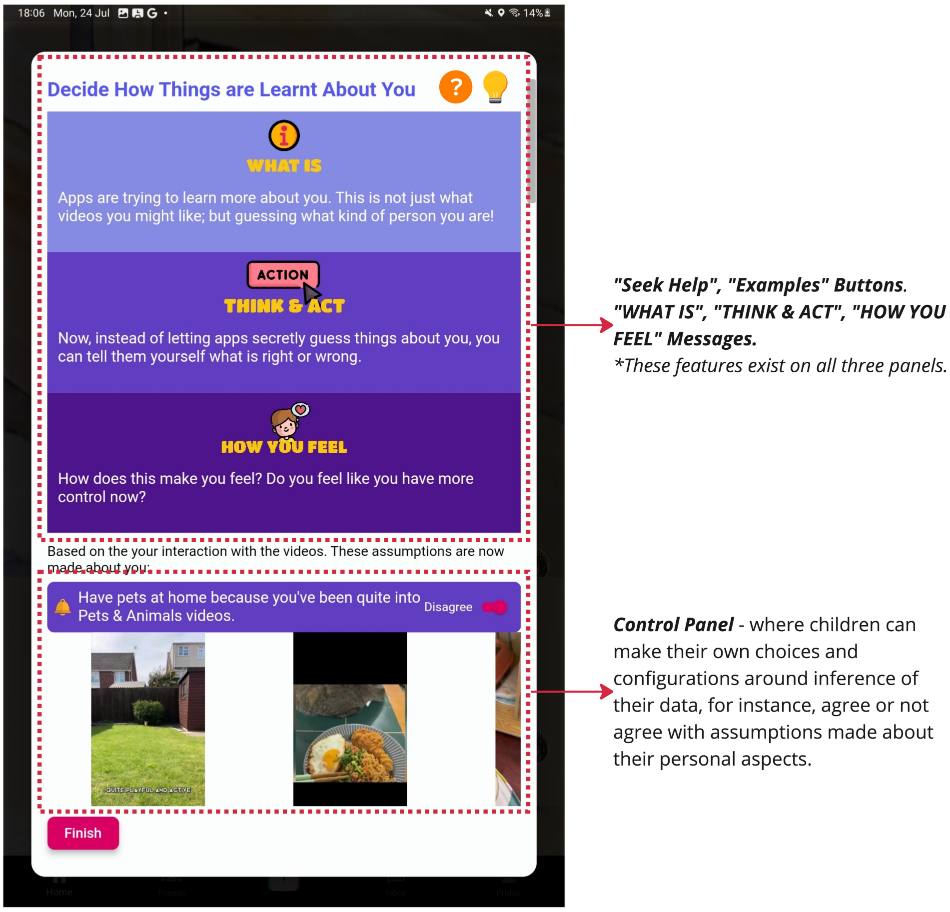
Data Processing Panel (Figure 7.3)

After users have scrolled through the 15 videos, the *Data Processing Panel* pops up, designed to support children's autonomy on the processing of their data. The features around ***cognitive autonomy*** and ***emotional autonomy*** stayed the same.

The **Control Panel**, which aims to support children's ***behavioural autonomy***, enables users to adjust two key features: 1) using **toggle buttons** to to modify how their data is processed. For example, the app analyzes tags from videos they engaged with and presents the top three interests, such as Gaming or Beauty & Fashion. Users can either accept these summaries or make adjustments; 2) using **slide bars** to customize CHAITok's recommendation algorithm by adjusting the importance of different data points, including personal information, declared interests, and video interactions. These slide bars enable users to set the weight of each data point in the algorithm to low, medium, or high. After configuring the settings, users can click **"Show me my videos"** to receive 15 new video recommendations according to their customisation. For instance, if they set a high weight for interactions data, the algorithm will focus more on their likes, comments, and watch time, suggesting videos similar to those that they engaged with. If they prioritised personal info, the algorithm will recommend more videos based on their personal info, like kid-friendly content or those relevant to their locale.

Data Inference Panel (Figure 7.4)

After scrolling through another set of 15 videos, a *Data Inference Panel* appears, designed to strengthen children's control over data-driven assumptions made about them. The features around ***cognitive autonomy*** and ***emotional autonomy*** stayed the same. The ***behavioral autonomy*** features include a **Control Panel** that allows users to review and control assumptions derived from their interactions with the last 30 recommended videos. For instance, if a user often likes, comments on, or stays longer on pet-related videos, we might assume, "You have pets at home because you've been quite into Pets & Animals videos". CHAITok computes these assumptions based on methods outlined in previous articles published by Google on ad personalisation (Google 2023). We do not claim these assumptions to be fully accurate or exactly representative of social media platforms' actions. Our goal is to provide simplified examples to children, helping them understand how assumptions could be made about them using their data. Users can use **toggle buttons** to



c). Data Inference Panel

Figure 7.4: Screenshot of Data Inference Panel.

agree or disagree with each assumption. Finally, users can click **"Finish"**, which marks the end of their activities on the app.

7.3.3 Implementation & Piloting

For CHAITok, I chose not to use TikTok's API, instead opting to build the app from scratch in Android Studio for Android 13. This approach allowed me to closely replicate the TikTok user experience while maintaining full control over content safety and recommendation algorithms. My work here underscores a commitment to innovation and the design of user-centric digital media platforms, emphasizing my original contribution. I integrated the app with Google Firebase to leverage its robust infrastructure for real-time data synchronization, user authentication, and to support the development of custom recommendation algorithms and content safety measures. I sourced video content from a meticulously curated database on Cloud Firestore, ensuring every piece of content adhered to strict appropriateness standards. This careful curation process highlights the originality and ethical responsibility I placed on content management. Additionally, I used Firebase to collect and store interaction data, which allowed me to continuously refine and enhance the app's recommendation algorithms, ensuring they remained dynamic and tailored to user preferences. Concerning privacy and safety, especially for children, I took proactive steps by setting up pre-registered anonymous accounts for child users. This strategy was designed to protect their identities while providing a secure environment for interaction with the app.

Our research team initially piloted the CHAITok app on our Android devices, conducting multiple rounds of configuration checks on the three panels and reviewing the recommended results. This process helped us identify several usability challenges, from font size issues to delays in updating recommendations, which were subsequently resolved. We then piloted CHAITok with three active TikTok users, aged 10, 11 and 13, who helped us identify further usability issues, including trouble locating configuration buttons and subtle changes in recommendation that were hard to notice. We addressed all these concerns to further refine the app's usability and functionality.

7.4 User Study

7.4.1 Participants

For our user study, we recruited 109 children aged between 10 to 13 through school recruitment. The participants were required to be active TikTok users. We selected this age group for several reasons. Firstly, reports have highlighted that this is a critical period during which children move from parent-supervised online activities to active social media participation (Ofcom 2023; Ofcom 2022; Ofcom 2021). Additionally, this age bracket corresponds to the period when many children across the globe progress from elementary to middle schools, a shift often leading to an increase in their online social interactions (Evangelio et al. 2022; Yao et al. 2022). It is thus essential to assess the experiences of children at this stage regarding social media datafication, as it often marks their entry into the digital world. Among the 109 participants (M = 11.8, SD = 0.95): 27 participants were 10, 27 were 11, 27 were 12, and 28 were 13 years old. 53 were girls and 56 were boys. We visited a diverse array of 8 schools: two state schools, two grammar schools, one private school, one Catholic-faith school, and one Muslim-faith school.

7.4.2 Procedure

We carried out 27 mixed-method studies involving 109 children between April and July 2023, post-IRB approval. Each study involved 3-5 children and was led by 2 researchers in school classrooms, using Android tablets with the CHAITok app pre-installed. The children in each session were usually classmates, fostering a more dynamic interaction and mitigating the cold start effect. Every study spanned a duration of 90 minutes and consisted of three sessions:

Warm up (10-min)

We kicked off the study with an icebreaker game. Children tossed a ball and the catcher named their favorite social media platform, followed by a brief thought on how that platform might handle their data. We asked children to elaborate on their responses without judging their answers as right or wrong. The questions aimed to provide initial insights into their understanding of their data on social media platforms.

Session 1: Tasks with CHAITok (40-min)

In this session, participants were introduced to CHAITok via a 2-minute video that summarized its key features and instructions. This video is also accessible in-app under the “Seek Help” button. They were then given 5 minutes to freely explore the app and ask any questions to the researchers. Next, children were presented with three activity sheets, each containing the tasks associated with each of the three panels (see Figure 7.5) that correspond to the datafication process. Though they were encouraged to record their thoughts and actions on these sheets, it wasn't mandatory. The task sheets were introduced to provide children with a structured framework to consolidate their thoughts (Arvanitis, Barrable, and Touloumakos 2023; Skinner and Belmont 1993), and children were reminded to verbalise their thought process for the audio recording. During tasks, children were encouraged to collaborate as a group. Researchers observed and noted points for discussion in the follow-up interview, intervening only when technical support is needed. Upon completion, children were provided with a 10-minute break during which they could hydrate and enjoy some snacks, which were conveniently arranged for us by the schools.

Session 2: Semi-structured Interview (30-min)

In this session, we carried out a semi-structured focus group interview. We began by prompting the children to share their most surprising findings from the session and identify the most exciting features. Next, we encouraged them to reflect on and articulate the concept of *data autonomy*. We directed the children to focus on three main questions: 1) After participating in the session, did they feel a greater sense of autonomy over their data on social media platforms compared to their prior experiences? 2) What additional features or support would make them feel greater sense of autonomy with their data? 3) How would they define 'data autonomy' in the context of social media, and did they feel they currently had this autonomy on existing platforms?

We tried to use the exact term *data autonomy* in our interview questions as much as possible to accurately capture children's expectation of this concept. While some children shared their thoughts insightfully, others found it abstract or confusing. To mitigate this, we used alternative phrasings like "in control", "empowered", or "have a say". These phrases have also been used in previous HCI studies on understanding users' sense of autonomy and agency, which were found to be more understandable and capable of capturing study participants' judgments (Benson 2010; Metcalfe and Greene 2007; Lukoff, Lyngs, Shirokova, et al. 2023). This also served as a great starting point for children in our study to elaborate on their

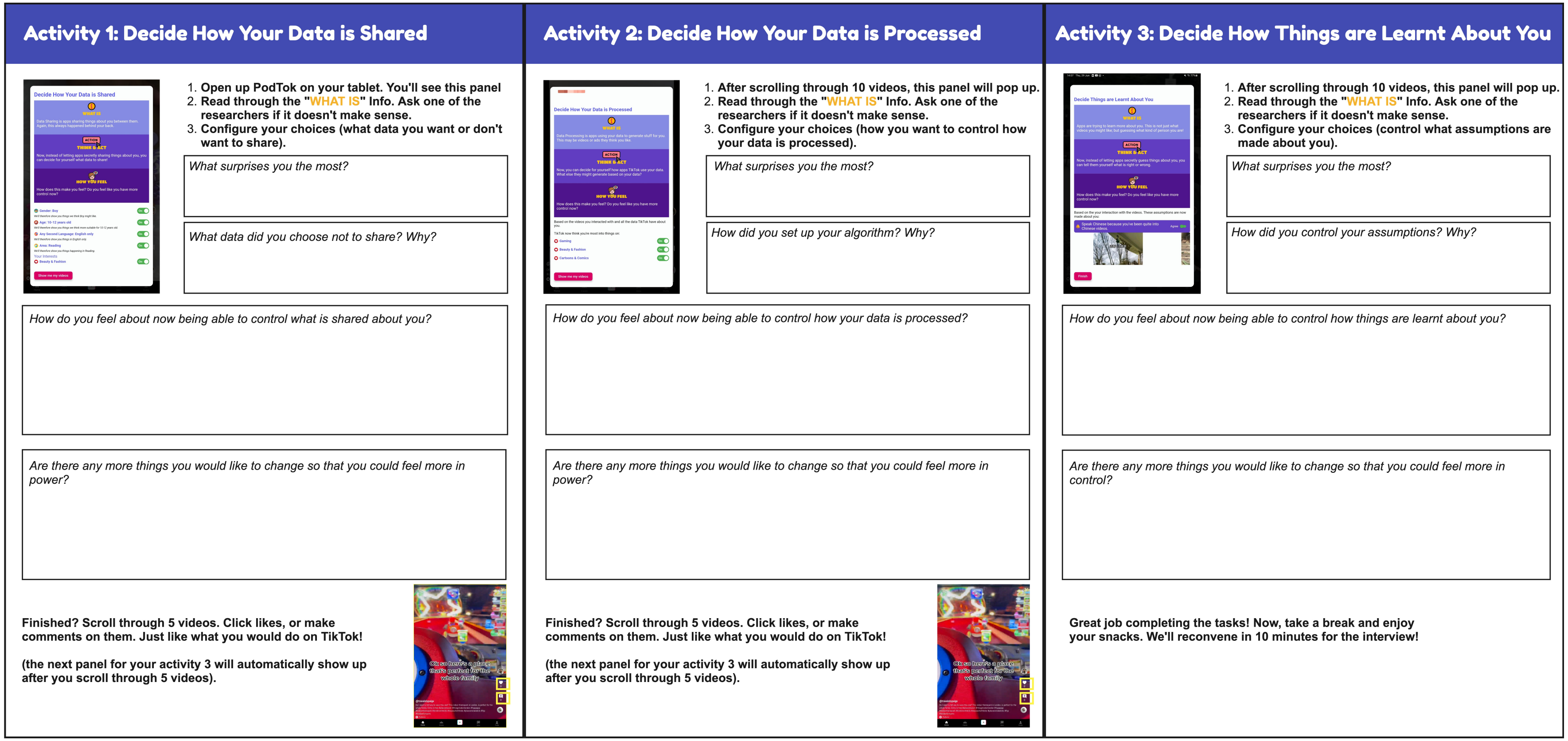


Figure 7.5: Activity sheets used for guiding children through the tasks on each panel.

thoughts about data autonomy. Any additional points noted in the observational notes from Session 1 were also addressed.

7.4.3 Data Analysis

All user studies were audio-recorded, yielding 2673 minutes of data that were transcribed for analysis. We employed thematic analysis (Braun and V. Clarke 2006) to generate a code book, and cross-referenced transcripts with observational notes and children's written responses on the worksheets to compliment our analysis process. For a comprehensive explanation of thematic analysis as a data analysis methodology, please refer to section 4.2.3.

7.5 Results

To address our research questions, we first outline children's general experiences and perceptions about data handling on social media platforms **(RQ1)**. Next, we discuss their user experience with the CHAITok app and how it influenced their sense of autonomy over data **(RQ2)**. Lastly, we explore their visions for gaining more autonomy over their data, including desired design features and expectation towards *data autonomy* **(RQ3)**. While our study is organized into three sessions, findings aren't confined to specific parts. For example, children might discuss their current social media experiences or interactions with CHAITok at any point. The study is best viewed as an integrated process where themes can emerge and be explored at any stage. Participant quotes are presented with their ID and age for context.

7.5.1 Children's Current Experience and Perception on How Social Media Platforms Handle Their Data

All 109 children reported using various social media platforms, with TikTok, YouTube, Snapchat, and WhatsApp being the most popular. Nearly all children knew they technically shouldn't be using these platforms due to age restrictions, but explained they still did, *"My mom set it up to stay connected."* (P2, age 10), or *"Everyone in class is using it."* (P38, age 12).

Although our primary focus wasn't on digital literacy, our findings do indicate that children between 10 and 13 have considerable understanding of how social media platforms use their data; and we did not find a significant difference in the understanding across this age bracket. The majority of the participants were aware that platforms **collected** their data, including personal details, such as mobile numbers and email addresses, and online behaviors, such as engaging with videos on TikTok, sharing content on Instagram, and connecting with friends on Snapchat. While many children had a vague idea that their data would be **shared** across platforms, only a few could clearly describe this process. Some referred to it as *"selling my data"* (P12, age 10), while others more insightfully called it their "digital footprint", explaining that *"Whatever you do online, it leaves a trace and it follows you."* (P37, age 13). We also discovered that most children in our study were well aware that social media platforms would **process** their data, such as TikTok using it to curate better videos for them. There were some knowledge gaps among some children regarding the potential real-world consequences that could be **inferred** from their data. Initially, about two-thirds of the children thought data was only used to improve their video experience or to target advertisements. However, interacting with CHAITok – by reading the "WHAT IS" information, viewing examples, and reflecting on their choices – led them to consider more nuanced implications of their data, such as external influence, nudging, and behavioral engineering, *"Your identity, your likes and dislikes, and everything in your life is being learned by them."* (P15, age 10).

While many children referred to social media platforms as "their favourites" and would spend *"anytime off school on them"* (P42, age 11); almost all expressed concerns with the ways their data is being handled. They shared experiences where they felt a lack of autonomy, or in their own words – *"don't have autonomy at all"* (P37, age 13) over their data.

A Passive and Disrespectful Experience

To start with, many children (around one in three) described their experience with data on social media platforms as quite a passive experience, *"I don't think we have autonomy. Once you enter your data, you can't change or delete it; it's already out there."* (P100, age 12). This perception extends beyond basic data collection to how algorithms shape their experience, essentially turning them into "zombies" by making all the decisions for them, *"The TikTok algorithm is deciding for you what you can like, and the more your data is collected, the more you just become a zombie."* (P40, age 13). Many children were concerned about being manipulated by the processing and interpretation of their data, *"If they just steal your information, it doesn't matter that much. There's way to stop it. But it's annoying if they use your data to put ideas into your head."* (P89, age 11). Some proposed that the handling of their data should be a 'two-way' thing, and otherwise "disrespectful": "*They've made us agree to their policies, but it's should be a two-way system. We have our conditions too, and neglecting that is disrespectful."* (P87, age 11).

As a result, children have shown great distrust in these social media platforms, describing practices as 'being tricked" and "behind your back". Interestingly, many children believed that platforms were intentionally undermining their autonomy by normalizing certain behaviors, *"[P103, age 13] The reason people are normalized to this is that companies are intentionally doing it. [P104, age 13] Yeah, like by pushing more and more 'allow all cookies' and using the exact same wordings and format on everything."* Due to this distrust, some children have developed various folk theories, ranging from believing their data might be sold to foreign governments or the dark web, to suspicions that their devices are always listening: *"[P42, age 11] Doesn't TikTok and Instagram do that, listen to you? [P43, age 11] Yes, my mom mentioned my sister's ballet classes and then we saw ballet shoes in her size on Instagram."*

A Detrimental and Helpless Experience

Many children also expressed concerns about their data-related well being on social media. About a third found it hard to disengage from these platforms, and some even reported sleep issues when using phones before bedtime. Many children connected these experiences to how social media processes their data, saying platforms only aim to capture attention: *"They don't care what's good for you. Like gaming? Let's send you more and who cares it keeps you up all night!"* (P18, age 10). Some children complained that their data is used to push content without regard for their feelings. *"If you search for Covid-19, they flood you with more news. Twitter aims to grab our attention, even with disturbing images. Do people really want to see that?"* (P93, age 12).

At the same time, children felt helpless against these practices. Many considered apps like WhatsApp and Snapchat essential for communicating with parents and friends. They also lacked confidence in "competing against big companies", noting *"they can piece together everything about you, and there's no undoing that"* (P63, age 11). Some children insightfully commented on the lack of 'good practice' in the industry, *"If even one company tries to do good, it would set a good example for others to follow."* (P73, age 13). Meanwhile, many felt they had been normalized to current datafication tactics, *"Honestly, I haven't given much thought to it since I was born into this system, it was there before me. If you were older and saw the increase in these things, you might have more to say."* (P31, age 12). Children recognised the problems but felt powerless, some even blamed themselves for lack of resilience, *"It feels like it's my fault for not reading their terms. Who am I to complain about options?"* (P23, age 10).

7.5.2 The CHAITok User Experience

To glean a comprehensive understanding of their experiences with CHAITok, we amalgamated data from diverse sources, including audio recordings of group activities, interviews, and observational notes, thereby facilitating a multifaceted insight into the participants' experiences.

Sense of Security

Initially, many children described their CHAITok experience as 'safe and secure,' a term noted by nearly one-third of them. Typically, they begin by reading the 'WHAT IS' section, which often triggers group discussions, and then scroll to view data might get collected, including activity and interests. This triggers various reactions, from surprise to concern, with a 10-year-old participant (P11) using the term *"reveal the reality"*. Most children were surprised by the amount of data collected about them and questioned why platforms like TikTok needed such information, leading to immediate discussions about the potential harms of extensive data collection.

Many children experimented with various data types, and some chose to turn off their online activity data, expressing fears of future misuse, *"I hadn't realized how freely they could use that data until I saw this myself. They would pretty much know every single thing we're doing when we get older."* (P16, age 10). We observed that children's inspired thinking often led them to promptly take protective actions, for instance, most children quickly deactivated personal information like name, age, gender, and location in the data collection panel, *"dangerous to give these out."* (P51, age 11). Similarly, many children chose to minimize personal information when setting up their algorithms, only keeping their declared interests, *"The only thing that should matter to the algorithm is our video interests, not who we are."* (P93, age 12).

When later asked about their feelings regarding their choices and decisions, nearly all children expressed feeling "safe", "secure", and "more self-assured". Some elaborated that they now have an "ease of mind" and are more willing to trust and use the app, *"With this app, I feel reassured that we're fine and protected. The fact that I can protect myself ease my mind a lot and definitely made me trust the app a lot more."* (P72, age 11).

Sense of Empowerment

Another significant theme reported by children was an increased feeling of empowerment and interest because they could choose what information to consume, which was believed to be a direct consequence of having control over their own data. Aside from immediate worries such as data leaks or password exposure, a large proportion of the children (about 80%) were also concerned about indirect harms such as data being used to subtly influence them. Children were surprised to see their data choices (e.g., toggling off data types) directly led to immediate changes in video recommendations. This sparked discussions on how different users could encounter completely unique content on the same platform; and due to the ubiquity of these platforms, companies could effortlessly manipulate global content: *"Maybe because platforms like TikTok are controlled by China, the Chinese government might try to push ideas that Americans might not find very good, but that the Chinese might think are acceptable."* (P76, age 12).

As a result, we noted that many children made choices based on subtle concerns about how their data might affect them. For instance, some didn't want their gender information collected, noting that it could lead to stereotypical content: *"Even if it knows that we're girls, it's kind of sexist because it's going to give us videos based on what a girl likes. But girls like all kinds of stuff."* (P57, age 11). Language-based concerns were also raised, *"Like what's happening between Russia and Ukraine. They might show you different things if you choose different language."* (P73, age 13). Moreover, certain kids deliberately opted out of categories like Cartoons to avoid time-consuming content, *"Typically, on TikTok, for cartoons, they would give you one from a whole series. And then it becomes addictive."* (P23, age 10).

We observed increased engagement when children could adjust settings and see real-time changes to their content feed. Even initially indifferent participants became more vocal in discussions after interacting with controls. Many reported feeling empowered, *"Now, I'm actively making choices, rather than just whatever is thrown at me."* (P97, age 13). Some children expressed feeling more prepared and comfortable with online content, *"To be honest, seeing all the beautiful girls on Instagram sometimes makes me feel bad. But now I realize the whole world isn't like that, and I can adjust my filters on this app (CHAITok) to see things differently."* (P35, age 11).

Sense of Respect

Another significant theme reported by children regarding their CHAITok user experience was a feeling of respect and equal treatment. For instance, while reviewing the assumptions inferred about them based on their past interactions, they quickly opposed many of them and engaged in critical discussions about the real-world implications, *"All this made me question if I'm just part of a data cycle. I give the data, they sell and use it, and then its consequences loop back to me. You think you're the customer online? Actually, you're the product."* (P46, age 12)

Children strongly wanted to be seen as equal individuals by social media platforms and to have a voice in their data experience. Their choices were often guided less by specific concerns like harm or influence, and more by a desire for control over their online presence. For example, some disabled data categories they felt they couldn't control such as demographic info, arguing for platforms to collect only modifiable data like activity. Almost all were proactive in accepting or rejecting assumptions about them, indicating a strong wish to shape their online image. We also saw many instances of children experimenting with settings to observe real-time effects and then reflecting on those choices, *"We tried reducing personal info and denying all interests. It's cool that it actually listens to us!"* (P51, age 11).

When reflecting on their choices, many children said they felt more respected and "special", *"It makes me feel special. It's asking my opinion, focusing on what I want and don't want."* (P75, age 12). An 11-year-old (P61) added *"I'm liking this app way more; it's not like YouTube or TikTok at all. It's designed with us in mind. Because it feels like I'm not being looked down in a way. Unlike the apps get to decide who I am and how they see me."*.

What Hinders Children and Why

While most children had positive experience with CHAITok, we did notice instances where the children were stuck. We carefully analyzed these situations to identify obstacles they faced:

Lack of supporting context. As elaborated previously, most children had a decent grasp of datafication. But for those with less background knowledge, like not knowing that activity data would be collected, they were often more hesitant to engage in discussions. For children who had a foundational understanding of various datafication aspects, the challenge often laid in bridging the gap between theoretical knowledge and real-world applications. For instance, while many recalled learning about algorithms in their ICT or computing classes, some struggled to relate this knowledge to how platforms like TikTok curate content, causing occasional pauses in initial conversations. A related obstacle was the absence of context to make sense of information presented in CHAITok. Even when children had the essential knowledge, they sometimes described the information as *"overwhelming"* (P58, age 13) or had trouble discerning *"what was important"* (P29, age 10), indicating a need for additional guidance to connect the dots.

Disinterest and lack of confidence. We noticed instances where some children showed disinterest in managing their data, which led to a lack of confidence in their capacity to make informed decisions. Although a small group (8 out of 109 participants) expressed this lack of interest, citing they *"don't see the point"* (P4, age 13), this disinterest often stemmed from a previously mentioned lack of context. For example, children who assumed their data was only used to improve services were generally indifferent to adjust their settings, *"It's good to have all these on? So that they can have all the information to give me the best videos and trends."* (P69, age 11). Among those who were somewhat aware, comments like *"I don't see any direct danger anyway."* (P7, age 13) revealed a level of disinterest. Some even felt resigned to datafication as inevitable, *"That's just how it is. There's nothing we can do."* (P35, age 11). Similarly, doubts about the efficacy of change were expressed, compounded by a lack of trust in all online platforms: *"This idea is great. But I doubt even if you asked them to implement CHAITok, there are still various ways they can get away."* (P99, age 12).

7.5.3 Children's Conceptualisation on What *Data Autonomy* Means for Them

Towards finishing of activities, children began discussing potential design improvements for greater autonomy. In follow-up interviews, we probed this topic further, asking what additional designs could increase their sense of autonomy over their data, and what *data autonomy* meant to them. In this section, we first outlined the children's practical suggestions for enhancing data autonomy, then their conceptual understanding and expectations of the term.

Children's Proposed Autonomy-Supportive Design Features

Among children's suggestions for enhancing their sense of autonomy over data, a recurring theme focused on ***enhancing data safety and security***. Ideas ranged from encrypting personal details to real-time alerts for suspicious activities on their social media accounts. Some children also considered the safety of data in transit, like a proposal for a "private web structure" aimed at *"restricting users' data from easily getting around across platforms and apps"* (P100, age 12).

About one in four children suggested features to ***support self-reflection of their data***, from basic screen-time tracking, to more innovative ideas like an adapted version of "screen data" that would quantify the amount of data they have shared and for what purpose. Some expressed a desire for tools that promote healthier social media habits, saying, *"Since they have so much of your data, they could use it to guide you into healthier behaviors. Support your autonomy rather than restricting it."* (P72, age 11). Beyond individual behavior, some also proposed features for societal benefit, *"For posts with crucial information, like those about vaccines, make sure they prompt users to pause and reflect."* (P105, age 13).

Many children looked beyond just interface design to consider what they saw as more autonomy-supportive algorithmic design. They imagined systems where users could ***create and control their own recommendation algorithms*** rather than rely on a one-size-fits-all platform algorithm. For instance, some suggested letting users input their own hashtags or keywords to influence content tagging (P71, age

11). Others proposed a more community-driven approach, allowing people to form groups and collaboratively design their own content recommendation algorithms. As one child put it, *'It would be good if you and your friends could create algorithms just for your group, because you know what each other likes better. Not just like a wide thing that's the same for everyone."* (P45, age 12).

Children's Conceptualisation and Expectation towards *Data Autonomy*

How do children conceptualise the concept of *Data Autonomy*, and what are their expectations then? In our exploration, we identified three major themes:

Data Autonomy as deciding for your own data. Most children conceptualize the term through the lens of control and awareness. One child expressed it as, *"Data autonomy means having the power to control what's happening with your data, and what are the consequences."* (P26, age 10). Some focused on the more practical aspects of how their data is used, *"It means having control over the content I see, so I can get what I actually want. I may watch a lot of gaming content, but that doesn't mean I really WANT it – it's just addictive."* (P56, age 11). Some children extended the concept of data autonomy to include a "mindset", *"It's not just about control but also feeling secure and safe in your choices."* (P57, age 11).

Data Autonomy as resilience over own data. Another major theme of data autonomy is resilience – the capacity to resist external control arising from the use of their own data, *"having free will and making your own decisions without being nudged into things using your own data."* (P56, age 11). Some extended it to the concept of "identity", *"It means having your own identity online. You are in control of your own landscape."* (P109, age 13). Some children also focused on data autonomy as a collective thing, *"Your autonomy is impacted by what your friends share or recommend, so it's not just a personal matter."* (P95, age 12). Alongside this concept of collective influence, several children emphasized the importance of self-reflection and personal responsibility, *"Data autonomy also involves you taking responsibility for your choices and understanding where that can take you."* (P23, age 10).

Data Autonomy as developmental competencies to be learnt. Children often express the view that the concept of data autonomy should be age-dependent and gradually acquired, *"It should be based on one's age and experience. Cause you don't want to give a 3-year-old too much autonomy."* (P87, age 11). Others describe expectations for how data autonomy should be taught, rather than simply handed over: *"I think data autonomy should be taught on how to approach things, not just given as 'here's your freedom, off you go."'* (P22, age 10). This reflects a broader sentiment among children that data autonomy is not an inherent trait but rather a skill or competency that needs to be cultivated. A 13-year-old (P108) offered an insightful perspective, *"Data autonomy isn't something you just have; it's a skill you learn, like Maths or Physics. You don't know it automatically at 13 or 14; it needs to be built up. So perhaps starting with simple exercises could help make you realise that you too can make own decisions, instead of just becoming normalized to things."*.

7.6 Discussion

7.6.1 Impact on Children's Sense of Autonomy Over Their Data

Before diving deeper, it's important to clarify our choice of terminology. We specifically used "autonomy" when describing working definitions and design goals, and carefully switched to "sense of autonomy" when discussing methods and findings as it often contains self-reported data. It's worth noting that using self-reported data from children may pose challenge as self-reported autonomy may be different from actual autonomy, potentially influenced by a false sense of autonomy (Killmister 2013; Dworkin 1981). However, in this study, we contend that "sense of autonomy" and actual "autonomy" can largely align due to various mitigating factors: Firstly, in addition to self-reporting data, our observations of children's discussions offer objective evidence regarding children's development of data autonomy, particularly in terms of their enhanced critical thinking and informed decision-making. Furthermore, we took care to ensure that children were not misled into a sense of autonomy when they had little actual control, for instance, the app allows their control to directly affect recommendation outcomes. We propose that by beginning with a "sense of autonomy" perspective, and then integrating the self-reports with direct observations of children's actual thoughts and actions, we can construct a more nuanced and robust indicator of children's perception and development of data autonomy.

Overall, our findings show that children often feel their autonomy is compromised on social media platforms. While CHAITok is not a complete solution, it has notably improved children's self-reported sense of data autonomy. Children felt safer, more empowered, and more respected – *"For the very first time, what I think actually matters on these apps."* (P31, age 12). In addition to these self-reported feelings of greater autonomy, we also observed grounded improvements, aligning with ***cognitive, behavioral***, and ***emotional*** forms of autonomy outlined earlier. For example, we observed that children spontaneously engaged in critical thinking about the potential uses of their data, beyond just service enhancement, demonstrating signs of improved ***cognitive*** understanding. This shift seemed to arise from a blend of features: scrutinizing the "WHAT IS" information an examples given by CHAITok, while exercising configurations of their settings, coupled with engaging in group discussions. Meanwhile, we found that children exhibited informed decision-making, reflecting enhanced sense of ***behavioral*** autonomy. For instance, they critically considered how social media might produce biased content based on certain data types and modified their choices to align with their values, such as opposing gender-specific recommendations or not wanting to see game-related content. Lastly, grounded in cognitive and behavioral autonomy, we observed that children were not only more confident but also increasingly interested. This newfound confidence made them feel more at ease for online challenges. For example, when faced with addictive content or content concerning body image, the children felt better prepared. They understood why such content is generated and knew they could exert some control over it by managing their own data – an indicator of ***emotional*** autonomy. Interestingly, this emotional autonomy seemed to reinforce and boost children's cognitive and behavioral autonomy in return. When feeling more confident, children were observed to become more proactive in critical thinking and feel more comfortable with making decisions for themselves.

It is important to note that our findings also indicate that children's autonomy development is not merely sequential; instead, various forms are interconnected and mutually reinforcing. Particularly, this research contributes to the existing literature by underscoring the importance of children's emotional autonomy, a dimension often eclipsed in previous studies (Dasgupta and Hill 2021; Druga, Vu, et al. 2019; G. Wang, J. Zhao, et al. 2023b). Our findings highlight the need to cultivate a mindset in children about their data. Instead of merely approaching data literacy as a skill-based development (like navigating to privacy settings), it's crucial to frame it as a *socio-technical* journey – by encouraging children to grasp its wider societal repercussions, evaluate the consequences of their choices, and consequentially make well-informed, balanced decisions.

7.6.2 Design Implications

As an initial proof-of-concept focused on enhancing children's data autonomy, our system has shown promising results based on children's responses. These outcomes offer key insights for future design considerations, suggesting how support for children's data autonomy can be integrated in ways that align with their values and behaviors.

Respect children's values: Our findings indicated that children place greater value on the alignment of data-driven outcomes with their personal values than on mere just control over data. They stress that algorithmic content should extend beyond relevance, and also reflect their ethical and social values, such as promoting vaccine awareness or encouraging healthier behaviors. This fresh perspective of respecting children's values should prompt immediate considerations in design practices. In today's social media, algorithms frequently equate relentless user engagement with success, and prioritize engagement over users' well-being. Our efforts to increase transparency and offer customisable settings are a good starting point. We advise future designers to prioritize children's well-being over engagement metrics, such as implementing adaptive feedback tools that ask simple questions like, "Is this something you want to see?" or "Does this content makes you feel happy?" to consistently align with children's evolving values and preferences.

Support children's evolving autonomy: In our study, children demonstrated a need for an adaptive approach to data autonomy, aligning with their evolving literacy and agency. Prior research highlights how autonomy needs to differ by life stage: younger children seek independence from caregivers (Rolfe 2020), whereas teenagers aim for independent life choices (Bassett, Chapman, and Beagan 2008). Legislation like the UK ICO Age Appropriate Design Code (*Age appropriate design code* 2020) and the IEEE Standard for an Age Appropriate Digital Services Framework (IEEE 2021) also emphasize accounting for children's age and developmental stage in designing of digital services. Designers and innovators must recognise this complex transition from childhood to adulthood, and put deliberate considerations into children's individual characteristics, vulnerabilities and circumstances.

In particular, by focusing on 10- to 13-year-olds, a crucial age for social media engagement (Ofcom 2023) and school transition (UNICEF 2023), our study found that this age group are especially concerned about external information impact, frequently emphasizing "resilience", which appears different from the traditional notion about autonomy that emphasizes on "agency" and "control" (Steinfeld 2021; Lukoff, Lyngs, Shirokova, et al. 2023) . Our research has also shown that an improved emotional autonomy in this age group is often associated with an improved behavioural autonomy. We recommend future designs for this age group incorporate features such as displaying shared data volume for self-reflection, while also fostering resilience-building towards autonomy by offering reflection points on their choices and the content they encounter. This approach could lay a critical foundation for more advanced capabilities, like exercising control and making informed decisions.

Design for children's digital rights: In line with prior research (G. Wang, J. Zhao, et al. 2022; G. Wang, Sun, et al. 2023), our study highlights the importance of children's digital and algorithmic literacy in fostering data autonomy. However, we underscore an urgent need to reevaluate the current focus of designs, which predominantly concentrate on developing children's digital skills rather than providing comprehensive support aligned with their digital rights. We have observed that current curriculum for children aged 10 to 13 typically cover basic topics related to data collection and algorithms, but predominantly focus on the technical aspects. Existing efforts also typically focus on skill-based knowledge in a defensive manner, like mastering how to turn off personalised adverts (Group 2021; Google 2021). Our findings show that these are inadequate to help children bridge the gap between their skills and practical challenges in real-world scenarios when it comes to exercising their data autonomy. This underscores the profound need to broaden current design approaches, aligning them more closely with children's fundamental rights and well-being. The Children's Rights by Design (S. Livingstone and K. Pothong 2023) is an exemplar initiative, based on the UNCRC (Nation 2023), it offers 11 Childs Rights by Design principles for innovators and designers. The initiative prioritizes placing children at the heart of the design process, urging reflection on digital product impacts while emphasizing children's voices, consultation, and upholding their best interests and agency. We underscore this important direction of future design approaches for supporting children's digital experiences online, focusing not only on providing guidance for children on how to control their data or opt out of data practices on platforms like YouTube, but also on providing mechanisms for them to exercise and develop their rights to their data online.

7.6.3 Towards Future Data Autonomy as Rights for Children

Our findings shed light on children's views about *data autonomy*, a term not yet well-understood, offering an empirical, HCI-centric look into children's desire for greater autonomy over their data on social media platforms. Interestingly though, many of children's expectations align well with existing philosophical theories. For instance, children seeing *data autonomy as deciding for their own data* aligns with the philosophical idea of *authorship and self-congruence* (Ryan and Deci 2006;

Ricoeur 1966), asserting that individuals act in tune with their core values and needs (Ryan and Deci 2004; Sartre 2022). Children's perspective on *data autonomy as resilience over data* reflects discussions on *susceptibility to control* (Deci, Eghrari, et al. 1994), emphasizing resistance to internal and external pressures. Lastly, their idea of *data autonomy as a learned developmental competency* aligns with the philosophy of autonomy as *interest-taking* (Hmel and Pincus 2002; Loevinger et al. 2014), focusing on continuous self-reflection and evolving understanding (Ryan and Deci 2006; Deci and Ryan 2013).

Why focus on *data* autonomy then? It's widely recognized that children can face harm on social media. Yet, most discussions have centered around topics related to children's "online safety", like inappropriate content; and "self-regulation", like excessive screen time. While these issues are crucial, the notion of data takes on a unique dimension. We observed that **among the various concerns children have expressed about their lack of autonomy on social media, data always emerges as the central issue.** Even seemingly surface-level aspects like click baits, and what might appear to be content-based issues like addictiveness, are fundamentally tied to the use of data. Children view their data experience on social media as "passive", "disrespectful", "harmful", and "helpless". In extreme instances, some children even blamed themselves for datafication consequences. **Children are struggling, and their rights are being neglected.** In 2021, the UNCRC (Nation 2023) outlined children's digital rights with four principles: i) the *right to non-discrimination*, ensuring all children have meaningful access to the digital world; ii) *the best interests of the child*, prioritizing their welfare in decisions; iii) the *right to life and development*, addressing online harms and highlighting tech education; and iv) *respect for the views of the child*, stressing their involvement in policy-making and freedom of thought. Regrettably, our findings indicate that none of these four essential digital rights for children are being fulfilled in the current datafication landscape: children are overwhelmed by algorithms, facing limited digital access, having their well-being sidelined in data decisions, and finding their voices suppressed on platforms regarding their data.

These findings indicate that it is critical to **reconsider the role of data autonomy as a fundamental digital right for children**. Data is at the core of the business models of many large online platforms. Examining children's online experiences with the lens of data autonomy reveals new pathways for research, policy, and intervention. Our research has highlighted that, in addition to the existing emphasis on supporting children's self-regulation and digital literacy development, the promotion of their "data" autonomy has received less attention despite being a crucial aspect of children's digital rights. Even the latest guidance on Child Rights by Design (S. Livingstone and K. Pothong 2023) has primarily focused on children's autonomy in terms of behavior and agency, rather than specially on the development of children's autonomy over their data. Our study serves as a starting point, aiming to deepen our understanding of data autonomy and promoting a reassessment of children's digital rights. It underscores the need to recognize the complex relationship between technology, social context, and children's specific rights and needs, which is fundamental for fostering strategies that can more effectively empower children in the online realm while upholding higher ethical standards. In this way, we could envision the creation of future platforms that prioritise children's voices, values and their capacity building.

7.7 Conclusion

In this chapter, we present the design, development and evaluation of *CHAITok*, an Android mobile application designed to support children's sense of autonomy over their data on social media. Our findings indicate that children generally have a passive data experience on social media and other AI-based platforms. However, our CHAITok has effectively enhanced their sense of autonomy over their data. We provide critical design recommendations, including designing to respect children's values, support children's evolving autonomy, as well as designing for children's digital rights. We emphasize *data autonomy* as a fundamental right for children and call for further research, design innovation, and policy changes focused on this critical issue.

Statement of Authorship for joint/multi-authored papers for PGR thesis

To appear at the end of each thesis chapter submitted as an article/paper

The statement shall describe the candidate's and co-authors' independent research contributions in the thesis publications. For each publication there should exist a complete statement that is to be filled out and signed by the candidate and supervisor **(only required where there isn't already a statement of contribution within the paper itself).**

Title of Paper	CHAITok: A Proof-of-Concept System Supporting Children's Sense of Data Autonomy on Social Media.
Publication Status	Published
Publication Details	Ge Wang, Jun Zhao, Samantha-Kaye Johnston, Zhilin Zhang, Max Van Kleek, and Nigel Shadbolt. 2024. CHAITok: A Proof-of-Concept System Supporting Children's Sense of Data Autonomy on Social Media. In Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (CHI '24). Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3613904.3642294

Student Confirmation

Student Name:	Ge Wang		
Contribution to the Paper	Ge Wang is the principal author who contributed the ideation, design and development of the CHAITok, conducted all the user studies and all data analysis, and wrote the paper. Jun Zhao, Samantha-Kaye Johnston and Zhilin Zhang each attended some of the user studies and helped as note takers. Max Van Kleek, and Nigel Shadbolt critically reviewed the manuscript and offered valuable feedback.		
Signature		Date	22/09/2023

Supervisor Confirmation

By signing the Statement of Authorship, you are certifying that the candidate made a substantial contribution to the publication, and that the description described above is accurate.

Supervisor name and title: Dr Jun Zhao, Senior Research Fellow			
Supervisor comments			
Signature		Date	19 March 2024

This completed form should be included in the thesis, at the end of the relevant chapter.

Publications arising from this chapter:

- *Ge Wang, Jun Zhao, Max Van Kleek, and Nigel Shadbolt. 2024. Challenges and Opportunities in Translating Ethical AI Principles into Practice for Children. Nature Machine Intelligence, pp.1-6.* `https://doi.org/10.1038/s42256-024-00805-x`

8

Discussion and Conclusions

Contents

Building on our findings from previous chapters, in this chapter, we now seek to answer the final thesis question: *How can we define data autonomy for children in* *the age of AI, and what open challenges remain to be addressed?* **(TQ3)**.

We first review the goals and findings of this thesis, before moving on to discuss the concept of *data autonomy*, followed by a child-centred perspective on data autonomy and its associated considerations and challenges. Finally, we discuss both potential future directions and the wider implications of our work.

8.1 Overview of Results

At the beginning of this thesis, we highlighted three core challenges underpinning our primary research goal: In the age of AI, **how to take a child-centred approach to support children's autonomy and development within the prevalent datafication landscape**.

Firstly, the current comprehension of the AI landscape as it relates to children is insufficient, especially when considering the intricate datafication practices involving them. Secondly, there's a discernible gap in understanding children's viewpoints, requirements, and the challenges they encountered in preserving their autonomy over data on these platforms. Lastly, the significance of data autonomy for children has been somewhat overlooked in much of the existing research, and that most existing research leans towards a conventional parent-led and protective stance.

Throughout this thesis, we tried to address these challenges through a structured four-step methodology: 1) Landscape analysis to comprehend the present scope of AI-based platforms for children and the prevalent challenges they encounter; 2) Conceptual review to elucidate the meaning of autonomy for children in the digital realm; 3) Empirical investigation focusing on children's perceptions, needs, and obstacles concerning data autonomy; and 4) Technical evaluation to assess the impact of technical interventions on children's sense of data autonomy.

8.1.1 Landscape Analysis of AI-Based Platforms for Children and Their Prevalent Challenges

In Chapter 2, we explored the current landscape of AI-based platforms designed for children and the digital tools supporting children navigating the various datafication practices inherent in these platforms. Through a systematic review of 188 AI-based platforms and 58 digital tools crafted for children's online safety, we shed light on the prevailing research gaps and highlighted potential design opportunities.

First, we identified **a critical necessity to integrate a child-centred perspective into current AI development for children.** In today's fast-paced AI landscape, many systems seem to have a tunnel vision, often prioritizing a select few principles. Yet, crucial aspects like *children's wellbeing* are frequently sidelined. This underscores the need for a broader child-centred perspective. It's essential not just to identify often-missed risks, but also to understand how AI uniquely affects children compared to adults. By emphasizing core values such as safety, fairness, inclusion, long-term impact, and the best interests of children, we can ensure a more holistic and responsible integration of AI into their lives.

Second, we recognised **an urgent call to move beyond traditional parent-guided protections, focusing instead on empowering children and nurturing their sense of autonomy.** In Chapter 2, we highlighted the complex datafication practices in AI-based platforms for children, which pose risks potentially hinder children's critical thinking and independent decision-making capabilities. Meanwhile, most existing digital tools for children overlook such datafication intricacies, focus narrowly on online safety, often with a parent-centric protective stance that neither children nor parents prefers. Our findings emphasised the value of designs that encourage open communication and transparency between parents and children. Thus, there's an urgent need to shift from merely parent-guided protections to solutions that both safeguard and empower children, truly nurturing their digital autonomy.

8.1.2 Conceptual Review of Children's Autonomy in the Digital Realm

Building on the advocacy for children's digital autonomy, Chapter 3 delves into the prevailing notions of autonomy for children in the digital space. Through a systematic review of 68 research papers centered on children's digital autonomy, we formulated a comprehensive taxonomy.

We discerned three primary themes: digital autonomy as the ability to develop intrinsic motivation and self-regulation; as the ability to make critical thinking and informed decisions; and as computational thinking and literacy development. We showed that in contrast to the general expectation that digital autonomy is largely about self-governance or behaviour change (Rafael A. Calvo et al. 2014; Peters, Rafael A Calvo, and Ryan 2018), the ability of computational thinking and critically acting on these information is just as important. By juxtaposing HCI interpretations with those from other disciplines, we underscored the need to approach digital autonomy as a multi-faceted construct. Recognizing its multi-dimensional nature is not only pivotal to children's developmental journey but also instrumental in shaping the design and creation of tools that empower their digital autonomy in the future.

Meanwhile, as we delved deeper into our review, **it's became increasingly clear that data plays a central role in all aspects of autonomy.** Children's online footprints aren't merely remnants of their activities; they actively shape and define their digital experiences. To nurture genuine digital autonomy in children, a foundational step is ensuring they have control over their own data. This revelation underscores the importance of equipping children with the skills and knowledge to navigate the complex landscape of datafication within AI-based platforms that characterizes their online interactions. Consequently, there's an emergent need to pivot our focus towards enhancing data autonomy specifically for children.

8.1.3 Empirical Investigation of Children's Perceptions, Needs and Obstacles Concerning Data Autonomy

In Chapters 4 and 5, we delved into children's perspectives on datafication, aiming to understand their perceptions, needs, and challenges when striving for autonomy over their data. Through interviews with 48 children and co-design workshops involving 53 children, we provided critical insights into how children currently perceive datafication and the support they require to navigate it effectively.

Our findings showed that not only do children care significantly about various aspects of datafication, but they already possess rudimentary conceptual understandings of it, and a significant willingness to take action to shape it to their desires. We found that children were well aware of the data collection practices on AI-based platforms. They largely knew that these collected data will be processed, and generally understood that such processing is not conducted by human, but through some kind of machine. They were somewhat aware that online datafication practices could "make guesses about them", and that the outcomes of these guesses might impact users like themselves, even though they didn't fully grasp the entire scope of these practices.

Building on this foundational knowledge, our co-design activities further underscored the necessity of fostering autonomy-supportive designs tailored for children. We identified possible autonomy-supportive design avenues, such as rethinking what 'transparency' means for children of different age groups; repositioning children as active participants than passive consumers online; and always approaching design with a lens of genuine care and respect for the young users.

8.1.4 Technical Evaluation on the Impact of Technical Interventions on Children's Sense of Data Autonomy

In Chapter 6 and 7, we designed, developed and assessed two technical prototypes as interventions to support children's sense of data autonomy. Through studies involving 27 families and 109 children, we discovered that these prototypes effectively enhanced children's critical thinking, informed decision-making, and amplified their sense of autonomy over their data.

Specifically, with our KOALA Hero Toolkit, we noted heightened awareness of datafication risks among families, an evolution in their thought processes, and a collaborative family decision-making approach. This emphasizes the potential of a child-centred approach which promotes a democratic interaction between children and parents as equal team members; which stands out against the traditional parent-led methods, especially in its ability to support a co-learning process between parents and children, therefore cultivating a heightened critical awareness of datafication.

Meanwhile, with our PodTok mobile app, children reported an enhanced sense of data autonomy, feeling safer and more respected. Beyond these feelings, we also observed grounded improvements: children actively engaged in critical thinking about their data's potential uses. They made informed decisions by critically analysing how social media might generate biased content based on certain data. Anchored in these improvements, children's confidence grew, preparing them for online challenges like dealing with addictive or body image related content.

These validations underscore the profound significance of adopting a child-centred approach in addressing data autonomy. By placing children at the heart of the solution, we not only elevate their digital experiences but also empower them to be informed and resilient navigators of their online journeys.

8.2 Towards Data Autonomy

Throughout the course of this thesis, we have conducted a comprehensive examination of autonomy, tracing its roots from philosophical foundations to its embodiment in practical scenarios such as HCI studies, and culminating in its specific contextualisation within the realm of data – the cornerstone of today's AI era. In this section, we seek to distil these insights into a refined conceptual framework on *Data Autonomy* in the age of AI.

Before delving into the framework, it's pivotal to revisit our path towards the concept of *data autonomy* throughout this thesis (see Figure 8.1). Our exploration commenced with a broad perspective rooted in the foundational principles of the UNCRC's *human rights* for children (*General comment No. 25 (2021) on children's rights in relation to the digital environment* 2020). These rights, including Non-discrimination (Article 2), the Best Interest of the Child (Article 3), the Right to Life, Survival, and Development (Article 6), and the Right to be Heard (Article 12), serve as universal tenets for child welfare.

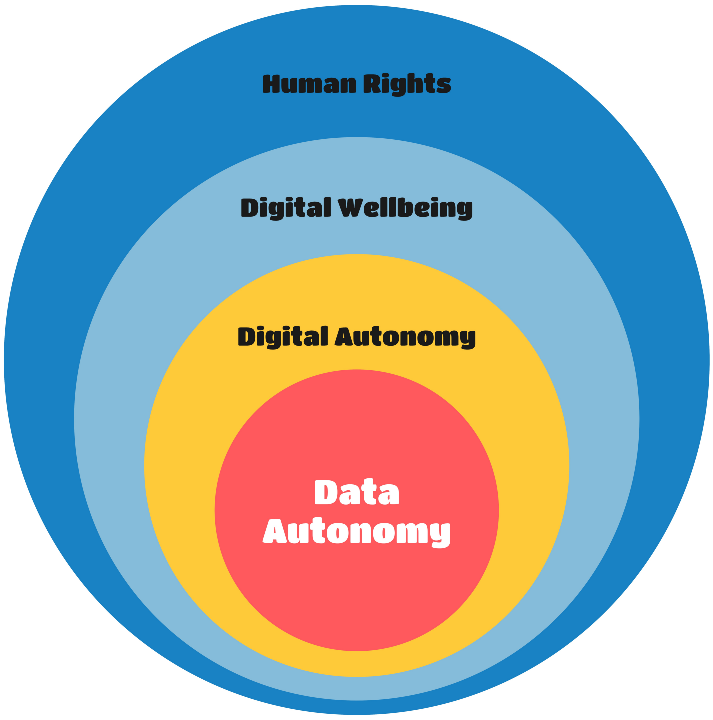


Figure 8.1: The role of *Data Autonomy*.

When these principles are adapted to the digital domain, they become beacons for fostering children's *digital wellbeing* (*General comment No. 25 (2021) on children's rights in relation to the digital environment* 2020). For instance, Digital Non-discrimination underscores the importance of equal access to digital resources for all children, regardless of their circumstances. The digital avatar of the Best Interest principle underscores the commitment to prioritize children's digital safety and welfare. Translating the Right to Life, Survival, and Development to the online realm means offering platforms that support a child's digital growth, knowledge, and safety. And the Right to be Heard in this context translates to empowering children with a voice online, ensuring their perspectives shape their digital journeys.

As this thesis progressed, the significance of a child-centred lens in the digital world became increasingly evident, culminating in the introduction of the concept of *digital autonomy* in Chapter 3. This concept captures a user's ability to have computational literacy, critically navigate, make well-informed decisions, and effectively manage both their actions and emotions in the digital environment.

In investigating the conceptualisations around digital autonomy, it became increasingly clear that data is at the heart of digital autonomy, consistently driving the profound influence of digital services. This led to an intensified focus on *data autonomy*. Rather than being a mere subset, this thesis identifies *data autonomy* as the nucleus of the entire concept of *digital autonomy*. It's the driving force, the core around which all other facets of digital autonomy revolve, thereby underpinning and shaping the entire discourse on children's digital wellbeing.

Now, drawing from our landscape analysis, conceptual review, empirical investigation, and technical evaluation, we are poised to articulate a definitive framework for *data autonomy*, anchored in three fundamental pillars:

8.2.1 Theme 1: Data Ownership

The first central theme of data autonomy is the concept of *data ownership*. It emphasizes users' inherent right to exercise full authority over their online personal data. This principle is not just foundational; it's the linchpin that upholds data autonomy, ensuring users can effectively control their data narratives.

Currently, the lack of genuine personal data ownership is glaringly evident on many online platforms. Users, instead of being data custodians, often find their digital footprints being captured in minute detail and analyzed using advanced techniques, further exploited by major platforms. Such a system doesn't merely hinder users from accessing their data; it denies them their rightful control, transforming their data into mere commodities traded without their consent.

Our research has highlighted a troubling trend among children. They increasingly feel detached from the idea of personal data ownership. To some of them, the notion that they might "own" their online data feels foreign. This detachment originates from feelings of powerlessness against the colossal data-driven entities shaping their digital experiences. The prevailing online architecture has taught them that their data is uncontrollable, a possession they can't genuinely own or influence.

It's evident that there's a pressing need to reshape this dynamic, ensuring all users, especially the young, to have genuine ownership of their own data. The data-centric strategies of many platforms, while challenging, also present an opportunity to redesign the future of data governance. It's time to evolve towards a data ecosystem rooted in ethics and user empowerment. Encouragingly, emerging trends already hint at this transformative shift (*Ethical Web And Data Architectures* 2021). Efforts are underway to craft innovative governance models, with a growing emphasis on decentralized data ownership and data sharing strategies – heralding a future where data autonomy is a reality for all.

8.2.2 Theme 2: Data Resilience

The secondary pillar in our exploration of data autonomy is encapsulated by the term *data resilience*. Our findings underscore that the current challenges faced by families go beyond the mere erosion of data ownership, but the consequences of this ownership loss can have deep-seated effects. Beyond merely possessing rights over their personal data; users, particularly children and their families, must be fortified against the potential harms arising from data misuse, which can range from attention diversion to behavioral manipulation.

While current educational frameworks attempt to instill resilience, the emphasis remains largely on immediate online dangers such as inappropriate content, stranger interactions, and excessive screen time. While these issues are undeniably crucial, they can overshadow more subtle, long-term dangers, ignoring the effects of data exploitation on children's holistic development and welfare.

To genuinely cultivate data resilience, we must arm children and their families with an enriched understanding and heightened awareness that will enable them to conscientiously steer their online journey, thereby asserting agency over their digital narratives. As we look forward, expanding the definition of "resilience" in our educational approaches becomes paramount. We need a more encompassing curriculum that thoroughly educates children about the intricate facets of data misuse and its potential lasting impacts. The ultimate aim is to shape a generation proficient and discerning in navigating the online landscape, endowed with the requisite knowledge and critical acumen to protect their digital wellbeing. Our research advocates for design strategies that support children's contextual reflections on data decisions, integrate their values into the design paradigm, and champion a forward-thinking exercise of autonomy. By adopting this proactive stance on data resilience, we can counteract the more covert manipulative strategies rooted in data misuse, paving the way for a fortified and empowered digital user experience.

8.2.3 Theme 3: Data Sovereignty

The third theme, *data sovereignty*, holds a critical position in today's digital domain, especially for children. Rooted in the ideals of rights and activism, it echoes the powerful sentiment: 'my data is my right' (*Age appropriate design code* 2020). Standing distinct from other themes, it moves beyond merely imparting skills. It endeavors to instill in children a deep-seated realization that they innately hold rights over their data. While much of the existing research has commendably equipped children with practical skills for online privacy, like password management and ad personalisation settings, a gap remains in nurturing foundational awareness about data rights. Such awareness forms the very essence upon which all other skills are constructed, motivating children to proactively curate their digital identities.

One of the key challenges in cultivating this sense of data sovereignty among children is its glaring absence from modern educational curriculum. Though the UNCRC General Comment 25 recognizes children's digital rights as fundamental, this acknowledgment largely remains within scholarly confines, barely trickling down to the educational content and tools on the market for children. This discrepancy underscores a critical need: to bridge the gap between academic discourse and real-world applications. This calls for a holistic revision of educational strategies, introducing modules that enlighten and empower children regarding their data rights. Moreover, our findings highlight crucial design opportunities to better acquaint children with their data rights. While it's pivotal to offer skill-based knowledge, such as managing privacy settings, it's equally crucial to embed these skills within a more contextualised narrative. Children should be encouraged to view data rights not as mere technicalities but as integral components of their overall well-being. By adopting a socio-technical methodology, the aim is to inculcate a deep appreciation for data sovereignty in children, guiding them towards a future where they traverse the digital realm with both proficiency and prudence.

8.3 A Child-Centred Perspective on Data Autonomy: Unique Considerations and Open Challenges

Given the proposed conceptual framework on data autonomy, what unique considerations and challenges arise when viewed through a child-centred perspective?

8.3.1 Considerations on the Literacy Development of Children

Although our primary intention was not to craft an educational piece, throughout this thesis, we've uncovered the crucial role of children's digital literacy in understanding, perceiving, and navigating datafication practices. Beyond our central exploration of data autonomy for children, our findings offer significant implications for the development of children's algorithmic literacy. Notably, our research emphasizes that such literacy transcends mere *cognitive* comprehension, and would be better supported when coupled with *situated* thought processes, enhancing children's grasp of the social intricacies of algorithmic systems. Furthermore, our findings underscore the important role of *critical* thinking in algorithmic literacy, which serves as a foundation for children to discern the relationship between their data and digital experiences, fueling activism and heightening their awareness of prevailing data power structures.

On the other hand, in the course of our research, we discovered that the ability to contextualize datafication within the wider digital society is seldom addressed in existing studies on algorithmic literacy (Aleman et al. 2021; Pangrazio and Cardozo-Gaibisso 2021; Long and Magerko 2020). At the same time, we noticed that while many current curricula for children touch upon basic topics related to data collection and algorithms, they often lean heavily towards technical details, emphasizing reactive, skill-based solutions (e.g., turning off cookies) rather than proactively empowering children to gain autonomy over their own data.

How to advocate for such literacy development in children, especially in educational settings, remains an open challenge and a critical point of investigation. We urge future researchers to delve deeper into curriculum design, pedagogical strategies, and cross-disciplinary collaborations that prioritize children's holistic understanding of datafication and its implications in their digital lives, laying the groundwork for their cultivation of data autonomy.

8.3.2 Considerations on the Role of Guardians in Childhood

While this thesis champions the development of autonomy for children, we must recognize the intrinsic distinctions between children and adults in terms of moral significance. A defining characteristic distinguishing children from other users is the involvement of parents or guardians, who carry significant legal and ethical decision-making roles for their children. It's therefore vital to acknowledge the clear moral boundary between 'competent children' and 'adults.' Irrespective of their intellectual capabilities, children do not possess the same level of autonomy and responsibilities as adults. Parents, both ethically and legally, shoulder these responsibilities. Thus, the roles of parents in this scenario warrant examination.

In our research, we underscored the unique influence parents and guardians had in the evolution of children's data autonomy. While a purely parent-led restrictive approach is not our recommendation, the indispensable role of parents in this landscape cannot be overlooked. Our findings suggest that a child's data autonomy can not be developed in silo but thrives more robustly when there's meaningful and timely parental/family involvement. Yet, a challenge we pinpointed early on in our thesis is the knowledge gap: often, parents may lack expertise in the AI-driven platforms their children frequent, while children sometimes emerge as the more knowledgeable party. This presents an opportunity for *co-learning* between parents and children, which can be both educative and empowering for both parties. Consequently, we advocate for future researchers to emphasize initiatives that promote age-appropriateness, open-ended play, and a balance of power between children and adults. Such initiatives should aim to nurture playful, investigative, and collaborative experiences within family settings.

8.3.3 Considerations on the Developmental Aspect of Childhood

As previously highlighted, many existing ethical AI frameworks either neglect the unique role of children or simply treat them as a monolithic group. The inclusion of children in such frameworks often feels more like a nod to compliance than a genuine effort to address their specific needs and perspectives. In many such frameworks, one could interchange 'children' with 'socially vulnerable groups' without altering the essence or context. To genuinely understand the ethical intricacies tied to involving children in research, it's vital to unpack the multifaceted concept of 'childhood'. One of its defining characteristics is the developmental trajectory – a journey from the profound dependence of infancy to the burgeoning independence of adolescence.

Although this thesis zeroes in on the specific age bracket of 7 to 14, our findings reveal that children's perceptions, needs, and challenges related to data autonomy don't neatly correlate with their age. Instead, relying solely on broad age categorizations falls short, given the significant variations in children's intellectual capacities, developmental pace, maturity, and personal experiences.

This prompts an interesting open discussion: Is data autonomy a dynamic concept? Intriguingly, our findings revealed that our child participants have already perceived data autonomy as *developmental competencies to be learnt*. This perspective stands in contrast to many philosophical definitions of autonomy, which tend to frame autonomy as a relatively static concept, emphasizing a predefined set of abilities. However, if children view autonomy as something developmental and evolving, this suggests a need to rethink and possibly expand our understanding. This raises questions about whether our traditional conceptions of autonomy are too rigid, especially in the context of the rapidly changing digital landscape. It invites scholars and practitioners to consider a more fluid and adaptable framework for data autonomy, one that evolves with individual experiences and developmental stages.

8.4 Directions for Future Work

This section summarises and expands on opportunities for future work that have been highlighted throughout the thesis.

8.4.1 Data Autonomy as a New Research Agenda

This thesis underscores the critical importance of children's data autonomy in determining their digital well-being. While traditional discourse in the digital realm has often focused on children's vulnerability and the passive impacts on them, it has largely overlooked the active empowerment and enhancement of their autonomy. Data, in our exploration, stands out as the central force driving this autonomy. Given its deep ties to the business models of online platforms, it's crucial to elevate data autonomy from a mere feature to a foundational digital right for children.

Venturing into the domain of children's online experiences with data autonomy as the guiding principle unveils novel avenues for academic inquiry, policy formulation, and targeted interventions. This not only paves the way for a more nuanced understanding of data autonomy as a dynamic and evolving field but also underscores the need for a recalibration of our perspectives on children's digital rights. The intrinsic relationship between technology, societal context, and children's unique rights and requirements demands attention, leading to more ethical and effective strategies for online empowerment.

Recognizing the untapped potential of children's data autonomy also paves the way for multidisciplinary collaborations. For instance, integrating the child-computer interaction community with the algorithmic and other related fields. Meanwhile, these collaborative endeavors should not be confined to academia alone: For academia, it calls for a paradigm shift from protection to empowerment; for developers, it emphasises an ethical reevaluation of design practices; and for legislators, it signifies a reconceptualization of children's rights in the age of data. Our work is but the first step in a long journey of ensuring children are not merely passive digital consumers but empowered and autonomous digital citizens.

8.4.2 Future Architectures for Ethical Data Governance

The body of evidence collated in this thesis underscores a pressing need to reevaluate our prevailing data governance structures, which are predominantly centralized and dominated by a handful of tech giants. This centralized paradigm not only undermines the core principle of data ownership central to our conceptualization of data autonomy, but also poses substantial challenges to user resilience and sovereignty in the digital realm.

As we navigate this transition towards more ethical data governance architectures, enhancing public awareness and education is crucial. Only with an informed public can we shape a governance model that genuinely honors user rights, enabling true autonomy in the digital world. At the moment, there have been several recent movements towards this, and researchers have been actively engaged in the development of new data governance structure. A wave of new decentralised paradigms for data sharing and ownership (*Ethical Web And Data Architectures* 2021) are being explored to expand individual data subjects' ability to control access to their data, by enabling collective access requests through representative intermediaries such as NGOs and trade unions, therefore increasing the agency of individual data subjects. However, while decentralization appears promising, it also opens up new challenges and questions: Who would oversee data in a decentralized environment? How can we guarantee accountability, particularly during disputes? Which safeguards would ensure data integrity and security? Specifically for children, who would be accountable for their data practices, and what protections would be available?

Looking forward, we champion further investigations and more research into these new alternative data governance models, with the aim of recalibrating the balance of power in the data domain, leading us into a fairer and more ethical age of AI.

8.5 Reflections and Limitations

The thesis presents several limitations worth unpacking, beginning with the selection of AI policies and reliance on literature reviews in earlier chapters. This approach, while foundational, restricts the exploration to a subset of frameworks and research-centric AI systems, potentially overlooking the broader spectrum of commercial practices. To mitigate this, a deliberate selection of comprehensive literature was made, aiming to anchor the analysis in some most relevant and impactful discussions.

As the focus shifts more prominently from the Chapter 4 onwards to children's digital interactions and perceptions of datafication, the thesis encounters further limitations. The geographical and demographic scope, mainly concentrated around urban and university-adjacent areas, might not capture the full diversity of children's experiences. Despite this, the research was designed to reflect a spectrum of digital literacy and awareness levels, though within the logistical confines of the study's scope.

The use of self-report data from children introduces the potential for response biases. A nuanced approach to question design and participant engagement was employed, creating an atmosphere where children could express their experiences honestly, without the constraint of providing "correct" answers. Additionally, the focus on specific platforms like YouTube and TikTok, while offering in-depth exploration of certain digital experiences, also potentially limit the findings' applicability across the digital landscape. However, these platform-specific analysis served as valuable case studies, providing insights into digital phenomena relevant to children but acknowledging the concentrated scope of this examination.

8.6 Concluding Remarks

The age of AI is a rapidly evolving and complex space for children, their deepening engagement with digital platforms subjects them to intensified datafication, wherein their personal details are not only gathered but often leveraged for both behavioral influence and commercial gains. This concealed manipulation by online services can subtly alter children's worldviews, challenging their autonomy. Although historically parents and guardians were considered the primary navigators for children in the digital realm, the rapid advancements in AI often outpace their understanding. Recognizing this disparity, there's an emergent shift towards a child-centred approach, which shifts from just protecting or limiting children with parents in charge, to actively guiding and empowering children to take a leading role in their digital environments.

In the evolving age of AI, this thesis delves into a pivotal question: How to take a child-centred approach to support children's autonomy and development within the prevalent datafication landscape. First, it offers a landscape review, shedding light on the design of AI-based platforms tailored for children and pinpointing the challenges tied to their datafication practices. Second, it provides a nuanced conceptualisation of autonomy for children in the digital space. Third, it provides empirical insights into children's perceptions, needs and obstacles regarding having autonomy with their data. Fourth, it contributes two technical prototypes as design interventions for supporting children's autonomy over their data. Finally, it makes important theoretical contributions to the concept of *data autonomy*, suggesting a fresh perspective on children's online experience and wellbeing. This work not only lays the foundation for future research on data autonomy as a novel research agenda, but also prompts a rethinking of existing data governance structures towards a more ethical data landscape.

Statement of Authorship for joint/multi-authored papers for PGR thesis

To appear at the end of each thesis chapter submitted as an article/paper

The statement shall describe the candidate's and co-authors' independent research contributions in the thesis publications. For each publication there should exist a complete statement that is to be filled out and signed by the candidate and supervisor **(only required where there isn't already a statement of contribution within the paper itself).**

Title of Paper	Challenges and Opportunities in Translating Ethical AI Principles into Practice for Children.
Publication Status	Published
Publication Details	Ge Wang, Jun Zhao, Max Van Kleek, and Nigel Shadbolt. 2024. Challenges and Opportunities in Translating Ethical AI Principles into Practice for Children. Nature Machine Intelligence, pp.1-6. https://doi.org/10.1038/s42256-024-00805-x

Student Confirmation

Student Name:	Ge Wang		
Contribution to the Paper	Ge Wang is the principal author who contributed the ideation, data collection, data analysis, and paper write-up. Jun Zhao, Max Van Kleek, and Nigel Shadbolt critically reviewed the manuscript and offered valuable feedback.		
Signature		Date	22/09/2023

Supervisor Confirmation

By signing the Statement of Authorship, you are certifying that the candidate made a substantial contribution to the publication, and that the description described above is accurate.

Supervisor name and title: Dr Jun Zhao, Senior Research Fellow			
Supervisor comments			
Signature		Date	19 March 2024

This completed form should be included in the thesis, at the end of the relevant chapter.

References

7 Heuristics For Humane Design (2020). URL: https://designlab.com/blog/heuristics-humane-design-designing-for-dignity/.

Abe, Kasumi et al. (2012). "Playmate robots that can act according to a child's mental state". In: *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 4660–4667.

Ackermann, Edith et al. (1996). *Constructionism in practice: Designing, thinking, and learning in a digital world*. Routledge.

Acquisti, A., L. Brandimarte, and G. Loewenstein (Jan. 2015). "Privacy and Human Behavior in the Age of Information". In: *Science* 347.6221, pp. 509–514. (Visited on 08/27/2020).

Adikara, Putra Pandu, Sigit Adinugroho, and Salsabila Insani (2020). "Detection of Cyber Harassment (Cyberbullying) on Instagram Using Naı¨ve Bayes Classifier with Bag of Words and Lexicon Based Features". In: *Proceedings of the 5th International Conference on Sustainable Information Engineering and Technology*. SIET '20. Malang, Indonesia: Association for Computing Machinery, pp. 64–68. URL: https://doi.org/10.1145/3427423.3427436.

Age appropriate design code (2020). URL: https://ico.org.uk/media/for-organisations/guide-to-data-protection/key-data-protection-themes/age-appropriate-design-a-code-of-practice-for-online-services-2-1.pdf.

Agency By Design (2020). URL: https://pz.harvard.edu/projects/agency-by-design.

AI and the future of media (2021). URL: https://www.ibm.com/thought-leadership/institute-business-value/report/cognitivemedia.

AI in the UK: No Room for Complacency (2020). URL: https://publications.parliament.uk/pa/ld5801/ldselect/ldliaison/196/196.pdf.

Akter, Mamtaj et al. (2022). "From Parental Control to Joint Family Oversight: Can Parents and Teens Manage Mobile Online Safety and Privacy as Equals?" In: *Proceedings of the ACM on Human-Computer Interaction* 6.CSCW1, pp. 1–28.

Aleman, Ezequiel et al. (2021). "Screening Bot: a Playground for Critical Algorithmic Literacy Engagement with Youth". In: *Extended Abstracts of the 2021 Annual Symposium on Computer-Human Interaction in Play*, pp. 198–202.

Alemany, Jose, Elena Del Val, and Ana Garcia-Fornes (2020). "Assessing the Effectiveness of a Gamified Social Network for Applying Privacy Concepts: An Empirical Study with Teens". In: *IEEE Transactions on Learning Technologies* 13.4, pp. 777–789.

Alemany, José et al. (2019). "Enhancing the privacy risk awareness of teenagers in online social networks through soft-paternalism mechanisms". In: *International Journal of Human-Computer Studies* 129, pp. 27–40.

Alemany Bordera, Jose, Elena Del Val Noguera, and Ana García-Fornes (Jan. 2020). "Empowering Users Regarding the Sensitivity of their Data in Social Networks through Nudge Mechanisms". In.

Alnajjar, Fady S. et al. (2019). “A Low-Cost Autonomous Attention Assessment System for Robot Intervention with Autistic Children”. In: *2019 IEEE Global Engineering Education Conference (EDUCON)*, pp. 787–792.

Amazon Privacy Notice (2023). URL: `https://www.amazon.com/gp/help/customer/`.

Andreswari, Rachmadita, Irfan Darmawan, and Warih Puspitasari (2018). “A Preliminary Study on Detection System for Assessing Children and Foster Parents Suitability”. In: *2018 6th International Conference on Information and Communication Technology (ICoICT)*, pp. 376–379.

Annex B: Age and developmental stages (2020). URL: `https://ico.org.uk/for-organisations/guide-to-data-protection/ico-codes-of-practice/age-appropriate-design-a-code-of-practice-for-online-services/annex-b-age-and-developmental-stages/`.

Ansari, Ahlam Shakeel et al. (2018). “Real-Time Child Abuse and Reporting System”. In: *2018 Fourth International Conference on Advances in Electrical, Electronics, Information, Communication and Bio-Informatics (AEEICB)*. IEEE, pp. 1–4.

Arrúa, Alejandra et al. (2017). “Influence of label design on children’s perception of 2 snack foods”. In: *Journal of Nutrition Education and Behavior* 49.3, pp. 211–217.

Arvanitis, Alexios, Alexia Barrable, and Anna K Touloumakos (2023). “The relationship between autonomy support and structure in early childhood nature-based settings: Practices and challenges”. In: *Learning Environments Research*, pp. 1–19.

Ashktorab, Zahra and Jessica Vitak (2016). “Designing Cyberbullying Mitigation and Prevention Solutions through Participatory Design With Teenagers”. In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. CHI ’16. San Jose, California, USA: Association for Computing Machinery, pp. 3895–3905. URL: `https://doi.org/10.1145/2858036.2858548`.

Auerbach, Carl and Louise B Silverstein (2003). *Qualitative data: An introduction to coding and analysis*. Vol. 21. NYU press.

Ayyappan, S. and S. Matilda (2020). “Criminals And Missing Children Identification Using Face Recognition And Web Scrapping”. In: *2020 International Conference on System, Computation, Automation and Networking (ICSCAN)*, pp. 1–5.

Azizi, Shahla, Farzad Towhidkhah, and Farshad Almasganj (2012). “Study of VTLN method to recognize common speech disorders in speech therapy of Persian children”. In: *2012 19th Iranian Conference of Biomedical Engineering (ICBME)*, pp. 246–249.

Baciu-Ureche, Ovidiu-Gabriel et al. (2019). “The Adventures of ScriptKitty: Using the Raspberry Pi to Teach Adolescents about Internet Safety”. In: *Proceedings of the 20th Annual SIG Conference on Information Technology Education*. SIGITE ’19. Tacoma, WA, USA: Association for Computing Machinery, pp. 118–123. URL: `https://doi.org/10.1145/3349266.3351399`.

Badillo-Urquiola, Karla et al. (2019). “Stranger Danger! Social Media App Features Co-Designed with Children to Keep Them Safe Online”. In: *Proceedings of the 18th ACM International Conference on Interaction Design and Children*. IDC ’19. Boise, ID, USA: Association for Computing Machinery, pp. 394–406. URL: `https://doi.org/10.1145/3311927.3323133`.

Bairavi, Krishna and K. B. Kumar Sundhara (2018). “EEG Based Emotion Recognition System for Special Children”. In: *Proceedings of the 2018 International Conference on Communication Engineering and Technology*. ICCET ’18. Singapore, Singapore: Association for Computing Machinery, pp. 1–4. URL: `https://doi.org/10.1145/3194244.3194245`.

Baldovino, Aika Patricia, Frances Neele Vergonio, and John Paul Tomas (2019). “Child Attention Detection through Facial Expression Recognition Using SVM Algorithm”. In: *Proceedings of the 2019 International Conference on Information Technology and Computer Communications*. ITCC 2019. Singapore, Singapore: Association for Computing Machinery, pp. 52–58. URL: https://doi.org/10.1145/3355402.3355411.

Ballagas, Rafael et al. (2013). “Electric agents: fostering sibling joint media engagement through interactive television and augmented reality”. In: *Proceedings of the 2013 conference on Computer supported cooperative work*, pp. 225–236.

Bandura, Albert (2005). “The evolution of social cognitive theory”. In: *Great minds in management*, pp. 9–35.

Barnes, Douglas (2008). “Exploratory talk for learning”. In: *Exploring talk in school*, pp. 1–15.

Bassett, Raewyn, Gwen E Chapman, and Brenda L Beagan (2008). “Autonomy and control: the co-construction of adolescent food choice”. In: *Appetite* 50.2-3, pp. 325–332.

Beckert, Troy E (2007). “Cognitive Autonomy and Self-Evaluation in Adolescence: A Conceptual Investigation and Instrument Development.” In: *North American Journal of Psychology* 9.3.

Bekker, Tilde, Saskia Bakker, et al. (2015). “Teaching children digital literacy through design-based learning with digital toolkits in schools”. In: *International Journal of Child-Computer Interaction* 5, pp. 29–38.

Bekker, Tilde, Linda De Valk, and Berry Eggen (2014). “A toolkit for designing playful interactions: The four lenses of play”. In: *Journal of Ambient Intelligence and Smart Environments* 6.3, pp. 263–276.

Benson, Phil (2010). “Measuring autonomy: Should we put our ability to the test”. In: *Testing the untestable in language education*, pp. 77–97.

Beyers, WIM and LUC Goossens (1999). “Emotional autonomy, psychosocial adjustment and parenting: Interactions, moderating and mediating effects”. In: *Journal of adolescence* 22.6, pp. 753–769.

Bhargava, Vikram R and Manuel Velasquez (2021). “Ethics of the attention economy: The problem of social media addiction”. In: *Business Ethics Quarterly* 31.3, pp. 321–359.

Bilal, Dania (2013). “Comparing google’s readability of search results to the flesch readability formulae: A preliminary analysis on children’s search queries”. In: *Proceedings of the American Society for Information Science and Technology* 50.1, pp. 1–9.

Billow, Richard M (1981). “Observing spontaneous metaphor in children”. In: *Journal of Experimental Child Psychology* 31.3, pp. 430–445.

Binns, Reuben et al. (2018). “Third party tracking in the mobile ecosystem”. In: *Proceedings of the 10th ACM Conference on Web Science*, pp. 23–31.

Bird, Steven, Ewan Klein, and Edward Loper (2009). *Natural language processing with Python: analyzing text with the natural language toolkit.* " O’Reilly Media, Inc."

Bokhorst, Marcel (2021). *NetGuard*. https://github.com/M66B/NetGuard.

Boteanu, Adrian and Sonia Chernova (2013). “Modeling Discussion Topics in Interactions with a Tablet Reading Primer”. In: *Proceedings of the 2013 International Conference on Intelligent User Interfaces*. IUI ’13. Santa Monica, California, USA: Association

for Computing Machinery, pp. 75–84. URL:
https://doi.org/10.1145/2449396.2449409.
Bowyer, Alex et al. (2018). "Understanding the Family Perspective on the Storage, Sharing and Handling of Family Civic Data". In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. CHI '18. Montreal QC, Canada: Association for Computing Machinery, pp. 1–13. URL:
https://doi.org/10.1145/3173574.3173710.
Braun, Virginia and Victoria Clarke (2006). "Using thematic analysis in psychology". In: *Qualitative research in psychology* 3.2, pp. 77–101.
Brazauskas, Justas et al. (2021). "DataMoves: Entangling Data and Movement to Support Computer Science Education". In: *Designing Interactive Systems Conference 2021*. DIS '21. Virtual Event, USA: Association for Computing Machinery, pp. 2068–2082. URL: https://doi.org/10.1145/3461778.3462039.
Brito, Rita et al. (2017). "Family dynamics in digital homes: The role played by parental mediation in young children's digital practices around 14 European countries". In: *Contemporary Family Therapy* 39.4, pp. 271–280.
Büchi, Moritz, Eduard Fosch-Villaronga, Christoph Lutz, Aurelia Tamò-Larrieux, and Shruthi Velidi (2021). "Making sense of algorithmic profiling: user perceptions on Facebook". In: *Information, Communication & Society*, pp. 1–17.
Büchi, Moritz, Eduard Fosch-Villaronga, Christoph Lutz, Aurelia Tamò-Larrieux, Shruthi Velidi, and Salome Viljoen (2020). "The chilling effects of algorithmic profiling: Mapping the issues". In: *Computer law & security review* 36, p. 105367.
Cabrera, Lautaro, John H. Maloney, and David Weintrop (2019). "Programs in the Palm of Your Hand: How Live Programming Shapes Children's Interactions with Physical Computing Devices". In: *Proceedings of the 18th ACM International Conference on Interaction Design and Children*. IDC '19. Boise, ID, USA: Association for Computing Machinery, pp. 227–236. URL: https://doi.org/10.1145/3311927.3323138.
Calvo, Rafael A. et al. (2014). "Autonomy in Technology Design". In: *CHI '14 Extended Abstracts on Human Factors in Computing Systems*. CHI EA '14. Toronto, Ontario, Canada: Association for Computing Machinery, pp. 37–40. URL:
https://doi.org/10.1145/2559206.2560468.
Cameron, Lynne (1996). "Discourse context and the development of metaphor in children". In: *Current Issues in Language & Society* 3.1, pp. 49–64.
Castellano, Ginevra et al. (June 2014). "Context-Sensitive Affect Recognition for a Robotic Game Companion". In: *ACM Trans. Interact. Intell. Syst.* 4.2. URL:
https://doi.org/10.1145/2622615.
Chandra, Shruti et al. (2018). "Do Children Perceive Whether a Robotic Peer is Learning or Not?" In: *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*. HRI '18. Chicago, IL, USA: Association for Computing Machinery, pp. 41–49. URL: https://doi.org/10.1145/3171221.3171274.
Chassin, L et al. (1995). *Adolescent health issues: handbook of pediatric psychology.*
Chen, He et al. (2017). "Mutual information analysis of EEG of children with attention-deficit/hyperactivity disorder". In: *2017 Chinese Automation Congress (CAC)*, pp. 2342–2347.
Cheng, Yushi et al. (July 2020). "Identifying Child Users via Touchscreen Interactions". In: *ACM Trans. Sen. Netw.* 16.4. URL: https://doi.org/10.1145/3403574.
Child Rights by Design (2023). URL:
https://childrightsbydesign.digitalfuturescommission.org.uk/.

Children, Committee for (2020). *Social-Emotional Learning, Digital Citizenship, and Media Literacy.* URL: https://www.cfchildren.org/wp-content/uploads/policy-advocacy/digital-citizenship-and-SEL-one-pager.pdf.

Children's Online Privacy Protection Rule ("COPPA") (1998). URL: https://www.ftc.gov/enforcement/rules/rulemaking-regulatory-reform-proceedings/childrens-online-privacy-protection-rule.

China releases regulations to protect minors in cyberspace (2023). URL: https://english.www.gov.cn/policies/latestreleases/202310/24/content_WS6537a5d2c6d0868f4e8e095e.html.

Choi, Da Hyeon (2019). "LYRA: an Interactive and Interactive Storyteller". In: *2019 IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC)*, pp. 148–153.

Coleman, Nick, Wendy Sykes, Carola Groom, et al. (2017). "Peer support and children and young people's mental health". In: *Independent Social Research. London, UK: Department of Education.*

Commision, European (2021). *General Data Protection Regulation (GDPR).* URL: https://gdpr-info.eu/.

Commission, European (2016). *The European Digital Competence Framework for Citizens.*

Corbett-Davies, Sam et al. (2017). "Algorithmic decision making and the cost of fairness". In: *Proceedings of the 23rd acm sigkdd international conference on knowledge discovery and data mining*, pp. 797–806.

Cukier, Kenneth and Viktor Mayer-Schoenberger (2013). "The rise of big data: How it's changing the way we think about the world". In: *Foreign Aff.* 92, p. 28.

Cushman, Fiery (2013). "Action, outcome, and value: A dual-system framework for morality". In: *Personality and social psychology review* 17.3, pp. 273–292.

Dadvar, Maral and Franciska de Jong (2012). "Cyberbullying Detection: A Step toward a Safer Internet Yard". In: *Proceedings of the 21st International Conference on World Wide Web*. WWW '12 Companion. Lyon, France: Association for Computing Machinery, pp. 121–126. URL: https://doi.org/10.1145/2187980.2187995.

Danby, Susan (2017). "Technologies, child-centred practice and listening to children". In: *Digital technologies and learning in the early years*, pp. 127–138.

Dasgupta, Sayamindu (2013). "From Surveys to Collaborative Art: Enabling Children to Program with Online Data". In: *Proceedings of the 12th International Conference on Interaction Design and Children.* IDC '13. New York, New York, USA: Association for Computing Machinery, pp. 28–35. URL: https://doi.org/10.1145/2485760.2485784.

Dasgupta, Sayamindu and Benjamin Mako Hill (2017). "Scratch community blocks: Supporting children as data scientists". In: *Proceedings of the 2017 CHI conference on human factors in computing systems*, pp. 3620–3631.

— (2021). "Designing for critical algorithmic literacies". In: *Algorithmic rights and protections for children.*

Daugherty, Lindsay et al. (2014). "Moving beyond Screen Time: Redefining Developmentally Appropriate Technology Use in Early Childhood Education. Policy Brief." In: *RAND Corporation.*

Deci, Edward L, Haleh Eghrari, et al. (1994). "Facilitating internalization: The self-determination theory perspective". In: *Journal of personality* 62.1, pp. 119–142.

Deci, Edward L and Richard M Ryan (2013). *Intrinsic motivation and self-determination in human behavior*. Springer Science & Business Media.
Dempsey, John et al. (2022). “Children designing privacy warnings: Informing a set of design guidelines”. In: *International Journal of Child-Computer Interaction* 31, p. 100446.
Deng, Xiaozhou et al. (2019). “ARCat: A Tangible Programming Tool for DFS Algorithm Teaching”. In: *Proceedings of the 18th ACM International Conference on Interaction Design and Children*. IDC ’19. Boise, ID, USA: Association for Computing Machinery, pp. 533–537. URL: https://doi.org/10.1145/3311927.3325308.
Dietz, Griffin et al. (2021). “StoryCoder: Teaching Computational Thinking Concepts Through Storytelling in a Voice-Guided App for Children”. In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. CHI ’21. Yokohama, Japan: Association for Computing Machinery. URL: https://doi.org/10.1145/3411764.3445039.
Dillhoff, Alex et al. (2019). “An Automated Assessment System for Embodied Cognition in Children: From Motion Data to Executive Functioning”. In: *Proceedings of the 6th International Workshop on Sensor-Based Activity Recognition and Interaction*. iWOAR ’19. Rostock, Germany: Association for Computing Machinery. URL: https://doi.org/10.1145/3361684.3361693.
Dindler, Christian et al. (2005). “Mission from Mars: a method for exploring user requirements for children in a narrative space”. In: *Proceedings of the 2005 conference on Interaction design and children*, pp. 40–47.
Disconnect.me and Mozilla (n.d.). *Firefox Blocklist*. https://github.com/mozilla-services/shavar-prod-lists.
Dombro, Amy Laura, Judy R Jablon, and Charlotte Stetson (2011). *Powerful interactions: How to connect with children to extend their learning*. National Association for the Education of Young Children Washington, DC.
Dowthwaite, Liz et al. (2020). “" It’s your private information. it’s your life." young people’s views of personal data use by online technologies”. In: *Proceedings of the Interaction Design and Children Conference*, pp. 121–134.
Draft Online Safety Bill, Joint Committee on the (2021). *Draft Online Safety Bill*. URL: https://committees.parliament.uk/publications/8206/documents/84092/default/.
Druga, Stefania and Amy J Ko (2021). “How do children’s perceptions of machine intelligence change when training and coding smart programs?” In: *Interaction Design and Children*, pp. 49–61.
Druga, Stefania, Sarah T Vu, et al. (2019). “Inclusive AI literacy for kids around the world”. In: *Proceedings of FabLearn 2019*, pp. 104–111.
Druga, Stefania, Randi Williams, et al. (2018). “How smart are the smart toys? Children and parents’ agent interaction and intelligence attribution”. In: *Proceedings of the 17th ACM Conference on Interaction Design and Children*, pp. 231–240.
Druga, Stefania, Jason Yip, et al. (2021). “The 4As: Ask, Adapt, Author, Analyze-AI Literacy Framework for Families”. In: *Algorithmic Rights and Protections for Children*. PubPub.
Druin, Allison (2002). “The role of children in the design of new technology”. In: *Behaviour and information technology* 21.1, pp. 1–25.

Duffy, Brian R et al. (1999). "What is a social robot?" In: *10th Irish Conference on Artificial Intelligence & Cognitive Science, University College Cork, Ireland, 1-3 September, 1999*.
Dworkin, Gerald (1981). "The concept of autonomy". In: *Grazer philosophische studien* 12, p. 203.
Eason, Kenneth D (1995). "User-centred design: for users or by users?" In: *Ergonomics* 38.8, pp. 1667–1673.
Education, Common Sense (2017). *Digital citizenship & social and emotional learning: Navigating life's digital dilemmas*. URL: https://www.commonsense.org/education/sites/default/files/tlr-blog/csedigitalcitizenship-sel.pdf.
ERSTAD, OLA et al. (2023). "Datafication in and of Education–a literature". In.
Ethical Web And Data Architectures (2021). URL: https://www.oxfordmartin.ox.ac.uk/ethical-web-and-data-architectures/.
EU AI Act (2021). URL: https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:52021PC0206.
Evangelio, Carlos et al. (2022). "Cyberbullying in elementary and middle school students: A systematic review". In: *Computers & Education* 176, p. 104356.
Exodus (n.d.). *Exodus Privacy*. https://exodus-privacy.eu.org/en/.
Fadhil, Ahmed and Adolfo Villafiorita (2017). "An Adaptive Learning with Gamification and Conversational UIs: The Rise of CiboPoliBot". In: *Adjunct Publication of the 25th Conference on User Modeling, Adaptation and Personalization*. UMAP '17. Bratislava, Slovakia: Association for Computing Machinery, pp. 408–412. URL: https://doi.org/10.1145/3099023.3099112.
Fagen, Robert M (2010). "Play and development". In.
Fasching, Joshua et al. (2012). "Detecting risk-markers in children in a preschool classroom". In: *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1010–1016.
Finck, Michèle and Frank Pallas (2020). "They who must not be identified—distinguishing personal from non-personal data under the GDPR". In: *International Data Privacy Law*.
Finegan, Colleen and Nancy Jo Austin (2002). "Developmentally appropriate technology for young children". In: *Information Technology in Childhood Education Annual* 2002.1, pp. 87–102.
Fisher, Eunice (1993). "Characteristics of children's talk at the computer and its relationship to the computer software". In: *Language and Education* 7.2, pp. 97–114.
Fletcher, Anne C and Bethany L Blair (2014). "Maternal authority regarding early adolescents' social technology use". In: *Journal of Family Issues* 35.1, pp. 54–74.
Foundation, 5 Rights (2021). *Playful by Design: Free play in a digital world*. URL: https://digitalfuturescommission.org.uk/wp-content/uploads/2021/11/A-Vision-of-Free-Play-in-a-Digital-World.pdf.
Frijda, Nico H and Batja Mesquita (1994). "The social roles and functions of emotions." In.
Frikha, Mohamed Heni, Tarek Zlitni, and Nadia Bouassida (2020). "Towards a Recommendation System for Children's Animated Movies Based on Chromatic Features". In: *Proceedings of the 2nd International Conference on Digital Tools*. DTUC '20. Virtual Event, Tunisia: Association for Computing Machinery. URL: https://doi.org/10.1145/3423603.3424058.

Fuchs, Christian (2011). “Web 2.0, prosumption, and surveillance”. In: *Surveillance & Society* 8.3, pp. 288–309.
Fuertes, Walter et al. (2015). “On the Development of Advanced Parental Control Tools”. In: *2015 1st International Conference on Software Security and Assurance (ICSSA)*, pp. 1–6.
Furnham, Philipp (2019). “Automation and autonomy of big data-driven algorithmic decision-making”. In: *Contemporary Readings in Law and Social Justice* 11.1, pp. 51–56.
Gamaethige, Chanuki et al. (2017). ““SenseA”-Autism Early Signs and Pre-Aggressive Detector through Image Processing”. In: *2017 Asia Modelling Symposium (AMS)*, pp. 125–130.
Garg, Radhika and Subhasree Sengupta (Mar. 2020). “He Is Just Like Me: A Study of the Long-Term Use of Smart Speakers by Parents and Children”. In: *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 4.1. URL: https://doi.org/10.1145/3381002.
Gauthier, Andrea et al. (2022). “Manipulating Interface Design Features Affects Children’s Stop-And-Think Behaviours in a Counterintuitive-Problem Game”. In: *ACM Transactions on Computer-Human Interaction* 29.2, pp. 1–22.
General comment No. 25 (2021) on children’s rights in relation to the digital environment (2020). URL: https://www.ohchr.org/EN/HRBodies/CRC/Pages/GCChildrensRightsRelationDigitalEnvironment.aspx.
Generation AI: Establishing Global Standards for Children and AI (2019). URL: http://www3.weforum.org/docs/WEF_Generation_AI_%5C%20May_2019_Workshop_Report.pdf.
Ghosh, Arup Kumar, Karla Badillo-Urquiola, et al. (2018). “Safety vs. surveillance: what children have to say about mobile apps for parental control”. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 1–14.
Ghosh, Arup Kumar, Charles E Hughes, and Pamela J Wisniewski (2020). “Circle of trust: a new approach to mobile online safety for families”. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1–14.
Goldstein, Daniel G et al. (2008). “Nudge your customers toward better choices”. In: *Harvard Business Review* 86.12, pp. 99–105.
Goodfellow, Ian, Yoshua Bengio, and Aaron Courville (2016). *Deep learning*. MIT press.
Google (2021). *Be Internet Legends*. URL: https://beinternetlegends.withgoogle.com/en_ie.
— (2023). *About privacy and personalised ads*. URL: https://myadcenter.google.com/home?sasb=true&ref=ad-settings.
Google Privacy&Terms (2023). URL: https://policies.google.com/privacy?hl=en-US.
“google-play-scraper 0.1.1” (2020). In.
Graesser, Arthur C, Xiangen Hu, and Robert Sottilare (2018). “Intelligent tutoring systems”. In: *International handbook of the learning sciences*. Routledge, pp. 246–255.
Gritschacher, Tobias and Wolfgang Slany (2012). “Standing on the shoulders of their peers: success factors for massive cooperation among children creating open source animations and games on their smartphones”. In: *Proceedings of the 11th International Conference on Interaction Design and Children*, pp. 264–267.
Grolnick, Wendy S, Edward L Deci, and Richard M Ryan (1997). “Internalization within the family: The self-determination theory perspective”. In: *Parenting and children’s internalization of values: A handbook of contemporary theory*, pp. 135–161.

Group, The Lego (2021). *The LEGO Group extends its Digital Citizenship programme to help families across the globe become digitally smart.* URL: https://www.lego.com/en-gb/aboutus/news/2021/october/digital-citizenship?locale=en-gb.

Groves, Philip M and Richard F Thompson (1970). "Habituation: a dual-process theory." In: *Psychological review* 77.5, p. 419.

Guha, Mona Leigh, Allison Druin, Gene Chipman, et al. (2004). "Mixing ideas: a new technique for working with young children as design partners". In: *Proceedings of the 2004 conference on Interaction design and children: building a community*, pp. 35–42.

Guha, Mona Leigh, Allison Druin, and Jerry Alan Fails (2013). "Cooperative Inquiry revisited: Reflections of the past and guidelines for the future of intergenerational co-design". In: *International Journal of Child-Computer Interaction* 1.1, pp. 14–23. URL: https://www.sciencedirect.com/science/article/pii/S2212868912000049.

Guidance for Regulation of Artificial Intelligence Applications (2020). URL: https://www.whitehouse.gov/wp-content/uploads/2020/11/M-21-06.pdf.

Guzman, Emitza and Walid Maalej (2014). "How do users like this feature? a fine grained sentiment analysis of app reviews". In: *2014 IEEE 22nd international requirements engineering conference (RE).* IEEE, pp. 153–162.

Hannum, Hurst (2011). *Autonomy, sovereignty, and self-determination.* University of Pennsylvania press.

Hashish, Yasmeen, Andrea Bunt, and James E Young (2014). "Involving children in content control: a collaborative and education-oriented content filtering approach". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1797–1806.

Hautea, Samantha, Sayamindu Dasgupta, and Benjamin Mako Hill (2017). "Youth perspectives on critical data literacies". In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pp. 919–930.

Hayes-Roth, Frederick (1985). "Rule-based systems". In: *Communications of the ACM* 28.9, pp. 921–932.

Hazman, Maryam and Amira M. Idrees (2015). "A healthy nutrition expert system for children". In: *2015 E-Health and Bioengineering Conference (EHB)*, pp. 1–4.

Heinze, Clint Andrew, Janet Haase, and Helen Higgins (2010). "An action research report from a multi-year approach to teaching artificial intelligence at the k-6 level". In: *First AAAI Symposium on Educational Advances in Artificial Intelligence.*

Hémar-Nicolas, Valerie et al. (2021). "How cartoon characters and claims influence children's attitude towards a snack vegetable–An explorative cross-cultural comparison between Indonesia and Denmark". In: *Food Quality and Preference* 87, p. 104031.

Hendrix, Susan L, Michael Eisenberg, et al. (2006). "Computer-assisted pop-up design for children: computationally enriched paper engineering". In: *Advanced Technology for Learning* 3.2, pp. 119–127.

Hiniker, Alexis, Sharon S Heung, et al. (2018). "Coco's videos: An empirical investigation of video-player design features and children's media use". In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 1–13.

Hiniker, Alexis, Bongshin Lee, et al. (2017). "Plan and Play: Supporting Intentional Media Use in Early Childhood". In: *Proceedings of the 2017 Conference on Interaction*

Design and Children. IDC '17. Stanford, California, USA: Association for Computing Machinery, pp. 85–95. URL: https://doi.org/10.1145/3078072.3079752.

Hiniker, Alexis, Kiley Sobel, and Bongshin Lee (2017). "Co-designing with preschoolers using fictional inquiry and comicboarding". In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pp. 5767–5772.

Hitron, Tom et al. (2019). "Can children understand machine learning concepts? The effect of uncovering black boxes". In: *Proceedings of the 2019 CHI conference on human factors in computing systems*, pp. 1–11.

Hmel, Beverly A and Aaron L Pincus (2002). "The meaning of autonomy: On and beyond the interpersonal circumplex". In: *Journal of personality* 70.3, pp. 277–310.

Hood, Deanna, Séverin Lemaignan, and Pierre Dillenbourg (2015). "When Children Teach a Robot to Write: An Autonomous Teachable Humanoid Which Uses Simulated Handwriting". In: *2015 10th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pp. 83–90.

Hou, Weimin et al. (2015). "Supporting children's online identity in international communities". In: *Behaviour & Information Technology* 34.4, pp. 375–391.

Hourcade, Juan Pablo et al. (2016). "Child-computer interaction SIG: New challenges and opportunities". In: *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, pp. 1123–1126.

Hsi, Sherry and Michael Eisenberg (2012). "Math on a sphere: Using public displays to support children's creativity and computational thinking on 3D surfaces". In: *Proceedings of the 11th International Conference on Interaction Design and Children*, pp. 248–251.

Hsieh, Hsiu-Fang and Sarah E Shannon (2005). "Three approaches to qualitative content analysis". In: *Qualitative health research* 15.9, pp. 1277–1288.

Hughes, Fergus P (2021). *Children, play, and development*. SAGE publications.

Humane by Design (2022). URL: https://humanebydesign.com/.

Hutchinson, Ben and Margaret Mitchell (2019). "50 Years of Test (Un)Fairness: Lessons for Machine Learning". In: *Proceedings of the Conference on Fairness, Accountability, and Transparency*. FAT* '19. Atlanta, GA, USA: Association for Computing Machinery, pp. 49–58. URL: https://doi.org/10.1145/3287560.3287600.

ICO (2021). *Designing data transparency for children*. URL: https://ico.org.uk/media/2620177/designing-data-transparency-for-children.pdf.

IEEE (2021). *IEEE Standard for an Age Appropriate Digital Services Framework Based on the 5Rights Principles for Children*. https://standards.ieee.org/ieee/2089/7633/.

Institute, DQ (2021). *Digital Intellegence Framework*. URL: https://www.dqinstitute.org/global-standards/.

Internet Safety, UK Council for (2020). *Education for a Connected World - 2020 Edition*. URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/896323/UKCIS_Education_for_a_Connected_World_.pdf.

Ivanov, Radoslav et al. (2015). "Early Detection of Critical Pulmonary Shunts in Infants". In: *Proceedings of the ACM/IEEE Sixth International Conference on Cyber-Physical Systems*. ICCPS '15. Seattle, Washington: Association for Computing Machinery, pp. 110–119. URL: https://doi.org/10.1145/2735960.2735962.

Jacq, Alexis et al. (2016). “Building successful long child-robot interactions in a learning context”. In: *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, pp. 239–246.

Jordan, Michael I and Tom M Mitchell (2015). “Machine learning: Trends, perspectives, and prospects”. In: *Science* 349.6245, pp. 255–260.

Kafai, Yasmin, Chris Proctor, and Debora Lui (2020). “From theory bias to theory dialogue: embracing cognitive, situated, and critical framings of computational thinking in K-12 CS education”. In: *ACM Inroads* 11.1, pp. 44–53.

Kahn, Peter H et al. (2006). “Robotic pets in the lives of preschool children”. In: *Interaction Studies* 7.3, pp. 405–436.

Kaushal, Rishabh et al. (2016). “KidsTube: Detection, characterization and analysis of child unsafe content promoters on YouTube”. In: *2016 14th Annual Conference on Privacy, Security and Trust (PST)*, pp. 157–164.

Kazai, Gabriella, Iskander Yusof, and Daoud Clarke (2016). “Personalised news and blog recommendations based on user location, Facebook and Twitter user profiling”. In: *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval*, pp. 1129–1132.

Keltner, Dacher and Jennifer S Lerner (2010). “Emotion.” In.

Khayat, Ghada A. El, Tamer F. Mabrouk, and Adel S. Elmaghraby (2012). “Intelligent serious games system for children with learning disabilities”. In: *2012 17th International Conference on Computer Games (CGAMES)*, pp. 30–34.

Kidron, Baroness, Alexandra Evans, and Jenny Afia (2018). “Disrupted Chldhood”. In.

Kids Online Safety Act: Legislation to impose responsibility on online platforms and equip children and parents with tools (2023). URL: `https://www.commonsensemedia.org/sites/default/files/featured-content/files/kosa-one-pager.pdf`.

Killmister, Suzy (2013). “Autonomy and false beliefs”. In: *Philosophical Studies* 164.2, pp. 513–531.

Ko, Minsam et al. (2015). “FamiLync: facilitating participatory parental mediation of adolescents’ smartphone use”. In: *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pp. 867–878.

Kochigami, Kanae et al. (2015). “Walking together hand in hand: Design and evaluation of autonomous robot system that a robot recognizes moving direction with a child’s assistance of pulling its hand”. In: *2015 IEEE/SICE International Symposium on System Integration (SII)*, pp. 738–743.

Kollnig, Konrad, Reuben Binns, et al. (2021). “Before and after GDPR: tracking in mobile apps”. In: *arXiv preprint arXiv:2112.11117*.

Kollnig, Konrad and Nigel Shadbolt (2022). “TrackerControl: Transparency and Choice around App Tracking”. In: *Journal of Open Source Software* 7.75, p. 4270.

Kumar, Priya et al. (2018). “Co-designing online privacy-related games and stories with children”. In: *Proceedings of the 17th ACM conference on interaction design and children*, pp. 67–79.

Kumaresamoorthy, N. and M.F.M. Firdhous (2018). “An APPROACH OF Filtering The Content Of Posts In Social Media”. In: *2018 3rd International Conference on Information Technology Research (ICITR)*, pp. 1–6.

Kuss, Daria J and Mark D Griffiths (2017). “Social networking sites and addiction: Ten lessons learned”. In: *International journal of environmental research and public health* 14.3, p. 311.

Kutrovátz, Kitti et al. (2022). “Parental mediation of adolescents’ technology use. Unequal parenting practices”. In: *Intersections. East European Journal of Society and Politics* 8.3, pp. 99–117.
Laal, Marjan and Seyed Mohammad Ghodsi (2012). “Benefits of collaborative learning”. In: *Procedia-social and behavioral sciences* 31, pp. 486–490.
Lamborn, Susie D and Laurence Steinberg (1993). “Emotional autonomy redux: revisiting Ryan and Lynch”. In: *Child development* 64.2, pp. 483–499.
Langford, Rachel (2010). “Critiquing child-centred pedagogy to bring children and early childhood educators into the centre of a democratic pedagogy”. In: *Contemporary Issues in Early Childhood* 11.1, pp. 113–127.
Lapidoth, Ruth (1997). *Autonomy: Flexible solutions to ethnic conflicts*. US Institute of Peace Press.
Lee, Charlotte P (2007). “Boundary negotiating artifacts: Unbinding the routine of boundary objects and embracing chaos in collaborative work”. In: *Computer Supported Cooperative Work (CSCW)* 16.3, pp. 307–339.
Lee, Kung Jin et al. (2021). “The show must go on: A conceptual model of conducting synchronous participatory design with children online”. In: *Proceedings of the 2021 CHI conference on human factors in computing systems*, pp. 1–16.
Lee, Tak Yeon et al. (2014). “CTArcade: Computational thinking with games in school age children”. In: *International Journal of Child-Computer Interaction* 2.1, pp. 26–33.
Lego (2021). *Captain Safety*. URL: `https://www.lego.com/en-gb/life/digital-safety`.
Leite, Iolanda et al. (2016). “Autonomous disengagement classification and repair in multiparty child-robot interaction”. In: *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pp. 525–532.
Leonard, Thomas C (2008). *Richard H. Thaler, Cass R. Sunstein, Nudge: Improving decisions about health, wealth, and happiness.*
Leslie, David (June 2019). *Understanding artificial intelligence ethics and safety: A guide for the responsible design and implementation of AI systems in the public sector*. URL: `https://doi.org/10.5281/zenodo.3240529`.
Lewis, Clayton and Cathleen Wharton (1997). “Cognitive walkthroughs”. In: *Handbook of human-computer interaction*. Elsevier, pp. 717–732.
Light, Ben, Jean Burgess, and Stefanie Duguay (2018). “The walkthrough method: An approach to the study of apps”. In: *New media & society* 20.3, pp. 881–900.
Lin, Pei-Chun et al. (2019). “Children Privacy Identification System in LINE Chatbot for Smart Toys”. In: *2019 IEEE 12th International Conference on Cloud Computing (CLOUD)*, pp. 86–90.
Litts, Breanne K et al. (2019). “I’m Drowning in Squirrels! How Children Embody and Debug Computational Algorithms Through Designing Mixed Reality Games”. In: *Proceedings of the 18th ACM International Conference on Interaction Design and Children*, pp. 267–273.
Livingstone, S. and K Pothong (2021). *Playful by Design: A Vision of Free Play in a Digital World*. Tech. rep. Digital Futures Commission.
— (2023). *Child Rights by Design: Guidance for Innovators of Digital Products and Services Used by Children*. Tech. rep. Digital Futures Commission, 5Rights Foundation.
Livingstone, Sonia (2018). “Children: a special case for privacy?” In: *Intermedia* 46.2, pp. 18–23.

Livingstone, Sonia and Jasmina Byrne (2018). "Parenting in the digital age: The challenges of parental responsibility in comparative perspective". In.

Livingstone, Sonia, Mariya Stoilova, and Rishita Nandagiri (2019). "Children's data and privacy online". In: *Technology* 58.2, pp. 157–65.

Loevinger, Jane et al. (2014). *Measuring ego development*. Psychology Press.

Long, Duri and Brian Magerko (2020). "What is AI literacy? Competencies and design considerations". In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1–16.

Lords House of Commons; Joint Committee on the Draft Online Safety Bill, House of (2021). "Draft Online Safety Bill". In: URL: `https://www.gov.uk/government/publications/draft-online-safety-bill`.

Lowry, Richard (2014). "Concepts and applications of inferential statistics". In.

Lukoff, Kai, Ulrik Lyngs, Karina Shirokova, et al. (2023). "SwitchTube: A Proof-of-Concept System Introducing "Adaptable Commitment Interfaces" as a Tool for Digital Wellbeing". In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. CHI '23. Hamburg, Germany: Association for Computing Machinery. URL: `https://doi.org/10.1145/3544548.3580703`.

Lukoff, Kai, Ulrik Lyngs, Himanshu Zade, et al. (2021). "How the Design of YouTube Influences User Sense of Agency". In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–17.

Lyngs, Ulrik et al. (2019). "Self-control in cyberspace: Applying dual systems theory to a review of digital self-control tools". In: *proceedings of the 2019 CHI conference on human factors in computing systems*, pp. 1–18.

M., Premasundari and Yamini C. (2019). "Food and Therapy Recommendation System for Autistic Syndrome using Machine Learning Techniques". In: *2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT)*, pp. 1–6.

Macedo, João, Filipe Costa, and Jefersson A. dos Santos (2018). "A Benchmark Methodology for Child Pornography Detection". In: *2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*, pp. 455–462.

Madrazo Azpiazu, Ion et al. (2018). "Looking for the Movie Seven or Sven from the Movie Frozen? A Multi-Perspective Strategy for Recommending Queries for Children". In: *Proceedings of the 2018 Conference on Human Information Interaction and Retrieval*. CHIIR '18. New Brunswick, NJ, USA: Association for Computing Machinery, pp. 92–101. URL: `https://doi.org/10.1145/3176349.3176379`.

Maharani, Septya et al. (2015). "Expert system applications for early diagnosis teeth and oral disease in children". In: *2015 2nd International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE)*, pp. 87–91.

Maqsood, Sana, Christine Mekhail, and Sonia Chiasson (2018). "A Day in the Life of Jos: A Web-based Game to Increase Children's Digital Literacy". In: *Proceedings of the 17th ACM Conference on Interaction Design and Children*. ACM, pp. 241–252.

Mascheroni, Giovanna (2020). "Datafied childhoods: Contextualising datafication in everyday life". In: *Current Sociology* 68.6, pp. 798–813.

Matamoros-Fernández, Ariadna (2017). "Platformed racism: The mediation and circulation of an Australian race-based controversy on Twitter, Facebook and YouTube". In: *Information, Communication & Society* 20.6, pp. 930–946.

McGrath, Cecile Hoareau et al. (2015). "Learning gain in higher education". In: *Santa Monica, CA: RAND Corporation*.

McLeod, Saul (2017). “Kolb’s learning styles and experiential learning cycle”. In: *Simply psychology* 5.
McNally, Brenna et al. (2018). “Co-designing mobile online safety applications with children”. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 1–9.
Media, Common Sense (2022). *Two Years Into the Pandemic, Media Use Has Increased 17% Among Tweens and Teens*. URL: `https://www.commonsensemedia.org/press-releases/two-years-into-the-pandemic-media-use-has-increased-17-among-tweens-and-teens`.
Mejias, Ulises A and Nick Couldry (2019). “Datafication”. In: *Internet Policy Review* 8.4.
Mercer, Neil and Rupert Wegerif (2002). *Is ‘exploratory talk’productive talk?* Routledge.
Meta Privacy Policy (2023). URL: `https://en-gb.facebook.com/privacy/policy/`.
Metcalfe, Janet and Matthew Jason Greene (2007). “Metacognition of agency.” In: *Journal of Experimental Psychology: General* 136.2, p. 184.
Michalski, Dana, Sau Yee Yiu, and Chris Malec (2018). “The Impact of Age and Threshold Variation on Facial Recognition Algorithm Performance Using Images of Children”. In: *2018 International Conference on Biometrics (ICB)*, pp. 217–224.
Miller, George A (1995). “WordNet: a lexical database for English”. In: *Communications of the ACM* 38.11, pp. 39–41.
Miller, Sara and Lisa Pennycuff (2008). “The power of story: Using storytelling to improve literacy learning”. In: *Journal of Cross-Disciplinary Perspectives in Education* 1.1, pp. 36–43.
Mitra, Avijit, Tahsin Mostafiz, and Raihan Ur Rashid (2017). “Photoplay: An android application to stimulate children’s cognitive development”. In: *2017 IEEE Region 10 Humanitarian Technology Conference (R10-HTC)*, pp. 424–427.
Mitrou, Lilian et al. (2014). “Social media profiling: A Panopticon or Omniopticon tool?” In: *Proc. of the 6th Conference of the Surveillance Studies Network*. Barcelona, Spain, pp. 1–15.
Mittelstadt, Brent (2016). “Automation, algorithms, and politics| auditing for transparency in content personalization systems”. In: *International Journal of Communication* 10, p. 12.
Moghadas, M. and H. Moradi (2018). “Analyzing Human-Robot Interaction Using Machine Vision for Autism screening”. In: *2018 6th RSI International Conference on Robotics and Mechatronics (IcRoM)*, pp. 572–576.
Mohanta, Abhijit and Vinay Kumar Mittal (2020). “Classifying Speech of ASD Affected and Normal Children Using Acoustic Features”. In: *2020 National Conference on Communications (NCC)*, pp. 1–6.
Morgan, Myfanwy et al. (2002). “Hearing children’s voices: methodological issues in conducting focus groups with children aged 7-11 years”. In: *Qualitative research* 2.1, pp. 5–20.
Muller, Michael J and Sarah Kuhn (1993). “Participatory design”. In: *Communications of the ACM* 36.6, pp. 24–28.
Mullin, Amy (2014). “Children, paternalism and the development of autonomy”. In: *Ethical Theory and Moral Practice* 17.3, pp. 413–426.
Nasi, Liora et al. (2019). “Pomelo, a collaborative education technology interaction robot”. In: *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, pp. 757–758.

Nation, United (2023). *The Convention on the Rights of the Child: The children's version.* URL: https://www.unicef.org/media/60981/file/convention-rights-child-text-child-friendly-version.pdf.

Neel, Elizabeth U, Susan Jay, and Iris F Litt (1985). "The relationship of self-concept and autonomy to oral contraceptive compliance among adolescent females". In: *Journal of adolescent health care* 6.6, pp. 445–447.

Nezhad, Mahshid Mehr and Maryam Mehrnezhad (2018). "A Child Recognition System Based on Image Selection Patterns". In: *Proceedings of the 7th Workshop on Socio-Technical Aspects in Security and Trust*. STAST '17. Orlando, Florida, USA: Association for Computing Machinery, pp. 76–81. URL: https://doi.org/10.1145/3167996.3168003.

Nissenbaum, Helen (2011). "A contextual approach to privacy online". In: *Daedalus* 140.4, pp. 32–48.

O'Keeffe, Gwenn Schurgin et al. (2011). "The impact of social media on children, adolescents, and families". In: *Pediatrics* 127.4, pp. 800–804.

Ofcom (2021). *Children and parents: media use and attitudes report 2020/21*. URL: https://www.ofcom.org.uk/__data/assets/pdf_file/0025/217825/children-and-parents-media-use-and-attitudes-report-2020-21.pdf.

— (2022). *Children and parents: media use and attitudes report 2022*. URL: https://www.ofcom.org.uk/__data/assets/pdf_file/0025/217825/children-and-parents-media-use-and-attitudes-report-2022.pdf.

— (2023). *Children and parents: media use and attitudes report 2023*. URL: https://www.ofcom.org.uk/__data/assets/pdf_file/0027/255852/childrens-media-use-and-attitudes-report-2023.pdf.

Ofcom.org (n.d.). *children and parents: media use and attitudes report 2019*. URL: https://www.ofcom.org.uk/__data/assets/pdf_file/0023/190616/children-media-use-attitudes-2019-report.pdf.

Ofer, Netta et al. (2019). "Coding for outdoor play: a coding platform for children to invent and enhance outdoor play experiences". In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pp. 1–12.

Pacurucu-Pacurucu, A. et al. (2016). "An intelligent ecosystem to support psychological evaluation and intervention of children under social risk: Towards graphics design validation and new findings". In: *IEEE CACIDI 2016 - IEEE Conference on Computer Sciences*, pp. 1–6.

Page, Matthew J et al. (2021). "The PRISMA 2020 statement: an updated guideline for reporting systematic reviews". In: *Systematic reviews* 10.1, pp. 1–11.

Pangrazio, Luci and Lourdes Cardozo-Gaibisso (2021). ""Your Data Can Go to Anyone": The Challenges of Developing Critical Data Literacies in Children". In: *Critical digital literacies: Boundary-crossing practices*. Brill, pp. 35–51.

Parker, Andrea Grimes et al. (2013). "I am what I eat: Identity & critical thinking in an online health forum for kids". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 2437–2446.

Parliament, UK et al. (2022). "Online Safety Bill (2022)". In.

Paterson, Ronald William Keith (2010). *Values, education and the adult*. Routledge.

Pera, Maria Soledad and Yiu-Kai Ng (2013). "What to Read next? Making Personalized Book Recommendations for K-12 Users". In: *Proceedings of the 7th ACM Conference on Recommender Systems*. RecSys '13. Hong Kong, China: Association for Computing Machinery, pp. 113–120. URL: https://doi.org/10.1145/2507157.2507181.

Peters, Dorian, Rafael A Calvo, and Richard M Ryan (2018). "Designing for motivation, engagement and wellbeing in digital experience". In: *Frontiers in psychology* 9, p. 797.
Peterson-Salahuddin, Chelsea and Nicholas Diakopoulos (2020). "Negotiated autonomy: The role of social media algorithms in editorial decision making". In: *Media and Communication* 8.3.
Piccolo, Lara Schibelsky Godoy, Pinelopi Troullinou, and Harith Alani (2021). "Chatbots to support children in coping with online threats: Socio-technical requirements". In: *Designing Interactive Systems Conference 2021*, pp. 1504–1517.
Poblet, Marta et al. (2017). "A co-regulatory approach to stay safe online: reporting inappropriate content with the MediaKids mobile app". In: *Journal of Family Studies* 23.2, pp. 180–197.
PwC's Global Artificial Intelligence Study: Exploiting the AI Revolution (2021). URL: https://www.pwc.com/gx/en/issues/data-and-analytics/publications/artificial-intelligence-study.html.
Rafiq, Rahat Ibn et al. (2018). "Scalable and Timely Detection of Cyberbullying in Online Social Networks". In: *Proceedings of the 33rd Annual ACM Symposium on Applied Computing*. SAC '18. Pau, France: Association for Computing Machinery, pp. 1738–1747. URL: https://doi.org/10.1145/3167132.3167317.
Rahman, Feisal Dani et al. (2014). "Automatic speech recognition system for Malay speaking children". In: *2014 Third ICT International Student Project Conference (ICT-ISPC)*, pp. 79–82.
Ramachandran, Aditi, Chien-Ming Huang, and Brian Scassellati (2017). "Give Me a Break! Personalized Timing Strategies to Promote Learning in Robot-Child Tutoring". In: *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*. HRI '17. Vienna, Austria: Association for Computing Machinery, pp. 146–155. URL: https://doi.org/10.1145/2909824.3020209.
Rao, Ashwini, Florian Schaub, and Norman Sadeh (2015). "What do they know about me? Contents and concerns of online behavioral profiles". In: *arXiv preprint arXiv:1506.01675*.
Rasheed, Osama, Aimal Rextin, and Mehwish Nasim (2019). "Adult or Child: Recognizing through Touch Gestures on Smartphones". In: *2019 Digital Image Computing: Techniques and Applications (DICTA)*, pp. 1–8.
Al-Razgan, Muna and Safa Alshaarri (2019). "Design and Development of a Mobile Spelling Game for Elementary Students Using Genetic Algorithms". In: *Proceedings of the 2019 11th International Conference on Education Technology and Computers*. ICETC 2019. Amsterdam, Netherlands: Association for Computing Machinery, pp. 205–209. URL: https://doi.org/10.1145/3369255.3369311.
Redrován-Reyes, E. et al. (2019). "An educational platform based on expert systems, speech recognition, and ludic activities to support the lexical and semantic development in children from 2 to 3 years". In: *2019 IEEE Colombian Conference on Communications and Computing (COLCOM)*, pp. 1–6.
Resnick, Mitchel, Robbie Berg, and Michael Eisenberg (2000). "Beyond black boxes: Bringing transparency and aesthetics back to scientific investigation". In: *The Journal of the Learning Sciences* 9.1, pp. 7–30.
Responsible AI for Social Empowerment and Education (2023). URL: https://raise.mit.edu/.

Richards, Deborah, Patrina HY Caldwell, and Henry Go (2015). "Impact of social media on the health of children and young people". In: *Journal of paediatrics and child health* 51.12, pp. 1152–1157.
Ricoeur, Paul (1966). *Freedom and nature: The voluntary and the involuntary.* Vol. 1. Northwestern University Press.
Rieder, Bernhard (2017). "Scrutinizing an algorithmic technique: The Bayes classifier as interested reading of reality". In: *Information, Communication & Society* 20.1, pp. 100–117.
Riehm, Kira E et al. (2019). "Associations between time spent using social media and internalizing and externalizing problems among US youth". In: *JAMA psychiatry* 76.12, pp. 1266–1273.
Ringeval, Fabien et al. (2011). "Automatic Intonation Recognition for the Prosodic Assessment of Language-Impaired Children". In: *IEEE Transactions on Audio, Speech, and Language Processing* 19.5, pp. 1328–1342.
Robles-Bykbaev, V., E. Andrade-Prieto, et al. (2018). "An educational support tool based on robotic assistants, mobile apps, and expert systems for children with Down syndrome". In: *2018 IEEE XXV International Conference on Electronics, Electrical Engineering and Computing (INTERCON)*, pp. 1–4.
Robles-Bykbaev, V., V. Velásquez-Angamarca, et al. (2018). "A proposal of a virtual robotic assistant and a rule-based expert system to carry out therapeutic exercises with children with Dyslalia". In: *2018 IEEE Third Ecuador Technical Chapters Meeting (ETCM)*, pp. 1–6.
Robles-Bykbaev, Vladimir E. et al. (2016). "An ontology-based expert system to generate therapy plans for children with disabilities and communication disorders". In: *2016 IEEE Ecuador Technical Chapters Meeting (ETCM)*, pp. 1–6.
Rode, Jennifer A. et al. (2015). "From Computational Thinking to Computational Making". In: *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing.* UbiComp '15. Osaka, Japan: Association for Computing Machinery, pp. 239–250. URL: https://doi.org/10.1145/2750858.2804261.
Rodriguez, M. Y., D. DePanfilis, and P. Lanier (2019). "Bridging the gap: Social work insights for ethical algorithmic decision-making in human services". In: *IBM Journal of Research and Development* 63.4/5, 8:1–8:8.
Rolfe, Sharne A (2020). *Rethinking attachment for early childhood practice: Promoting security, autonomy and resilience in young children.* Routledge.
Rowe, Meredith L, Şeyda Özçalışkan, and Susan Goldin-Meadow (2008). "Learning words by hand: Gesture's role in predicting vocabulary development". In: *First language* 28.2, pp. 182–199.
Rubegni, Elisa et al. (2022). "Raising Awareness of Stereotyping Through Collaborative Digital Storytelling: Design for Change with and for Children". In: *International Journal of Human-Computer Studies* 157, p. 102727.
Ryan, Richard M and Edward L Deci (2004). "Autonomy Is No Illusion: Self-Determination Theory and the Empirical Study of Authenticity, Awareness, and Will." In.
— (2006). "Self-regulation and the problem of human autonomy: Does psychology need choice, self-determination, and will?" In: *Journal of personality* 74.6, pp. 1557–1586.
Sahebi, Siavosh and Paul Formosa (2022). "Social media and its negative impacts on autonomy". In: *Philosophy & Technology* 35.3, p. 70.

Samonte, Mary Jane C. et al. (2019). “Assistive Gamification and Speech Recognition E-Tutor System for Speech Impaired Students”. In: *Proceedings of the 2nd International Conference on Image and Graphics Processing.* ICIGP ’19. Singapore, Singapore: Association for Computing Machinery, pp. 37–41. URL: `https://doi.org/10.1145/3313950.3313968`.

Sangal, Nikita et al. (2021). “Teen-alyse: A Mobile Application for Parental control, Teen Self-Monitoring and Active Mediation”. In: *2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO).* IEEE, pp. 1–5.

Sangi, Mehrdad, Benjamin Thompson, and Jason Turuwhenua (2015). “An Optokinetic Nystagmus Detection Method for Use With Young Children”. In: *IEEE Journal of Translational Engineering in Health and Medicine* 3, pp. 1–10.

Santarcangelo, Joseph and Xiao-Ping Zhang (2014). “Automatic age recommendation system for children’s video content”. In: *2014 IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. 750–753.

Sartre, Jean-Paul (2022). *Being and nothingness: An essay in phenomenological ontology.* Taylor & Francis.

Sax, Marijn (2016). “Big data: Finders keepers, losers weepers?” In: *Ethics and Information Technology* 18.1, pp. 25–31.

Scalise, Kathleen, Michelle Douskey, and Angelica Stacy (2018). “Measuring learning gains and examining implications for student success in STEM”. In: *Higher Education Pedagogies* 3.1, pp. 183–195.

Schiano, Diane J and Christine Burg (2017). “Parental controls: Oxymoron and design opportunity”. In: *HCI International 2017–Posters’ Extended Abstracts: 19th International Conference, HCI International 2017, Vancouver, BC, Canada, July 9–14, 2017, Proceedings, Part II 19.* Springer, pp. 645–652.

Schön, Martin, Martin Ebner, and Georg Kothmeier (2012). “It’s Just about Learning the Multiplication Table”. In: *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge.* LAK ’12. Vancouver, British Columbia, Canada: Association for Computing Machinery, pp. 73–81. URL: `https://doi.org/10.1145/2330601.2330624`.

Sendari, Siti, Triyanna Widyaningtyas, and Nur Amelia Maulidia (2019). “Classification of Toddler Nutrition Status with Anthropometry using the K-Nearest Neighbor Method”. In: *2019 International Conference on Electrical, Electronics and Information Engineering (ICEEIE).* Vol. 6, pp. 1–5.

Shahin, Mostafa, Usman Zafar, and Beena Ahmed (2020). “The Automatic Detection of Speech Disorders in Children: Challenges, Opportunities, and Preliminary Results”. In: *IEEE Journal of Selected Topics in Signal Processing* 14.2, pp. 400–412.

Sharma, Kavya et al. (2021). “Impact of digital nudging on information security behavior: an experimental study on framing and priming in cybersecurity”. In: *Organizational Cybersecurity Journal: Practice, Process and People.*

Shifflet-Chila, Erica D et al. (2016). “Adolescent and family development: Autonomy and identity in the digital age”. In: *Children and Youth Services Review* 70, pp. 364–368.

Shin, Hyunjin and Gahgene Gweon (2020). “Supporting preschoolers’ transitions from screen time to screen-free time using augmented reality and encouraging offline leisure activity”. In: *Computers in Human Behavior* 105, p. 106212.

Shipman, Frank M. and Catherine C. Marshall (Apr. 2020). “Ownership, Privacy, and Control in the Wake of Cambridge Analytica: The Relationship between Attitudes

and Awareness”. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Honolulu HI USA: ACM, pp. 1–12. (Visited on 06/26/2020).
Shklovski, Irina et al. (2014). “Leakiness and Creepiness in App Space: Perceptions of Privacy and Mobile App Use”. In: *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems - CHI '14*. Toronto, Ontario, Canada: ACM Press, pp. 2347–2356. (Visited on 02/14/2020).
Shukla, Smita, Craig H Kennedy, and Lisa Sharon Cushing (1998). “Adult influence on the participation of peers without disabilities in peer support programs”. In: *Journal of Behavioral Education* 8.4, pp. 397–413.
Siegler, Robert S (1991). *Children's thinking*. Prentice-Hall, Inc.
Silverberg, SB and DM Gondoli (1996). *Autonomy in adolescence. A contextualized perspective. En GR Adams, R. Montemayor, & TP Gullotta (Eds.) Psychosocial development during adolescence. Thousand Oak*.
Skinner, Ellen A and Michael J Belmont (1993). “Motivation in the classroom: Reciprocal effects of teacher behavior and student engagement across the school year.” In: *Journal of educational psychology* 85.4, p. 571.
Sobel, Kiley et al. (2017). “EduFeed: A Social Feed to Engage Preliterate Children in Educational Activities”. In: *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. CSCW '17. Portland, Oregon, USA: Association for Computing Machinery, pp. 491–504. URL: `https://doi.org/10.1145/2998181.2998231`.
Solove, Daniel J (2002). “Conceptualizing privacy”. In: *California law review*, pp. 1087–1155.
Souza, Jerffeson Teixeira de et al. (2020). “An Agent Program in an IoT System to Recommend Activities to Minimize Childhood Obesity Problems”. In: *Proceedings of the 35th Annual ACM Symposium on Applied Computing*. SAC '20. Brno, Czech Republic: Association for Computing Machinery, pp. 654–661. URL: `https://doi.org/10.1145/3341105.3373927`.
Souza, Paulo V. C. et al. (2019). “Using hybrid systems in the construction of expert systems in the identification of cognitive and motor problems in children and young people”. In: *2019 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, pp. 1–6.
Spaulding, Samuel and Huili Chen (2018). “A social robot system for modeling children's word pronunciation”. In: *autonomous agents and multi agent systems 2018*.
Spaulding, Samuel, Goren Gordon, and Cynthia Breazeal (2016). “Affect-aware student models for robot tutors”. In: *Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems*, pp. 864–872.
Spear, Hila J and Pamela Kulbok (2004). “Autonomy and adolescence: A concept analysis”. In: *Public Health Nursing* 21.2, pp. 144–152.
Spear, Hila J and Pamela A Kulbok (2001). “Adolescent health behaviors and related factors: A review”. In: *Public health nursing* 18.2, pp. 82–93.
Speer, Samantha et al. (2021). “MindfulNest: Strengthening Emotion Regulation with Tangible User Interfaces”. In: *Proceedings of the 2021 International Conference on Multimodal Interaction*, pp. 103–111.
Srinivas, Nisha et al. (2019). “Exploring Automatic Face Recognition on Match Performance and Gender Bias for Children”. In: *2019 IEEE Winter Applications of Computer Vision Workshops (WACVW)*, pp. 107–115.

Steinfeld, Nili (2021). “Parental mediation of adolescent Internet use: Combining strategies to promote awareness, autonomy and self-regulation in preparing youth for life on the web”. In: *Education and Information Technologies* 26.2, pp. 1897–1920.

Stoilova, Mariya, Sonia Livingstone, and Rishita Nandagiri (2020). “Digital by default: Children’s capacity to understand and manage online data and privacy”. In: *Media and Communication*.

Susser, Daniel, Beate Roessler, and Helen Nissenbaum (2019). “Technology, autonomy, and manipulation”. In: *Internet Policy Review* 8.2.

Sutton, Richard S and Andrew G Barto (2018). *Reinforcement learning: An introduction.* MIT press.

Tahir, Rashid et al. (2019). “Bringing the Kid back into YouTube Kids: Detecting Inappropriate Content on Video Streaming Platforms”. In: *2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, pp. 464–469.

The state of AI in 2020 (2020). URL: https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/global-survey-the-state-of-ai-in-2020.

Thida, Aye et al. (2020). “VOIS: The First Speech Therapy App Specifically Designed for Myanmar Hearing-Impaired Children”. In: *2020 23rd Conference of the Oriental COCOSDA International Committee for the Co-ordination and Standardisation of Speech Databases and Assessment Techniques (O-COCOSDA)*, pp. 151–154.

THORN (2021). *Responding to Online Threats: Minors’ Perspectives on Disclosing, Reporting, and Blocking.* URL: https:%20//info.thorn.org/hubfs/Research/Responding%5C%20to%5C%20Online%5C%20Threats_2021-Full-Report.pdf.

Thorson, Kjerstin et al. (2021). “Algorithmic inference, political interest, and exposure to news and politics on Facebook”. In: *Information, Communication & Society* 24.2, pp. 183–200.

TikTok (2023). *Privacy Policy.* URL: https://www.tiktok.com/legal/page/eea/privacy-policy/en.

Tomasello, Michael and Katharina Hamann (2012). “The 37th sir frederick bartlett lecture: Collaboration in young children”. In: *Quarterly Journal of Experimental Psychology* 65.1, pp. 1–12.

Tsiakas, Konstantinos et al. (2020). “BrainHood: Towards an Explainable Recommendation System for Self-Regulated Cognitive Training in Children”. In: *Proceedings of the 13th ACM International Conference on PErvasive Technologies Related to Assistive Environments.* PETRA ’20. Corfu, Greece: Association for Computing Machinery. URL: https://doi.org/10.1145/3389189.3398004.

UNESCO (2019). *Digital Kids: Asia-Pacific competence framework.* URL: https://www.gcedclearinghouse.org/sites/default/files/resources/190165eng.pdf.

UNICEF (2020). “Policy guidance on AI for children”. In.

— (2023). *Secondary education.* URL: https://data.unicef.org/topic/education/secondary-education/.

Unicef et al. (2012). “Children’s rights and business principles”. In: *UNICEF, Save the Children, UN Global Compact.*

University, De Montfort (2022). *DMU research suggests 10-year-olds lose sleep to check social media.* URL: https://www.dmu.ac.uk/about-dmu/news/2022/september/dmu-research-suggests-10-year-olds-lose-sleep-to-check-social-media.aspx#:~:text=Dr%20Shaw%20said%3A%20%E2%80%9CPrimary%20school,is%20directly%20affecting%20their%20sleep..

Van Brummelen, Jessica, Viktoriya Tabunshchyk, and Tommy Heng (2021). "“Alexa, Can I Program You?”: Student Perceptions of Conversational Artificial Intelligence Before and After Programming Alexa". In: *Interaction Design and Children*, pp. 305–313.

Van Mechelen, Maarten (2016). "Designing technologies for and with children: Theoretical reflections and a practical inquiry towards a co-design toolkit". In.

Veale, Michael and Frederik Zuiderveen Borgesius (2021). "Demystifying the Draft EU Artificial Intelligence Act—Analysing the good, the bad, and the unclear elements of the proposed approach". In: *Computer Law Review International* 22.4, pp. 97–112.

Vidal-Hall, Charlotte, Rosie Flewitt, and Dominic Wyse (2020). "Early childhood practitioner beliefs about digital media: integrating technology into a child-centred classroom environment". In: *European Early Childhood Education Research Journal* 28.2, pp. 167–181.

Wachter, Sandra (2020). "Affinity Profiling and Discrimination by Association in Online Behavioral Advertising". In: *Berkeley Tech. LJ* 35, p. 367.

Wang, Ge, Kaiwen Sun, et al. (2023). "Child-Centred AI Design: Definition, Operation, and Considerations". In: *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*, pp. 1–6.

Wang, Ge, Jun Zhao, et al. (2021). "Protection or punishment? relating the design space of parental control apps and perceptions about them to support parenting for online safety". In: *Proceedings of the ACM on Human-Computer Interaction* 5.CSCW2, pp. 1–26.

— (2022). "'Don't make assumptions about me!': Understanding Children's Perception of Datafication Online". In: *Proceedings of the ACM on Human-Computer Interaction* 6.CSCW2, pp. 1–24.

— (2023a). "'Treat Me as Your Friend, Not a Number in Your Database': Co-Designing with Children to Cope with Datafication Online". In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. CHI '23. Hamburg, Germany: Association for Computing Machinery. URL: https://doi.org/10.1145/3544548.3580933.

— (2023b). "12 Ways to Empower: Designing for Children's Digital Autonomy". In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. CHI '23. Hamburg, Germany: Association for Computing Machinery. URL: https://doi.org/10.1145/3544548.3580935.

Wanglavan, Penpitcha et al. (2019). "The Development of Attention Detection Model from Child Behavior for Robot-Assisted Autism Therapy". In: *2019 19th International Conference on Control, Automation and Systems (ICCAS)*, pp. 775–780.

Wegerif, Rupert and Neil Mercer (1996). "Computers and reasoning through talk in the classroom". In: *Language and education* 10.1, pp. 47–64.

Wertsch, James V (1984). "The zone of proximal development: Some conceptual issues". In: *New Directions for Child and Adolescent Development* 1984.23, pp. 7–18.

Westlund, Jacqueline M Kory and Cynthia Breazeal (2016). "Transparency, teleoperation, and children's understanding of social robots". In: *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, pp. 625–626.

What do national AI strategies say about children? (2020). URL: https://www.unicef.org/globalinsight/stories/what-do-national-ai-strategies-say-about-children.

What is automated individual decision-making and profiling? (2020a). URL: https://ico.org.uk/for-organisations/guide-to-data-protection/guide-tothe-general-data-protection-regulation-gdpr/automated-decision-makingand-profiling/what-is-automated-individual-decision-making-and-profiling/.

What is automated individual decision-making and profiling? (2020b). URL: https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/automated-decision-making-and-profiling/what-is-automated-individual-decision-making-and-profiling/.

What is personal data? (2021). URL: https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/key-definitions/what-is-personal-data/.

Willard, Aiyana K et al. (2019). "Explain this, explore that: A study of parent–child interaction in a children's museum". In: *Child Development* 90.5, e598–e617.

Williams, Meredydd, Jason RC Nurse, and Sadie Creese (2019). "(Smart) Watch Out! encouraging privacy-protective behavior through interactive games". In: *International Journal of Human-Computer Studies* 132, pp. 121–137.

Winpenny, Eleanor M, Theresa M Marteau, and Ellen Nolte (2014). "Exposure of children and adolescents to alcohol marketing on social media websites". In: *Alcohol and alcoholism* 49.2, pp. 154–159.

Wisniewski, Pamela et al. (2017). "Parental control vs. teen self-regulation: Is there a middle ground for mobile online safety?" In: *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, pp. 51–69.

Wohlin, Claes (2014). "Guidelines for snowballing in systematic literature studies and a replication in software engineering". In: *Proceedings of the 18th international conference on evaluation and assessment in software engineering*, pp. 1–10.

Wolff, Annika, Michel Wermelinger, and Marian Petre (2019). "Exploring design principles for data literacy activities to support children's inquiries from complex data". In: *International Journal of Human-Computer Studies* 129, pp. 41–54.

Wöllmer, Martin et al. (Aug. 2011). "Tandem Decoding of Children's Speech for Keyword Detection in a Child-Robot Interaction Scenario". In: *ACM Trans. Speech Lang. Process.* 7.4. URL: https://doi.org/10.1145/1998384.1998386.

Wood, Elizabeth (2007). "Reconceptualising Child-Centred Education: Contemporary Directions in Policy, Theory and Practice in Early Childhood." In: *FORUM: for promoting 3-19 comprehensive education*. Vol. 49. 1. ERIC, pp. 119–134.

Woodward, Julia, Shaghayegh Esmaeili, et al. (Nov. 2018). "Investigating Separation of Territories and Activity Roles in Children's Collaboration around Tabletops". In: *Proc. ACM Hum.-Comput. Interact.* 2.CSCW. URL: https://doi.org/10.1145/3274454.

Woodward, Julia, Zari McFadden, Nicole Shiver, Amir Ben-Hayon, et al. (2018). "Using co-design to examine how children conceptualize intelligent interfaces". In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 1–14.

Woodward, Julia, Zari McFadden, Nicole Shiver, Amir Ben-hayon, et al. (2018). "Using Co-Design to Examine How Children Conceptualize Intelligent Interfaces". In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. CHI '18. Montreal QC, Canada: Association for Computing Machinery, pp. 1–14. URL: https://doi.org/10.1145/3173574.3174149.

Worden, Kirsten J (2019). “Disengagement in the digital age: A Virtue ethical approach to epistemic sorting on social media”. In: *Moral Philosophy and Politics* 6.2, pp. 235–259.
Yao, Liang-Shuang et al. (2022). “Parental mediation moderates the association between social media exposure and tobacco and alcohol use: Differences between elementary and middle school students”. In: *Journal of Studies on Alcohol and Drugs* 83.2, pp. 267–275.
Yap, Christine Ee Ling and Jung-Joo Lee (2020). “’Phone apps know a lot about you!’ educating early adolescents about informational privacy through a phygital interactive book”. In: *Proceedings of the Interaction Design and Children Conference*, pp. 49–62.
Yeung, Karen (2020). “Recommendation of the council on artificial intelligence (oecd)”. In: *International Legal Materials* 59.1, pp. 27–34.
Yim, Jee Bin, Donghyeon Ko, and Woohun Lee (2021). “Romi: Screen Peripheral Companion for Mediating Children’s Healthy Screen Use”. In: *Interaction Design and Children*, pp. 456–460.
YouTube (2021). *Terms of Service.* URL: https://kids.youtube.com/t/terms.
Yuhana, Umi Laili et al. (2017). “Predicting Math performance of children with special needs based on serious game”. In: *2017 IEEE 5th International Conference on Serious Games and Applications for Health (SeGAH)*, pp. 1–5.
Yulina, Syefrida and Dewi Hajar (2017). “Kids Menu Care: An Application for Food Menu Scheduling with Caloric Balance”. In: *Proceedings of the 2017 International Conference on Information Technology.* ICIT 2017. Singapore, Singapore: Association for Computing Machinery, pp. 253–257. URL: https://doi.org/10.1145/3176653.3176732.
Zakaria, Nasriah et al. (2011). “Protecting privacy of children in social networking sites with rule-based privacy tool”. In: *8th International Conference on High-capacity Optical Networks and Emerging Technologies*, pp. 253–257.
Zaman, Bieke and Marije Nouwen (2016). “Parental controls: advice for parents, researchers and industry”. In: *EU Kids Online*, pp. 1–9.
Zaman, Bieke, Marije Nouwen, et al. (2016). “A qualitative inquiry into the contextualized parental mediation practices of young children’s digital media use at home”. In: *Journal of Broadcasting & Electronic Media* 60.1, pp. 1–22.
Zeniarja, Junta et al. (2018). “Search Engine for Kids with Document Filtering and Ranking Using Naive Bayes Classifier”. In: *2018 International Seminar on Application for Technology of Information and Communication*, pp. 560–564.
Zhang, Facun, Lijun Cui, and HuaiJun Wang (2017). “Research on Children’s Fall Detection by Characteristic Operator”. In: *Proceedings of the International Conference on Advances in Image Processing.* ICAIP 2017. Bangkok, Thailand: Association for Computing Machinery, pp. 1–5. URL: https://doi.org/10.1145/3133264.3133268.
Zhao, Jun, Blanche Duron, and Ge Wang (2022). “KOALA Hero: Inform Children of Privacy Risks of Mobile Apps”. In: *Interaction Design and Children.* IDC ’22. Braga, Portugal: Association for Computing Machinery, pp. 523–528. URL: https://doi.org/10.1145/3501712.3535278.
Zhao, Jun, Ge Wang, et al. (2019). “‘I Make up a Silly Name’: Understanding Children’s Perception of Privacy Risks Online”. In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems.* CHI ’19. Glasgow, Scotland Uk: Association

for Computing Machinery, pp. 1–13. URL:
https://doi.org/10.1145/3290605.3300336.

Zhao, Richard et al. (2019). "Knowledge assessment: game for assessment of symptoms of child physical abuse". In: *Proceedings of the 14th International Conference on the Foundations of Digital Games*, pp. 1–7.

Zheng, Zhi et al. (2014). "Autonomous robot-mediated imitation learning for children with autism". In: *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 2707–2712.

Zhu, Kening et al. (2016). "How different input and output modalities support coding as a problem-solving process for children". In: *Proceedings of the The 15th International Conference on Interaction Design and Children*, pp. 238–245.

Zuboff, Shoshana (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power: Barack Obama's books of 2019*. Profile books.

Appendices

A
CUREC Forms and Ethics Approval Letters

A.1 CUREC forms and ethics approval letters for Chapter 4

CENTRAL UNIVERSITY RESEARCH ETHICS COMMITTEE (CUREC)

CUREC 2 Form

Higher-risk research involving human participants and/or their data

SECTION A. RESEARCH DETAILS	
1. Full title of research	Children's perceptions of and desires for datafication online
2. Short title of research	
3. Principal Investigator (PI) / Student Supervisor	Ge Wang, Dphil in Computer Science
4. PI's training in research ethics and/or research integrity	
5. Student name and degree programme (if applicable)	Ge Wang, Dphil in Computer Science
6. Department/Institute name	Department of Computer Science
7. University email address	Ge.wang@cs.ox.ac.uk
8. University telephone number	01865 273 875
9. Funding Source (required for ethics team use)	COVID-19 Rebuilding Research (Oxford University Internal Grant)
10. State any conflicts of interest and explain how these will be addressed	None

SECTION B. RESEARCHERS	
1. **Researcher title and name**	Jun Zhao
2. **Department / Institute name or affiliation**	Department of Computer Science
3. **Role in Research**	DPhil co-supervisor
4. **Training in Research Ethics and/or research integrity**	Research integrity online training - 28 September 2020
5. **Researcher title and name**	Pro. Nigel Shadbolt
6. **Department / Institute name or affiliation**	Department of Computer Science
7. **Role in Research**	DPhil co-supervisor
8. **Training in Research Ethics and/or research integrity**	Extensive experience in HCI research

SECTION C. BASIC INFORMATION	
1. **Provide a brief lay summary of the aims and objectives of the research. This should cover the questions it will answer and any potential benefits.** **(Maximum 300 words)**	We will be investigating children's understanding and perceptions of datafication practices online. To start with, 1) we will be looking at their existing perceptions and their experience with data collection and recommendation practices on YouTube, and observing to what extent they may be aware of the personalisation aspect of their experiences on YouTube and how they perceive it. 2) We will explore how children perceive different design options that could motivate/influence children's awareness of such data collection/personalisation practices 3) The research would be concluded with open questions to identify what children would need for a data transparency tool.
2. **List all places where research will be conducted (including any other countries and online)**	The study will take place online, through university approved platforms (Microsoft Teams). The study participants will be residents of the UK.
3. **Anticipated research start date**	16 August 2021
4. **Anticipated research end date**	31 December 2021

(n.b. A maximum of 5 years approval can be granted)	
5. Please list any CUREC Approved Procedure(s) you will follow	AP25 Non-invasive research methods with children
6. Please list any CUREC Best Practice Guidance used to develop your research	BPG 05 Payments and incentives in research BPG 06 Internet-mediated research BPG 09 Data collection, protection and management BPG 10 Conducting research interviews
7. Please list any Professional Guidelines used	Association for Computing Machinery (ACM) Code of Ethics and Professional Conduct
8. Name of departmental / peer reviewer (if applicable)	N/A
9. Will you submit, or have you submitted, this research for ethical review or consideration elsewhere (e.g. local or collaborator's ethics committee, or other local approval)?	No

SECTION D. PARTICIPANTS (n.b. where there is no contact with human participants (in person or virtual) and no observation of them, but only use of data about them, please omit this section, and complete section I instead)	
1. Age range of participants	7-13
2. Are research participants classed as people whose ability to give free and informed consent is in question? e.g. those under 18, prisoners, or adults 'at risk'	Yes, children under 18.
3. Anticipated number of participants	30-40
4. How was the number of participants decided?	Remote interviews with children from two age groups (7-9, 10-13), with 15-20 from each group

5. Inclusion Criteria	Children between 7-13 who use YouTube platform regularly (minimum 2 hours per week)
6. Exclusion criteria	
7. Please mark 'X' against all planned recruitment methods Provide copies of all recruitment material for review	Poster advert ☒ Flyer ☒ Email circulation ☒ In-person approach ☒ Website ☒ Social media (e.g. twitter, Facebook) ☒ Snowball sampling (recruiting through contacts of existing participants) ☒ Newspapers ☐ Research recruitment sites (e.g. Prolific Academic, Amazon Turk) ☒ Existing departmental contacts or volunteer database ☐ Other (please specify below) ☐
8. How will potential participants be identified and approached?	We will use Prolific Academic as our main recruitment channel, and we will share flyers and poster avert via our project website. We will verify the suitability of the potential participants via email or over the phone. The criteria for recruitment is that children needs to be between 7 and 13, and have some experience of using YouTube (e.g. minimum 2 hours per week). Once the participants are verified via email/phone, we will ensure that both the parents and guardians have read and fully understood the information sheet and children have also read and fully understood the information sheet written in a language appropriate for them before agreeing on participating in the study.
9. Will informed consent be obtained from the research participants or their parents/ guardians? If not, please explain why not.	Yes, we will obtain consent from their parents and guardians, and assent from the children participants. Parents, guardians and children will be fully acknowledged that participation is voluntary and that they are free to withdraw without giving any reason. The deadline by which they can withdraw any information the child has contributed to the research is [*31 August 2021*].
10. For each activity or group of participants, explain how informed consent will be obtained from the participants themselves and/or their parents/guardians, if applicable. How will their consent be recorded?	Due to the covid situation, consent will be collected through a digital form, with parents/guardians and children reply in an email with the following sentence: I, _____, confirm that I have read and understood the details of the study, and have had the opportunity to ask questions and discuss the study with others. I have received satisfactory answers to my questions. I understand that the project has received ethics clearance through the University of Oxford's ethical approval process for research involving human participants, and I understand who will have access to the data, how it will be stored and what will happen to the data at the end of the study. I understand that participation is voluntary and that my

	child and I are free to withdraw at any time without giving any reason. I understand how to raise a concern or make a complaint. I, _____, I have read and understood the details of the study, and understand what this study is about. I have had the opportunity to ask questions that I want and my questions have been answered in a way that I understand. I understand that it's OK to stop taking part at any time. I am happy to take part and for my video and voice to be recorded.

SECTION E. RESEARCH METHODOLOGY

1. Please mark 'X' against the methods that will be used in your research

Ensure you address each method you will use in your informed consent documents and on this form

Use of casual or local workers (e.g. interpreters)	☐	Audio recording of participant	☒
Interview (refer to guidance in BPG 10: Conducting research interviews)	☒	Video recording of participant	☒
Focus group	☐	Photography of participant	☐
Participant completes questionnaire in hard copy	☐	Physiological recording from participant	☐
Participant completes online questionnaire or other online task (refer to guidance in BPG 06: Internet-mediated research)	☒	Taking a sample of blood or other bodily fluid from a participant	☐
Use of social media to recruit or interact with participants (refer to guidance in BPG 06: Internet-mediated research)	☒	Participant observation	☐
Analysis of existing records	☐	Covert observation	☐
Participant performs verbal or aural task	☐	Systematic observation	☐
Participant performs paper and pencil task	☐	Observation of specific organisational practices	☐
Participant performs computer based task	☐	Other (please specify below)	☐
Measurement/recording of motor behaviour	☐		

2. Provide a lay description of the research design and methods. In particular, describe clearly what participants in the research will be asked to do.

Through a short online survey (15 mins) on JISC online survey platform followed by a short online interview (25-30 mins), we would like to learn more about children's perceptions of the data collection and processing practices on the YouTube platform and seek design opportunities to improve children's experiences with YouTube. In the online survey, children participants will be provided with a list of (10-15) structured questions, expressed in a language appropriate for their age and development, asking information about their current usage and experience of the YouTube platform and how they perceive the recommended videos on the platform. In the online interview, we will present children with several (6-9) scenarios of data practices on, which are designed to make the data-based algorithmic recommendation more explicit for children in different formats, and provide different discussion points for them to think about how to control these data-driven. In this way, we will gain a deeper understanding about how children of different age groups (7-9; 10-13) are aware of the data-based algorithmic operations on the YouTube platforms and how this affects their viewing experiences. Parents and guardians are not required to be present in the study, as long as the children are competent with the technology and are comfortable with the arrangement.
3. Will the research include any audio, video or photographic recordings?
Yes. Video and audio recordings of the interviews with children will be recorded for the sake of data collection. No specific analysis will be performed on the video recordings, but they are collected as part of the recording function of the remote meeting platforms. However, the participants will always have the option to either turn off their camera, or a Teams "call" if they prefer. After the initial study warm-up, children participants will be offered an option to turn their videos off if that may make them more comfortable.
4. Please detail any expenses or gifts that will be offered to participants Guidance is available in Best Practice Guidance: 05 Payments and incentives in research.
A £15 Amazon gift card will go to each family, and stickers will be given to children.

SECTION F. ETHICAL CONSIDERATIONS For guidance on ethical issues, please see http://researchsupport.admin.ox.ac.uk/governance/ethics/resources *(N.B. To complete, double click on the check boxes and select 'checked')*		
1. Will the research involve any participants considered vulnerable in the context of the research (e.g. children, elderly, prisoners, adults "at risk")? **If yes,** please describe how they are defined as vulnerable and detail any CUREC Approved Procedures or guidance that will be applied to the research (for current documents and templates see https://researchsupport.admin.ox.ac.uk/governance/ethics/resources). In particular, please state why either CUREC Approved Procedure 15 or 25 cannot be applied wholly to your research.	Yes ☒	No ☐
Children under 18 will be involved. AP25 can be applied.		

2. **Will [unequal relationships] exist between participants and those obtaining informed consent?** **If yes,** describe the nature of the unequal relationship and how arising ethical issues will be addressed	Yes ☐	No ☒
3. **Will the research involve questions and/or discussions of contentious and/or sensitive issues (e.g. information relating to ethnicity, political opinions, religious beliefs, physical/mental health or sexual life)?** **If yes,** please justify why this is required and provide the questions (or an outline of them) raising the issues that will be used in your research.	Yes ☐	No ☒
4. **Will taking part in the research put participants under any particular burden and/or risk (including risk of prosecution)?** **If yes,** describe how risks will be mitigated. If there is a risk of prosecution to the participant, justify why incriminating data are sought	Yes ☐	No ☒
5. **Will the research involve deliberate [deception] of participants outside the scope of [CUREC Approved Procedure 07]?** **If yes**, justify why deception is used, describe deception and debriefing process, and include debriefing documents in the application	Yes ☐	No ☒
6. **Could the proposed research affect your own physical and/or psychological safety as a researcher?** **If yes,** describe how you will manage this.	Yes ☐	No ☒

7. How will you ensure the research is conducted according to the details given in this form?

All members involved in the research have received extensive research ethics training and have extensive experience of conducting research with human/children participants. Furthermore, we will run weekly project meetings to discuss project progress and address any potential concerns that could be raised during the study.

Risks to researchers:

Researchers will mainly carry out activities (such as interviews, surveys or setting up software tools) online, from their home/office environment. Researchers are not expected to physically interact with participants. This reduces risks to the researchers. Further, we will not study a risky or sensitive topic, and therefore we do not anticipate any expression of distress from the participants during researchers' interactions with them. To further minimise risks to the researchers, all research staff would read through the University's Safeguarding Code of Practice before conducting the research study. We will also carry out a risk assessment, following the attached risk assessment form, prior to the study.

Dr Zhao will be appointed as the "Designated Safeguarding Lead" of the study and work together with the University Safeguarding Officers to ensure all safeguarding processes are in place and the University's Code of Practice is thoroughly followed through. The following procedures will be carried out:

- The Safeguarding Lead will carry out a risk assessment according to the attached risk assessment form for all members of staff involved in the study, and identify the need for any pre-employment or pre-activity checks in accordance with the University's guidance.

- The lead will also ensure that all members of staff will be under appropriate supervision at all time. The likelihood of researchers working alone with young children will be minimised.

- Both the contact details of the University's Safeguarding Officer and the nominated Safeguarding Lead will be made available to the children at the start of the study, in case any complaints made against any member of the team arise. The procedure outlined in the University's Safeguarding Code of Practice (Section 4) will be followed through to ensure that issues are dealt with in a timely and thorough manner. Any allegations will be reported to the relevant University Safeguarding Officers ***without delay***. Any such allegations that may need onward referral to external agencies will be dealt with within one working day. The Designated Safeguarding Lead must not investigate the matter, and must refer as promptly as possible.

- The Designated Lead completed OSCB's (Oxfordshire Safeguarding Children Board) online training on Safeguarding Children in July 2017 (see certificate attached), and received an enhanced DBS check clearance in April 2018, which has been recently renewed.

Risks to the participants:

We do not anticipate any ethical, emotional, physical or political risks to our participants, given that our topic focuses on understanding technology usage patterns and education of technology usage. As outlined above, we have ensured that all research staff have received appropriate training, including 1) the use of appropriate research methods, 2) conduct relating to engagement with children, 3) recognition of and dealing with ethical issues, and 4) recognition of and dealing with situations where abuse and/ or serious risk is identified.

The questions asked in the studies do not concern personal matters, and thus are unlikely to cause participants undue stress or carry any other potential psychological risks.

The data to be collected will be kept on a secure server with strict access control in place. They will not be shared with any third parties or external organisations. Participants will have ample opportunities to request to delete data recordings at any time during the project. Once a study is completed, all recordings will be transcribed as soon as possible, with all names removed. The recordings which contain any personally identifiable information will be deleted from our secure desk immediately. All the anonymised transcriptions will be kept on our secure and encrypted hard drive, and any linkage to our pseudonymisation token will be kept accessible only to the lead researchers. The participants would have the right to look at the survey results and interview transcriptions of their own study if they want to, but not any other data collected throughout the study (e.g. transcriptions of others, data analysis).

Participants will be reminded that they can withdraw from the study at any time, and studies will be stopped immediately when young children show any sign of stress or unwillingness to cooperate.The deadline by which they can withdraw any information the child has contributed to the research is [*31 August 2021*].

8. Please give details of any other ethical and/or safety considerations

N/A

9. How do you propose to deal with / handle any incidental findings?

There is a possibility that during the study we discovered that children participants have been exposed to inappropriate content. The implications of such incidence can vary on a case by case basis, depending on the age of the children and the content. If our researcher recognises that such an incident may pose serious risks to children's safety and well-being, we would stop the interview immediately and have a group discussion, with the presence of our Safeguarding Lead, to decide the most appropriate response to the situation. We would always obtain the children's assent before raising our concerns to their parents/guardians or police forces.
10. Will any data or information from this study be provided to individual participants?
No, but participants will be provided with an option if they would like to be kept informed of the study results.

SECTION G. OTHER CONSIDERATIONS

1. Is any part of this research being conducted overseas? **If yes,** please give details below. Ensure you complete and submit a travel risk assessment form to your departmental safety officer, if your department requires this. (This is necessary to ensure the travel/fieldwork is covered by the University's travel insurance – see http://www.admin.ox.ac.uk/finance/insurance/travel/) Please also address any physical or psychological risks for Oxford researchers and local fieldworkers in the 'Ethical Considerations' section above and discuss these with your safety officer.	Yes ☐	No ☒
2. Does your research raise issues relevant to the Counter-Terrorism and Security Act (the Prevent Duty), which seeks to prevent people from being drawn into terrorism? **If yes,** please say how you plan to address any related risks. Please see advice on this on our Best Practice Guidance Web Page.	Yes ☐	No ☒

SECTION H. DATA MANAGEMENT AND HANDLING

All information provided by participants is considered **research data** for the purpose of this form. Any research data from which participants can be identified is known as **personal data**; any personal data which is sensitive is considered **special category data**.

Management of personal data, either directly or via a third party, must comply with the requirements of the General Data Protection Regulation (GDPR) and the Data Protection Act 2018, as set out in the University's Guidance on Data Protection and Research. In answering the questions below, please also consider the points raised in the Data Protection Checklist and whether, for higher-risk data processing, a separate Data Protection Impact Assessment (DPIA) may also be required for the research. Advice on research data management and security is available from Research Data Oxford and your local IT department. Advice on data protection is available from the Information Compliance team.

1. Please mark 'X' against the data you will collect for your research

Screening documents	☒	Audio recordings	☒
Consent records including participant name or other identifiers (e.g. written consent forms, audio-recorded consent, assent forms)	☒	Video recordings	☒
Consent obtained anonymously (e.g. via online survey)	☐	Transcript of audio/video recordings	☒
Opt-out forms	☐	Photographs	☐
Contact details for the purpose of this research only	☒	Information about the health of the participant (including mental health)	☐
Contact details for future use	☒	Physiological test results / measurements	☐
Field notes	☐	MRI scans	☐
Task results (e.g. questionnaires, diary completion)	☒	IP addresses (refer to Best Practice Guidance 09: Data collection, protection and management for guidance)	☐
Data already in the public domain. Specify the source of the data:	☐	Other (please specify below)	☐
Previously collected (secondary) data	☐		

2. How and where will each type of data be stored whilst the research is ongoing (until the end of all participant involvement)?

List each type of data selected above, and explain how each will be physically transferred (including movement/sharing of audio files, paper records, electronic downloads etc.) from where it is collected to a suitable storage site (e.g. Nexus365 OneDrive for Business, SharePoint, University servers). State the storage location for each. Do not store unencrypted data in freely available cloud services or unprotected USB drives.

Refer to Best Practice Guidance on data collection, protection and management (BPG09).

Screening documents, consent records, opt-out forms, contact details (for the purpose of this research & for future use) will be collected from participants online, and stored electronically in password-protected folders on the research team's OneDrive within the University network.

Survey data will be downloaded from the online survey provider (JISC online survey platform), and transferred electronically to storage in password-protected folders on the research team's OneDrive within the University network.

Video and audio recordings will be transferred from the recording device to be stored as password-protected files on folders on the research team's OneDrive within the University network. They will then be deleted from the original recording device. A researcher from the team will transcribe the audio recordings. The audio recording held by the researchers will then be deleted. The transcription will be stored as a Word file on encrypted OneDrive folders within the University network

3. Will you use a unique participant number on research data instead of participant name?

If yes, state whether or not you will retain a list of participant names against numbers (pseudonymisation via a linkage list). Where will the list be stored, and when will it be destroyed?

Yes, this linkage file will be kept in a separate encrypted folder on an encrypted hard drive, which will be kept in a locked cabinet in the PI's office.
4. Who will have access to the research data?
Researchers listed on this form will have access to the research data. Access will be granted to the MS IDREC for the purposes of monitoring and/or audit of the research.
5. If research data is to be shared with another organisation, how will it be transferred / disclosed securely?
n/a
6. When and how will identifiable data (including audio/video recordings & photos) be destroyed or deleted? N.B. If any identifiable data will be retained beyond the end of the study and/or indefinitely, please state what data this is, and the reasons for retention (e.g. contact details for future studies; photos used in publication). This must be clearly stated on participant information, and specific consent obtained.
Video and audio recordings will be transferred from the recording device to be stored as password-protected files on folders on the research team's OneDrive within the University network. They will then be deleted from the original recording device. A researcher from the team will transcribe the audio recordings. The audio recording held by the researchers will then be deleted as soon as possible (within a week). The transcription will be stored as a Word file on encrypted OneDrive folders within the University network

7. Please confirm that you will store other (non-identifiable) research data safely for at least 3 years after final publication or public release and adhere to any additional research funder policies. For more information about the University policies, please see the University's web pages on research data management. **If 'Yes'**, please give details of who will store the data and on storage format, location and security. **If 'No'**, please provide further details.	Yes ☒	No ☐
Data will be stored electronically in password-protected files on encrypted computers within the University network		

SECTION I. RESEARCH INVOLVING SECONDARY USE OR DISCLOSURE OF PERSONAL DATA OR SPECIAL CATEGORY DATA

This section of the form is **only to be completed** for research activity (as part or all of the research) where there is no contact with human participants (in person or virtual) and no observation of them, only use of data about them.

Your research must meet the standards laid down in the Data Protection Act 2018 with respect to the collection, use, and storage of personal data about human participants.

1. Will you seek data access agreements for these data? **If yes,** ● List the individual(s) or organisation(s) from which the information will be sourced ● Attach a copy of the agreement with the individual(s) or organisations in question ● Provide details of any conditions imposed by the organisation(s) concerning the release of the information **If no,** please explain how and when the agreement of the disclosing organisation(s) will be obtained	Yes ☐	No ☐
2. Could these data be linked back to an individual or individuals? **If yes,** ● Please explain why data cannot be collected in a way that prevents linkage with an individual/individuals ● Say how individual consent was obtained for the collection, use or disclosure of linkable data **If no,** you do not need to complete the rest of this section	Yes ☐	No ☐
3. How will any personally identifiable data be transferred to you? Please describe the arrangements for any physical transfer of personal data (including paper records and data captured electronically via portable media) from where you are obtaining it to local storage		
4. Where, and for how long, will personally identifiable data be stored during and after the research? Please outline procedures for ensuring confidentiality, e.g. security arrangements, pseudonymisation etc.		
5. Who will have access to the personally identifiable data? If data is to be shared with another organisation, other than the researchers listed, how will it be transferred / disclosed securely		
6. When and how will personally identifiable data be destroyed?		

SECTION J. PUBLICATION AND DISSEMINATION OF RESULTS

1. How will you disseminate project outcomes at the end of the research?	Research results will be disseminated as public reports, technical reports, as well as academic publications. Public reports will be written in a language accessible to the general public. We will ask all participants for their permission to use direct quotes, and all quotes will have all identifiable names and places removed. All research outcomes will be made available through the project's website and open access wherever possible. Study participants can also opt-in to receive updates of research publications directly at the end of the study.

SECTION K. ADDITIONAL QUESTIONS FOR APPLICATIONS TO THE MEDICAL SCIENCES IDREC	
1. List any standardised questionnaires that will be utilised (there is no need to send a copy)	
2. List any additional questionnaires designed by the researchers – a copy of these must be sent to the MS IDREC for review	
3. Give details of any biological sample(s) that will be taken (e.g. blood, urine, saliva, faeces)	

DECLARATIONS AND SIGNATURES

In providing signatures, the IDRECs accept either:

Option 1: Email confirmations sent from a University of Oxford email address. Separate emails should be sent by each of the relevant signatories as outlined below, indicating acceptance of their responsibilities.

Option 2: That the form be fully-signed with handwritten (wet-ink) signatures. Please scan these and the rest of the form pages to create a single PDF document and email to us.

The form should be sent with Word versions of all documents by email to:

ethics@medsci.ox.ac.uk (for applications from the Medical Sciences and MPLS divisions)

ethics@socsci.ox.ac.uk (for applications from the Social Sciences and Humanities divisions)

Applications from departments with a departmental research ethics committee (DREC) should first be sent for initial review to the relevant DREC.

Pasted images of signatures cannot be accepted

PRINCIPAL INVESTIGATOR (AND STUDENT IF APPLICABLE) I/We, the researcher(s): ● Understand our responsibilities as outlined on this form and in the CUREC glossary and guidance ● Agree to start this research only after obtaining approval from the IDREC; ● Understand that the Principal Investigator must ensure that all researchers are suitably qualified and trained to conduct the research described, or are appropriately supervised until deemed qualified/trained; ● Agree to provide additional information as requested by the IDREC before approval is secured and as research progresses; ● Agree to maintain the confidentiality of all data collected from or about participants; ● Agree to notify the IDREC in writing immediately of any proposed change to the research, and await approval before proceeding with the proposed change; ● Agree to notify the IDREC if the Principal Investigator changes and supply the name of the successor; ● Will use the data collected only for the research for which approval has been given; ● Will grant access to data only to authorised persons; and ● Have made arrangements to ensure that personal data collected from participants will be held in compliance with the requirements of the GDPR and the Data Protection Act 2018.	
Principal Investigator (Name)	
Principal Investigator (Signature) (Wet-ink signature, not pasted electronic image)	

Date	
Student (Name)	Ge Wang
Student (Signature) **(**Wet-ink signature, not pasted electronic image)	
Date	11-06-2021

ACCEPTANCE BY HEAD OF DEPARTMENT/FACULTY OR DESIGNATED NOMINEE*

*Another senior member of the department may sign where the head of department is the Principal Investigator, or where the head of department has appointed a nominee. Example nominees include Deputy Head of Department, Director of Research, and Director of Graduate/Undergraduate Studies.

I have read the research proposal above. On the basis of the information available to me, I:

- consider the Principal Investigator/Supervisor and student researcher (if applicable) to be aware of their ethical responsibilities in regard to this research;
- I am satisfied that the proposed design and scientific methodology are sound; the research has been subject to appropriate peer review and is likely to contribute to existing knowledge and/or to the education and training of the researcher(s) and that it is in the public interest.

Head of Department or designated nominee (Name)	
Head of Department or designated nominee (Signature) Wet-ink signature (not pasted electronic image) *or* The Head of Department/nominee can send an email (including PI name and study title) to ethics@medsci.ox.ac.uk or ethics@socsci.ox.ac.uk confirming the above	
Date	

RE: curec application

Computer Science Departmental Research Ethics Committee <ethics@cs.ox.ac.uk>
Fri 08/10/2021 23:19
To:Ge Wang <ge.wang@jesus.ox.ac.uk>

Dear Ge,

The above application has been considered by the Computer Science Departmental Research Ethics Committee (DREC), on behalf of the Social Sciences and Humanities Inter-divisional Research Ethics Committee (IDREC) in accordance with the procedures laid down by the University for ethical approval of all research involving human participants.

I am pleased to inform you that, on the basis of the information provided to the DREC, the proposed research has been judged as meeting appropriate ethical standards, and accordingly APPROVAL has been granted. Please use the reference number **CS_C1A_021_028** on all relevant documentation.

Should there be any subsequent changes to the project, which raise ethical issues not covered in the original application, you should submit details to the DREC for consideration.

Kind regards

Oliver

UNIVERSITY OF OXFORD

Dr. Oliver Sampson
DREC Secretary, Room 113
Research Support, Professional Service
Department of Computer Science
Wolfson Building, Parks Road
Oxford OX1 3QD
Tel: 01865 610761

A.2 CUREC forms and ethics approval letters for Chapter 5

CENTRAL UNIVERSITY RESEARCH ETHICS COMMITTEE (CUREC)

CUREC 2 Form

Higher-risk research involving human participants and/or their data

The University of Oxford places a high value on the knowledge, expertise, and integrity of its members and their ability to conduct research to high standards of scholarship and ethics. The research ethics clearance procedures have been established to ensure that the University is meeting its obligations as a responsible institution. They start from the presumption that all members of the University will take their responsibilities and obligations seriously, and will ensure that their research involving human participants is conducted according to the established principles and good practice in their field and in accordance, where appropriate, with legal requirements.

ONLY TYPE-WRITTEN FORMS WILL BE ACCEPTED

Before completing this application, please work through the guidance on our website to ensure that you do not need to submit a CUREC 1/1A (minimal-risk review) or CUREC 3 (full Committee review for certain applications to the MS IDREC) application instead.

SECTION A. RESEARCH DETAILS	
1. Full title of research	Designing for age-appropriate YouTube
2. Short title of research	YouTube Design
3. Principal Investigator (PI) / Student Supervisor	Ge Wang, Dphil in Computer Science
4. PI's training in research ethics and/or research integrity	Research integrity online training - 28 September 2020
5. Student name and degree programme (if applicable)	Ge Wang, Dphil in Computer Science
6. Department/Institute name	Department of Computer Science
7. University email address	Ge.wang@cs.ox.ac.uk
8. University telephone number	01865 273 875

9. Funding Source (required for ethics team use)	COVID-19 Rebuilding Research (Oxford University Internal Grant)
10. State any conflicts of interest and explain how these will be addressed	None

SECTION B. RESEARCHERS	
1. Researcher title and name	Jun Zhao
2. Department / Institute name or affiliation	Department of Computer Science
3. Role in Research	DPhil co-supervisor
4. Training in Research Ethics and/or research integrity	Research integrity online training - 28 September 2020
5. Researcher title and name	Pro. Nigel Shadbolt
6. Department / Institute name or affiliation	Department of Computer Science
7. Role in Research	DPhil co-supervisor
8. Training in Research Ethics and/or research integrity	Extensive experience in HCI research
9. Researcher title and name	Pro. Max Van Kleek
10. Department / Institute name or affiliation	Department of Computer Science
11. Role in Research	DPhil co-supervisor
12. Training in Research Ethics and/or research integrity	Extensive experience in HCI research

SECTION C. BASIC INFORMATION	
1. Provide a brief lay summary of the aims and objectives of the research. This should cover the questions it will answer and any potential benefits. **(Maximum 300 words)**	Following our study on children's perceptions of data practices on YouTube in November 2021 - February 2022, we will now be investigating their desires for a new YouTube probe. Based on our existing findings from the previous stage of the study, we will be looking at 1). How children perceive and react to different design options, and 2). What kind of control and transparency mechanism would be useful.
2. List all places where research will be conducted (including any other countries and online)	The study will take place in the Department of Computer Science, through university approved protocols. The study participants will be residents of the UK.
3. Anticipated research start date	15 February 2022
4. Anticipated research end date (n.b. A maximum of 5 years approval can be granted)	30 June 2022
5. Please list any CUREC Approved Procedure(s) you will follow	AP25 Non-invasive research methods with children
6. Please list any CUREC Best Practice Guidance used to develop your research	BPG 05 Payments and incentives in research BPG 09 Data collection, protection and management BPG 10 Conducting research interviews
7. Please list any Professional Guidelines used	Association for Computing Machinery (ACM) Code of Ethics and Professional Conduct
8. Name of departmental / peer reviewer (if applicable)	N/A
9. Will you submit, or have you submitted, this research for ethical review or consideration elsewhere (e.g. local or collaborator's ethics committee, or other local approval)?	No

Commented [OS1]: Please confirm that relevant risk assessments (including COVID safety) have been filed for holding workshops in the department

SECTION D. PARTICIPANTS (n.b. where there is no contact with human participants (in person or virtual) and no observation of them, but only use of data about them, please omit this section, and complete section I instead)		
1. Age range of participants	7-15 (we found that from 7 onwards, children will more capable of articulating their thoughts on more advanced topic such as data practices online)	
2. Are research participants classed as people whose ability to give free and informed consent is in question? e.g. those under 18, prisoners, or adults 'at risk'	Yes, children under 18.	
3. Anticipated number of participants	10-20	
4. How was the number of participants decided?	Design workshops with children from two age groups (7-10, 11-15), with 5-10 from each group	
5. Inclusion Criteria	Children between 7-15 who use YouTube platform regularly (minimum 2 hours per week)	
6. Exclusion criteria		
7. Please mark 'X' against all planned recruitment methods Provide copies of all recruitment material for review	Poster advert	☒
	Flyer	☒
	Email circulation	☒
	In-person approach	☒
	Website	☒
	Social media (e.g. twitter, Facebook)	☒
	Snowball sampling (recruiting through contacts of existing participants)	☒
	Newspapers	☐
	Research recruitment sites (e.g. Prolific Academic, Amazon Turk)	☒
	Existing departmental contacts or volunteer database	☐
	Other (please specify below)	☐
8. How will potential participants be identified and approached?	We will use Prolific Academic as our main recruitment channel, and we will share flyers and poster avert via our project website and Twitter account, as well as our existing pool of participants (from our previous project). We will also directly send emails to school head teachers. We will verify the suitability of the potential participants via email or over the phone. The criteria for recruitment is that children needs to be between 7 and 15, and have some experience of using YouTube (e.g. minimum 2 hours per week).	

Commented [OS2]: Why is this many participants required? Why are children segmented into these age groups?

Commented [JB3R2]: Please answer this question.

Commented [JB4]: Please could you answer this question for all the recruitment methods indicated in your answer to the previous question. For example, which social media platforms will you use?

Commented [OS5]: Also advertising directly to schools as per the attached letter to head teachers?

Commented [JB6R5]: Please could you answer this?

	Once the participants are verified via email/phone, we will ensure that both the parents and guardians have read and fully understood the information sheet and children have also read and fully understood the information sheet written in a language appropriate for them before agreeing on participating in the study.
9. Will informed consent be obtained from the research participants or their parents/ guardians? If not, please explain why not.	Yes, we will obtain consent from their parents and guardians, and assent from the children participants. Parents, guardians and children will be fully acknowledged that participation is voluntary and that they are free to withdraw without giving any reason. The deadline by which they can withdraw any information the child has contributed to the research is [*30 June 2022*].
10. For each activity or group of participants, explain how informed consent will be obtained from the participants themselves and/or their parents/guardians, if applicable. How will their consent be recorded?	Due to the covid situation, consent will be collected through a digital form, with parents/guardians and children reply in an email with the following sentence. The consent emails will be stored on an encrypted drive as detailed in section H2: I, ____, confirm that I have read and understood the details of the study, and have had the opportunity to ask questions and discuss the study with others. I have received satisfactory answers to my questions. I understand that the project has received ethics clearance through the University of Oxford's ethical approval process for research involving human participants, and I understand who will have access to the data, how it will be stored and what will happen to the data at the end of the study. I understand that participation is voluntary and that my child and I are free to withdraw at any time without giving any reason. I understand how to raise a concern or make a complaint. I, ____, I have read and understood the details of the study, and understand what this study is about. I have had the opportunity to ask questions that I want and my questions have been answered in a way that I understand. I understand that it's OK to stop taking part at any time. I am happy to take part and for my video and voice to be recorded.

Commented [OS7]: This sounds good

Commented [OS8]: Consent emails to be stored on an encrypted drive as detailed in section H2

SECTION E. RESEARCH METHODOLOGY

1. Please mark 'X' against the methods that will be used in your research

Ensure you address each method you will use in your informed consent documents and on this form

Use of casual or local workers (e.g. interpreters)	☐	Audio recording of participant	☒
Interview (refer to guidance in BPG 10: Conducting research interviews)	☐	Video recording of participant	☒
Focus group	☒	Photography of participant	☐
Participant completes questionnaire in hard copy	☐	Physiological recording from participant	☐
Participant completes online questionnaire or other online task (refer to guidance in BPG 06: Internet-mediated research)	☐	Taking a sample of blood or other bodily fluid from a participant	☐
Use of social media to recruit or interact with participants (refer to guidance in BPG 06: Internet-mediated research)	☒	Participant observation	☐

Analysis of existing records	☐	Covert observation	☐
Participant performs verbal or aural task	☒	Systematic observation	☐
Participant performs paper and pencil task	☒	Observation of specific organisational practices	☐
Participant performs computer based task	☐	Other (please specify below)	☐
Measurement/recording of motor behaviour	☐		

2. Provide a lay description of the research design and methods. In particular, describe clearly what participants in the research will be asked to do.

Through a series of design workshops (90 minutes) held in-person in the department of Computer Science, we would like to learn more about children's needs for a more age-appropriate YouTube-like video platform. In the workshop, children participants will team up and be instructed on completing two sets of design activities, set in a child-friendly manner for their age and development.

During the design workshop (see attached study design), we will first ask the children to finish a fictional inquiry session. We will create an original story describing how a cartoon character finds trouble in watching YouTube videos. After reading the story aloud as a group, children will be asked to sketch the screens of a new video-watching platform.

After children have completed the sketches, they will then participate in 3 different comicboarding exercises. Each comicboard will be designed with panel-style scaffolding, such that the comic is complete except for a single missing panel for children to fill in. The comicboards will extend the storyline of our fictional inquiry and bring in elements of planning that were not probed in the original story. When a child is ready, a facilitator will read the comicboard out loud with the child, along with a prompt at the bottom of the page explaining the missing panel. Children will sit with their age-defined group as they sketch but work individually.

Each group of children will be accompanied by at least two researchers. Parents and guardians are not required to be present in the study, as long as the children are comfortable with the arrangement.

3. Will the research include any audio, video or photographic recordings?

Yes. Audio recordings of the interviews with children will be recorded for the sake of data collection. Audio recordings were mandatory for participating in the study. Photographic recordings will be taken for the designs and comic boards children produced, but not on children themselves.

4. Please detail any expenses or gifts that will be offered to participants

Guidance is available in [Best Practice Guidance: 05 Payments and incentives in research](#).

A £15 Amazon gift card will go to each family, and stickers will be given to children.

Commented [OS9]: Held in-person in the department of Computer Science

Commented [OS10]: (see attached study design)

Commented [JB11]: If children are happy to take part but don't want to be recorded would that be ok or do they have to agree to be recorded in order to take part?

Commented [OS12]: This seems acceptable given that images will only be stored of the final product of the comicborading activity

Commented [OS13]: This fits with the expectations of the living wage for a 90 minute activity

SECTION F. ETHICAL CONSIDERATIONS

For guidance on ethical issues, please see http://researchsupport.admin.ox.ac.uk/governance/ethics/resources

(N.B. To complete, double click on the check boxes and select 'checked')

1. Will the research involve any participants considered vulnerable in the context of the research (e.g. children, elderly, prisoners, adults "at risk")? **If yes,** please describe how they are defined as vulnerable and detail any CUREC Approved Procedures or guidance that will be applied to the research (for current documents and templates see https://researchsupport.admin.ox.ac.uk/governance/ethics/resources). In particular, please state why either CUREC Approved Procedure 15 or 25 cannot be applied wholly to your research.	Yes ☒	No ☐
Children under 18 will be involved. AP25 can be applied.		
2. Will unequal relationships exist between participants and those obtaining informed consent? **If yes,** describe the nature of the unequal relationship and how arising ethical issues will be addressed	Yes ☒	No ☐
We will make sure with both the parents and children that they are free to stop at anytime before and during the study, and they have till 30 June 2022 to inform us if they do not want their data to be stored and used by this study.		
3. Will the research involve questions and/or discussions of contentious and/or sensitive issues (e.g. information relating to ethnicity, political opinions, religious beliefs, physical/mental health or sexual life)? **If yes,** please justify why this is required and provide the questions (or an outline of them) raising the issues that will be used in your research.	Yes ☐	No ☒
4. Will taking part in the research put participants under any particular burden and/or risk (including risk of prosecution)? **If yes,** describe how risks will be mitigated. If there is a risk of prosecution to the participant, justify why incriminating data are sought	Yes ☐	No ☒
5. Will the research involve deliberate deception of participants outside the scope of CUREC Approved Procedure 07? **If yes**, justify why deception is used, describe deception and debriefing process, and include debriefing documents in the application	Yes ☐	No ☒

Commented [JB14]: Shouldn't the answer to this question be yes? It would be helpful if you could add a sentence or two to explain how you'll make sure the participants realise they don't have to take part and how you'll make it easy to withdraw if they change their mind part way through.

6. **Could the proposed research affect your own physical and/or psychological safety as a researcher?** **If yes,** describe how you will manage this.	Yes ☐	No ☒

7. How will you ensure the research is conducted according to the details given in this form?

All members involved in the research have received extensive research ethics training and have extensive experience of conducting research with human/children participants. Furthermore, we will run weekly project meetings to discuss project progress and address any potential concerns that could be raised during the study.

Risks to researchers:

Researchers will carry out activities within the department of Computer Science, following pre-established protocols and risk assessments. This reduces risks to the researchers. Further, we will not study a risky or sensitive topic, and therefore we do not anticipate any expression of distress from the participants during researchers' interactions with them. To further minimise risks to the researchers, all research staff would read through the University's Safeguarding Code of Practice before conducting the research study. We will also carry out a risk assessment (including Covid-19 related risks), following the attached risk assessment form, prior to the study.

Commented [JB15]: Will this cover any Covid-19 related risks?

Dr Zhao will be appointed as the "Designated Safeguarding Lead" of the study and work together with the University Safeguarding Officers to ensure all safeguarding processes are in place and the University's Code of Practice is thoroughly followed through. The following procedures will be carried out:

- The Safeguarding Lead will carry out a risk assessment according to the attached risk assessment form for all members of staff involved in the study, and identify the need for any pre-employment or pre-activity checks in accordance with the University's guidance.

- The lead will also ensure that all members of staff will be under appropriate supervision at all time. The likelihood of researchers working alone with young children will be minimised.

- Both the contact details of the University's Safeguarding Officer and the nominated Safeguarding Lead will be made available to the children and parents/guardians at the start of the study, in case any complaints made against any member of the team arise. The procedure outlined in the University's Safeguarding Code of Practice (Section 4) will be followed through to ensure that issues are dealt with in a timely and thorough manner. Any allegations will be reported to the relevant University Safeguarding Officers ***without delay***. Any such allegations that may need onward referral to external agencies will be dealt with within one working day. The Designated Safeguarding Lead must not investigate the matter, and must refer as promptly as possible.

Commented [OS16]: And parents/guardians

- The Designated Lead completed OSCB's (Oxfordshire Safeguarding Children Board) online training on Safeguarding Children in July 2017 (see certificate attached), and received an enhanced DBS check clearance in April 2018, which has been recently renewed.

Risks to the participants:

We do not anticipate any ethical, emotional, physical (for Covid-19 related risk please see attached risk assessment form) or political risks to our participants, given that our topic focuses on understanding technology usage patterns and education of technology usage. As outlined above, we have ensured that all research staff have received appropriate training, including 1) the use of appropriate research methods, 2) conduct relating to engagement with children, 3) recognition of and dealing with ethical issues, and 4) recognition of and dealing with situations where abuse and/ or serious risk is identified.

Commented [JB17]: Please also consider how, if you conduct the research in person, you'll manage the Covid risks to yourself and your participants.

The questions asked in the studies do not concern personal matters, and thus are unlikely to cause participants undue stress or carry any other potential psychological risks.

The data to be collected will be kept on a secure server with strict access control in place. They will not be shared with any third parties or external organisations. Participants will have ample opportunities to request to delete data recordings at any time during the project. Once a study is completed, all recordings will be transcribed as soon as possible, with all names removed. The recordings which contain any personally identifiable information will be deleted from our secure desk immediately. All the anonymised transcriptions will be kept on our secure and encrypted hard drive, and any linkage to our pseudonymisation token will be kept accessible only to the lead researchers. The participants would have the right to look at the results and transcriptions of their own study if they want to, but not any other data collected throughout the study (e.g. transcriptions of others, data analysis).

Participants will be reminded that they can withdraw from the study at any time, and studies will be stopped immediately when young children show any sign of stress or unwillingness to cooperate.The deadline by which they can withdraw any information the child has contributed to the research is [*31 June 2022*].

8. Please give details of any other ethical and/or safety considerations

N/A

9. How do you propose to deal with / handle any incidental findings?

There is a possibility that during the study we discovered that children participants have been exposed to inappropriate content. The implications of such incidence can vary on a case by case basis, depending on the age of the children and the content. If our researcher recognises that such an incident may pose serious risks to children's safety and well-being, we would stop the interview immediately and have a group discussion, with the presence of our Safeguarding Lead, to decide the most appropriate response to the situation. We would always obtain the children's assent before raising our concerns to their parents/guardians or police forces.

10. Will any data or information from this study be provided to individual participants?

No, but participants will be provided with an option if they would like to be kept informed of the study results.

SECTION G. OTHER CONSIDERATIONS

1. Is any part of this research being conducted overseas? **If yes,** please give details below. Ensure you complete and submit a travel risk assessment form to your departmental safety officer, if your department requires this. (This is necessary to ensure the travel/fieldwork is covered by the University's travel insurance – see http://www.admin.ox.ac.uk/finance/insurance/travel/) Please also address any physical or psychological risks for Oxford researchers and local fieldworkers in the 'Ethical Considerations' section above and discuss these with your safety officer.	Yes ☐	No ☒
2. Does your research raise issues relevant to the Counter-Terrorism and Security Act (the Prevent Duty), which seeks to prevent people from being drawn into terrorism? **If yes,** please say how you plan to address any related risks. Please see advice on this on our Best Practice Guidance Web Page.	Yes ☐	No ☒

SECTION H. DATA MANAGEMENT AND HANDLING

All information provided by participants is considered **research data** for the purpose of this form. Any research data from which participants can be identified is known as **personal data**; any personal data which is sensitive is considered **special category data**.

Management of personal data, either directly or via a third party, must comply with the requirements of the General Data Protection Regulation (GDPR) and the Data Protection Act 2018, as set out in the University's Guidance on Data Protection and Research. In answering the questions below, please also consider the points raised in the Data Protection Checklist and whether, for higher-risk data processing, a separate Data Protection Impact Assessment (DPIA) may also be required for the research. Advice on research data management and security is available from Research Data Oxford and your local IT department. Advice on data protection is available from the Information Compliance team.

1. Please mark 'X' against the data you will collect for your research

Screening documents	☐	Audio recordings	☒
Consent records including participant name or other identifiers (e.g. written consent forms, audio-recorded consent, assent forms)	☒	Video recordings	☒
Consent obtained anonymously (e.g. via online survey)	☐	Transcript of audio/video recordings	☒

Commented [JB18]: What will these involve?

Opt-out forms	☐	Photographs	☒
Contact details for the purpose of this research only	☒	Information about the health of the participant (including mental health)	☐
Contact details for future use	☒	Physiological test results / measurements	☐
Field notes	☐	MRI scans	☐
Task results (e.g. questionnaires, diary completion)	☐	IP addresses (refer to Best Practice Guidance 09: Data collection, protection and management for guidance)	☐
Data already in the public domain. Specify the source of the data:	☐	Other (please specify below)	☐
Previously collected (secondary) data	☐		

2. How and where will each type of data be stored whilst the research is ongoing (until the end of all participant involvement)?

List each type of data selected above, and explain how each will be physically transferred (including movement/sharing of audio files, paper records, electronic downloads etc.) from where it is collected to a suitable storage site (e.g. Nexus365 OneDrive for Business, SharePoint, University servers). State the storage location for each. Do not store unencrypted data in freely available cloud services or unprotected USB drives.

Refer to Best Practice Guidance on data collection, protection and management (BPG09).

Consent records, and contact details (for the purpose of this research & for future use) will be collected from participants online, and stored electronically in password-protected folders on the research team's OneDrive within the University network.

Audio recordings, as well as photographs of children's drawings will be transferred from the recording device to be stored as password-protected files on folders on the research team's OneDrive within the University network. They will then be deleted from the original recording device (encrypted audio recorder with password protection, and encrypted phone). A researcher from the team will transcribe the audio recordings. The audio recording held by the researchers will then be deleted. The transcription will be stored as a Word file on encrypted OneDrive folders within the University network

Commented [JB19]: What will these involve?

Commented [JB20]: Will you be collecting these? If not please delete this.

Commented [JB21]: How will interviews be recorded? Will you use the built-in recording function, or a recording device? If you're using a separate device, please make sure it's encrypted and password-protected if possible, and isn't automatically backing up recordings to the cloud.

3. Will you use a unique participant number on research data instead of participant name?

If yes, state whether or not you will retain a list of participant names against numbers (pseudonymisation via a linkage list). Where will the list be stored, and when will it be destroyed?

Yes, this linkage file will be kept in a separate encrypted folder on an encrypted hard drive, which will be kept in a locked cabinet in the PI's office.

4. Who will have access to the research data?

Researchers listed on this form will have access to the research data.

5. If research data is to be shared with another organisation, how will it be transferred / disclosed securely?

n/a

6. When and how will identifiable data (including audio/video recordings & photos) be destroyed or deleted?

N.B. If any identifiable data will be retained beyond the end of the study and/or indefinitely, please state what data this is, and the reasons for retention (e.g. contact details for future studies; photos used in publication). This must be clearly stated on participant information, and specific consent obtained.

Video and audio recordings will be transferred from the recording device to be stored as password-protected files on folders on the research team's OneDrive within the University network. They will then be deleted from the original recording device.

A researcher from the team will transcribe the audio recordings. The audio recording held by the researchers will then be deleted as soon as possible (within a week). The transcription will be stored as a Word file on encrypted OneDrive folders within the University network

7. Please confirm that you will store other (non-identifiable) research data safely for at least 3 years after final publication or public release and adhere to any additional research funder policies. For more information about the University policies, please see the University's web pages on research data management. **If 'Yes'**, please give details of who will store the data and on storage format, location and security. **If 'No'**, please provide further details.	Yes ☒	No ☐
Data will be stored electronically in password-protected files on encrypted computers within the University network		

SECTION J. PUBLICATION AND DISSEMINATION OF RESULTS	
1. How will you disseminate project outcomes at the end of the research?	Research results will be disseminated as public reports, technical reports, as well as academic publications. Public reports will be written in a language accessible to the general public. We will ask all participants for their permission to use direct quotes, and all quotes will have all identifiable names and places removed. All research outcomes will be made available through the project's website and open access wherever possible. Study participants can also opt-in to receive updates of research publications directly at the end of the study.

DECLARATIONS AND SIGNATURES

In providing signatures, the IDRECs accept either:

Option 1: Email confirmations sent from a University of Oxford email address. Separate emails should be sent by each of the relevant signatories as outlined below, indicating acceptance of their responsibilities.

Option 2: That the form be fully-signed with handwritten (wet-ink) signatures. Please scan these and the rest of the form pages to create a single PDF document and email to us.

The form should be sent with Word versions of all documents by email to:

ethics@medsci.ox.ac.uk (for applications from the Medical Sciences and MPLS divisions)
ethics@socsci.ox.ac.uk (for applications from the Social Sciences and Humanities divisions)

Applications from departments with a departmental research ethics committee (DREC) should first be sent for initial review to the relevant DREC.

Pasted images of signatures cannot be accepted

PRINCIPAL INVESTIGATOR (AND STUDENT IF APPLICABLE)

I/We, the researcher(s):

- Understand our responsibilities as outlined on this form and in the CUREC glossary and guidance
- Agree to start this research only after obtaining approval from the IDREC;
- Understand that the Principal Investigator must ensure that all researchers are suitably qualified and trained to conduct the research described, or are appropriately supervised until deemed qualified/trained;
- Agree to provide additional information as requested by the IDREC before approval is secured and as research progresses;
- Agree to maintain the confidentiality of all data collected from or about participants;
- Agree to notify the IDREC in writing immediately of any proposed change to the research, and await approval before proceeding with the proposed change;
- Agree to notify the IDREC if the Principal Investigator changes and supply the name of the successor;
- Will use the data collected only for the research for which approval has been given;
- Will grant access to data only to authorised persons; and
- Have made arrangements to ensure that personal data collected from participants will be held in compliance with the requirements of the GDPR and the Data Protection Act 2018.

Principal Investigator (Name)	
Principal Investigator (Signature) **(**Wet-ink signature, not pasted electronic image)	
Date	
Student (Name)	Ge Wang
Student (Signature) **(**Wet-ink signature, not pasted electronic image)	
Date	26-11-2021

ACCEPTANCE BY HEAD OF DEPARTMENT/FACULTY OR DESIGNATED NOMINEE*

*Another senior member of the department may sign where the head of department is the Principal Investigator, or where the head of department has appointed a nominee. Example nominees include Deputy Head of Department, Director of Research, and Director of Graduate/Undergraduate Studies.

I have read the research proposal above. On the basis of the information available to me, I:

- consider the Principal Investigator/Supervisor and student researcher (if applicable) to be aware of their ethical responsibilities in regard to this research;
- I am satisfied that the proposed design and scientific methodology are sound; the research has been subject to appropriate peer review and is likely to contribute to existing knowledge and/or to the education and training of the researcher(s) and that it is in the public interest.

Head of Department or designated nominee (Name)	Andrew Martin (DREC Chair)
Head of Department or designated nominee (Signature) Wet-ink signature (not pasted electronic image) *or* The Head of Department/nominee can send an email (including PI name and study title) to ethics@medsci.ox.ac.uk or ethics@socsci.ox.ac.uk confirming the above	Confirmed by email
Date	16^{th} Dec 2021

SOCIAL SCIENCES & HUMANITIES
INTERDIVISIONAL RESEARCH ETHICS COMMITTEE

Research Governance, Ethics & Assurance Team, Research Services, University of Oxford,
Boundary Brook House, Churchill Drive, Headington, Oxford OX3 7GB, UK
Tel: +44(0)1865 289881 Email: ethics@socsci.ox.ac.uk

Ge Wang
Department of Computer Science
University of Oxford

Friday 25 February 2022

Dear Ge,

Research ethics approval

Research title: Designing for age-appropriate YouTube

Research ethics reference: R76800/RE002

The above application has been considered by the Social Sciences and Humanities Interdivisional Research Ethics Committee (SSH IDREC) in accordance with the University's procedures for ethical approval of all research involving human participants.

I am pleased to confirm that, on the basis of the information provided to the IDREC, ethics approval has now been granted for this study.

Please note the following:

Personal data: It is the responsibility of the PI to ensure that all personal data collected during the project is managed in accordance with the University's guidance and legal requirements.

In-person activities: Any data collection involving in-person interactions with participants must have an up-to-date Covid-19 fieldwork risk assessment in place; further guidance is available from the Safety Office's website.

Amendments: Please notify the committee if you intend to make any amendments to the information in your ethics application as submitted at date of this approval, as all changes must receive ethical approval prior to implementation. The amendment form is available on the SSH IDREC webpage.

Adverse events: The SSH IDRECs must be notified within seven days of any unexpected adverse consequences to the research participants or other people involved in this research project.

Audit: The SSH IDREC audits a sample of projects each year to enable the Committee to monitor the ethical aspects of research in progress.

We welcome feedback on your experience of the ethical review process and suggestions for improvement. Please email any comments to ethics@socsci.ox.ac.uk.

Yours sincerely

DocuSigned by:
Alison Monk
162CBC83B879410...

Alison Monk
Research Ethics Administrator

for

Jennifer Blaikie
Research Ethics Manager (SSH IDREC)

cc: Dr. Jun Zhao

A.3 CUREC forms and ethics approval letters for Chapter 6

Central University Research Ethics Committee (CUREC)

CUREC 2 Form

Higher-risk research involving human participants and/or their data

The University of Oxford places a high value on the knowledge, expertise, and integrity of its members and their ability to conduct research to high standards of scholarship and ethics. The research ethics clearance procedures have been established to ensure that the University is meeting its obligations as a responsible institution. They start from the presumption that all members of the University will take their responsibilities and obligations seriously, and will ensure that their research involving human participants is conducted according to the established principles and good practice in their field and in accordance, where appropriate, with legal requirements.

Before completing this application, please work through the guidance on our website to ensure that you do not need to submit a CUREC 1/1A (minimal-risk review) or CUREC 3 (full Committee review for certain applications to the MS IDREC) application instead. Only type-written forms will be accepted.

Advisory text is highlighted in yellow and should be deleted before finalising the document.

Section A. Research Details	
1. Full title of research	A Field study of KOALA Hero for assisting family data privacy literacy development
2. Short title of research	A Field Study of KOALA Hero App
3. Principal Investigator (PI) / Student Supervisor	Tiffany (Ge) Wang, 3rd CS DPhil student
4. PI's training in research ethics and research integrity Information about online training	Research Integrity Online Training - 28 September 2020
5. PI - date of completion of Information Security training	
6. Student name and degree programme (if applicable)	Tiffany (Ge) Wang, 3rd CS DPhil student
7. Department/Institute name	Department of Computer Science
8. University email address	Jun.zhao@cs.ox.ac.uk; ge.wang@jesus.ox.ac.uk
9. University telephone number	01865 273 875

10. Funding Source (required for ethics team use)	COVID-19 Rebuilding Research (Oxford University Internal Grant)
11. State any conflicts of interest and explain how these will be addressed	None

Section B. Researchers

Copy and paste the below 4 rows as necessary to complete for each member of the research team, including student(s) named above, then delete this entire row. Note that **the PI does not need to be entered again** in this section.

1. Researcher title and name	Dr Jun Zhao
2. Department / Institute name or affiliation	Department of Computer Science
3. Role in Research	Dphil Co-supervisor
4. Training in Research Ethics and/or research integrity Information about online training	Dr Zhao received her research ethics training at Oxford University in 2016 and completed her refresher course in September 2022. She also completed OSCB's online training on "Safeguarding Children" in July 2017.
5. Date of completion of Information Security training	September 2022
6. Researcher title and name	Konrad Kollnig / Adrien Zier / Blanche Duron / Zhilin Zhang
7. Department / Institute name or affiliation	Department of Computer Science
8. Role in Research	Student researcher / Summer Intern / Summer Intern / Student researcher, helping with tracker app testing at pilot study.
9. Training in Research Ethics and/or research integrity 10. Information about online training	Research ethics training at Oxford University.
11. Date of completion of Information Security training	September 2021

Section C. Basic information	
1. Provide a brief lay summary of the aims and objectives of the research. This should cover the questions it will answer, any potential benefits and what you will do to address the question. **(Maximum 300 words)**	Tablet computers are becoming the primary means for young children to go online. Although our research has shown how much children aged between 6 and 10 care about their personal privacy online, there is little technological support for families to have a transparent view of the data privacy risks associated with mobile devices and make informed decisions about them. Based on several years of research, we have designed and implemented an Android mobile app that can help families to find safer apps and discuss app safety together. We would like to invite 10 local families to install our KOALA Hero mobile app on their home tablet devices, so that we can assess the extent to which the technologies are effective in helping families make more informed choices of apps for their child. The study includes a pre-study survey to the participants, a diary study carried out by the participants at their home, and an exit (hybrid) survey and interview for parents and children.
2. List all places where research will be conducted (including any other countries and online)	University department, for initial setup on participants' devices; and participants' home, which researchers will not visit.
3. Anticipated research start date	5 December 2022
4. Anticipated research end date (n.b. A maximum of 5 years approval can be granted)	30 September 2023
5. Please list any CUREC Approved Procedure(s) you will follow	AP25 Non-invasive research methods with children
6. Please list any CUREC Best Practice Guidance used to develop your research	BPG 05 Payments and incentives in research BPG 09 Data collection, protection and management BPG 10 Conducting research interviews
7. Please list any Professional Guidelines used	Association for Computing Machinery (ACM) Code of Ethics and Professional Conduct
8. Name of departmental / peer reviewer (if applicable)	n/a
9. Will you submit, or have you submitted, this research for ethical review or consideration elsewhere (e.g. local or collaborator's ethics committee, or other local	n/a

approval)?	

Section D. Participants

(n.b. where there is no contact with human participants (in person or virtual) and no observation of them, but only use of data about them, please omit this section, and complete section I instead)

1. Age range of participants	7-11
2. Are research participants people who may not be able to give free and informed consent? e.g. those under 18, prisoners, or adults 'at risk'	Yes, children under 18.
3. Anticipated number of participants	10 children and their parents
4. How was the number of participants decided?	The number of participants should be sufficient to provide the depth and width of qualitative observations we need
5. Inclusion Criteria	Families with a primary child participant aged 7-11, with an Android device and fluent in English
6. Exclusion criteria	Families do not use English at home or primarily use Apple devices as our technology can only work on iPads

7. Please mark 'X' against all planned recruitment methods Provide copies of all recruitment material for review	Poster advert	☒
	Flyer	☒
	Email circulation	☒
	In-person approach	☒
	Website	☒
	Social media (e.g. twitter, Facebook)	☒
	Snowball sampling (recruiting through contacts of existing participants)	☒
	Newspapers	☐
	Research recruitment sites (e.g. Prolific Academic, Amazon Turk)	☒
	Existing departmental contacts or volunteer database	☒
	Other (please specify below)	☐

8. How will potential participants be identified and approached?	We will use Prolific Academic as our main recruitment channel, and we will share flyers and poster avert via our project website and Twitter account, as well as our existing pool of participants (from our previous project). We will also directly send emails to school head teachers. We will verify the suitability of the potential participants via email or over the phone. The criteria for recruitment are that children needs to be between 7 and 11, and use Android devices.

	Once the participants are verified via email/phone, we will ensure that both the parents and guardians have read and fully understood the information sheet and children have also read and fully understood the information sheet written in a language appropriate for them before agreeing on participating in the study.
9. Will informed consent be obtained from the research participants or their parents/ guardians? If not, please explain why not.	Yes, we will obtain consent from their parents and guardians, and assent from the child participants. Parents, guardians and children will be fully acknowledged that participation is voluntary and that they are free to withdraw without giving any reason. The deadline by which they can withdraw any information the child has contributed to the research is [*30 September 2023*].
10. For each activity or group of participants, explain how [informed consent] will be obtained from the participants themselves and/or their parents/guardians, if applicable. How will their consent be recorded?	Consent will be collected through a digital form, with parents/guardians and children reply in an email with the following sentence. The consent emails will be stored on an encrypted drive as detailed in section H2: I, _____, confirm that I have read and understood the details of the study, and have had the opportunity to ask questions and discuss the study with others. I have received satisfactory answers to my questions. I understand that the project has received ethics clearance through the University of Oxford's ethical approval process for research involving human participants, and I understand who will have access to the data, how it will be stored and what will happen to the data at the end of the study. I understand that participation is voluntary and that my child and I are free to withdraw at any time without giving any reason. I understand how to raise a concern or make a complaint. I, _____, I have read and understood the details of the study, and understand what this study is about. I have had the opportunity to ask questions that I want and my questions have been answered in a way that I understand. I understand that it's OK to stop taking part at any time. I am happy to take part and for my voice to be recorded.

Section E. Research Methodology

1. Please mark 'X' against the methods that will be used in your research

Ensure you address each method you will use in your informed consent documents and on this form

Use of casual or local workers (e.g. interpreters)	☐	[Audio recording] of participant	☒
Interview (refer to guidance in [BPG 10: Conducting research interviews])	☒	[Video recording] of participant	☐
Focus group	☐	Photography of participant	☐
Participant completes questionnaire in hard copy	☐	Physiological recording from participant	☐
Participant completes online questionnaire or other online task (refer to guidance in [BPG 06: Internet-mediated research])	☒	Taking a sample of blood or other bodily fluid from a participant	☐
Use of social media to recruit or interact with participants (refer to guidance in [BPG 06: Internet-mediated research])	☒	Participant observation	☐
Analysis of existing records	☐	Covert observation	☐

Participant performs verbal or aural task	☐	Systematic observation	☐
Participant performs paper and pencil task	☐	Observation of specific organisational practices	☐
Participant performs computer based task	☒	Other (please specify below)	☐
Measurement/recording of motor behaviour	☐		

2. Provide a lay description of the research design and methods. In particular, describe clearly what participants in the research will be asked to do.

Participants will be involved in the following activities by joining in the study:

1. Pre-survey and study setup (20 mins): which can take place either in our research lab on Parks Road (Oxford) or via Teams. During this process, participants (parents and children) will be invited to complete a short survey (see attached), which covers topics about their family's demographics, their child's usage of the device, their approach to mediating their child's choice of apps.
2. Activities with the Koala Hero Toolkit (40 mins): Each participant family is expected to engage with the Koala Hero toolkit. They will be asked to complete tasksheets and interact with the data cards and make joint configurations on the tracker app.
3. Exit interview (30 mins): At the end of the study, parents and children will be for a 30-min interview about their experience of using KOALA Hero. The interview will include questions such as what they liked and disliked about the app, and whether they will continue using the app after the study ended, etc. They can opt to have the experiment app and all local data to be deleted from their home devices.

3. Will the research include any audio, video or photographic recordings?

Yes. Audio recordings of the interviews and home conversations while using the KOALA app will be recorded for the sake of data collection.

4. Please detail any expenses or gifts that will be offered to participants

Guidance is available in Best Practice Guidance: 05 Payments and incentives in research.

A £20 Amazon gift card will go to each family.

Section F. Ethical Considerations

For guidance on ethical issues, please see http://researchsupport.admin.ox.ac.uk/governance/ethics/resources

(N.B. To complete, double click on the check boxes and select 'checked')

1. Will the research involve any participants considered vulnerable in the context of the research (e.g. children, elderly, prisoners, adults "at risk")? **If yes,** please describe how they are defined as vulnerable and detail any CUREC Approved Procedures or guidance that will be applied to the research (for current documents and templates see https://researchsupport.admin.ox.ac.uk/governance/ethics/resources). For research involving children, please state why either CUREC Approved Procedure 15 or 25 cannot be applied wholly to your research.	Yes ☒	No ☐

We will involve child participants aged under 18. We cannot follow AP15 or 25 because part of the study will be taking place in the participants' home environment as a diary study, although no researchers will need to visit

these private homes.		
2. Will unequal relationships exist between participants and those obtaining informed consent? **If yes,** describe the nature of the unequal relationship and how arising ethical issues will be addressed	Yes ☒	No ☐
We will make sure with both the parents and children that they are free to stop at anytime before and during the study, and they have till 30 September 2023 to inform us if they do not want their data to be stored and used by this study.		
3. Will the research involve questions and/or discussions of contentious and/or sensitive issues (e.g. information relating to ethnicity, political opinions, religious beliefs, physical/mental health or sexual life)? **If yes,** please justify why this is required and provide the questions (or an outline of them) raising the issues that will be used in your research.	Yes ☐	No ☒
4. Will taking part in the research put participants under any particular burden and/or risk (including risk of prosecution)? **If yes,** describe how risks will be mitigated. If there is a risk of prosecution to the participant, justify why incriminating data are sought. During the consent process, participants should be made aware of the risks of disclosing potentially illegal information and understand what the researchers would do if they were to receive that information.	Yes ☐	No ☒
5. Will the research involve deliberate deception of participants beyond that covered by CUREC Approved Procedure 07? **If yes**, justify why deception is used, describe deception and debriefing process, and include debriefing documents in the application	Yes ☐	No ☒
6. Could the proposed research affect your own physical and/or psychological safety as a researcher? **If yes,** describe how you will manage this. Explain what safety procedures, structured mentoring or other ongoing support will be in place during this research. Include details of lone working procedures, if applicable.	Yes ☐	No ☒
7. How will you ensure the research is conducted according to the details given in this form?		

All members involved in the research have received extensive research ethics training and have extensive experience of conducting research with human/children participants. Furthermore, we will run weekly project meetings to discuss project progress and address any potential concerns that could be raised during the study.

Risks to researchers:

Researchers will carry out activities within the department of Computer Science, following pre-established protocols and risk assessments. This reduces risks to the researchers. Further, we will not study a risky or sensitive topic, and therefore we do not anticipate any expression of distress from the participants during researchers' interactions with them. To further minimise risks to the researchers, all research staff would read through the University's Safeguarding Code of Practice before conducting the research study. We will also carry out a risk assessment (including Covid-19 related risks), following the attached risk assessment form, prior to the study.

Dr Zhao will be appointed as the "Designated Safeguarding Lead" of the study and work together with the University
Safeguarding Officers to ensure all safeguarding processes are in place and the University's Code of Practice is thoroughly followed through. The following procedures will be carried out:

- The Safeguarding Lead will carry out a risk assessment according to the attached risk assessment form for all members of staff involved in the study, and identify the need for any pre-employment or pre-activity checks in accordance with the University's guidance.

- The lead will also ensure that all members of staff will be under appropriate supervision at all time. The likelihood of researchers working alone with young children will be minimised.

- Both the contact details of the University's Safeguarding Officer and the nominated Safeguarding Lead will be made available to the children and parents/guardians at the start of the study, in case any complaints made against any member of the team arise. The procedure outlined in the University's Safeguarding Code of Practice (Section 4) will be followed through to ensure that issues are dealt with in a timely and thorough manner. Any allegations will be reported to the relevant University Safeguarding Officers ***without delay***. Any such allegations that may need onward referral to external agencies will be dealt with within one working day. The Designated Safeguarding Lead must not investigate the matter, and must refer as promptly as possible.

- The Designated Lead completed OSCB's (Oxfordshire Safeguarding Children Board) online training on Safeguarding Children in July 2017 (see certificate attached), and received an enhanced DBS check clearance in April 2018, which has been recently renewed.

Risks to the participants:

We do not anticipate any ethical, emotional, physical (for Covid-19 related risk please see attached risk assessment form) or political risks to our participants, given that our topic focuses on understanding technology usage patterns and education of technology usage. As outlined above, we have ensured that all research staff have received appropriate training, including 1) the use of appropriate research methods, 2) conduct relating to engagement with children, 3) recognition of and dealing with ethical issues, and 4) recognition of and dealing with situations where abuse and/ or serious risk is identified.

The questions asked in the studies do not concern personal matters, and thus are unlikely to cause participants undue stress or carry any other potential psychological risks.

The data to be collected will be kept on a secure server with strict access control in place. They will not be shared with any third parties or external organisations. Participants will have ample opportunities to request to delete data recordings at any time during the project. Once a study is completed, all recordings will be transcribed as soon as possible, with all names removed. The recordings which contain any personally identifiable information will be deleted from our secure desk immediately. All the anonymised transcriptions will be kept on our secure and encrypted hard drive, and any linkage to our pseudonymisation token will be kept accessible only to the lead researchers. The participants would have the right to look at the results and transcriptions of their own study if they want to, but not any other data collected throughout the study (e.g. transcriptions of others, data analysis).

Participants will be reminded that they can withdraw from the study at any time, and studies will be stopped at any time when children show any sign of stress or unwillingness to cooperate. The deadline by which they can withdraw any information the child has contributed to the research is [*30 September 2023*].

8. Please give details of any other ethical and/or safety considerations, including whether there might be any risks or benefits to the wider community.

n/a

9. How do you propose to deal with / handle any incidental findings?

There is a possibility that during the study we discovered that children participants have been exposed to inappropriate content or personal safety risks. The implications of such incidence can vary on a case by case basis, depending on the age of the children and the content. If our researcher recognises that such an incident may pose serious risks to children's safety and well-being, we would stop the interview immediately and have a group discussion, with the presence of our Safeguarding Lead, to decide the most appropriate response to the situation. We would always obtain the children's assent before raising our concerns to their parents/guardians or police forces.

10. Will any data or information from this study be provided to individual participants?

No, but participants will be provided with an option if they would like to be kept informed of the study results.

Section G. Other considerations

1. Is any part of this research being conducted overseas? **If yes,** please give details below. Explain how you will address any ethical issues specific to the local context. Please provide details of the local review, approval or permission obtained or required. If there will be no local review, explain why not. You may find it helpful to refer to CUREC's BPG 16: Social science research conducted outside the UK. Ensure you complete and submit a travel risk assessment to your departmental safety officer, if your department requires this. (This is necessary to ensure the travel/ fieldwork is covered by the University's travel insurance – see http://www.admin.ox.ac.uk/finance/insurance/travel)Please also address any physical or psychological risks for Oxford researchers and local fieldworkers in the 'Ethical Considerations' section above and discuss these with your safety officer.	Yes ☐	No ☒

2. Please list any stakeholder or community engagement that has been, or will be, undertaken in relation to the research.

3. Does your research raise issues relevant to the Counter-Terrorism and Security Act (the Prevent Duty), which seeks to prevent people from being drawn into terrorism? **If yes,** please say how you plan to address any related risks. Please see advice on this on our Best Practice Guidance Web Page.	Yes ☐	No ☒

Section H. Data management and handling

All information provided by participants is considered **research data** for the purpose of this form. Any research data from which participants can be identified is known as **personal data**; any personal data which is sensitive is considered **special category data**.

Management of personal data, either directly or via a third party, must comply with the requirements of the UK General Data Protection Regulation (UK GDPR) and the Data Protection Act 2018, as set out in the University's Guidance on Data Protection and Research. In answering the questions below, please also consider the points raised in the Data Protection Checklist and whether, for higher-risk data processing, a separate Data Protection Impact Assessment (DPIA) may also be required for the research. Advice on research data management and security is available from Research Data Oxford and your local IT department. Advice on data protection is available from the Information Compliance team.

1. **Please mark 'X' against the data you will collect for your research**

Screening documents	☐	Audio recordings	☒
Consent records including participant name or other identifiers (e.g. written consent forms, audio-recorded consent, assent forms)	☒	Video recordings	☐
Consent obtained anonymously (e.g. via online survey)	☐	Transcript of audio/video recordings	☒
Opt-out forms	☐	Photographs	☐
Contact details for the purpose of this research only	☒	Information about the health of the participant (including mental health)	☐
Contact details for future use	☒	Physiological test results / measurements	☐
Field notes	☐	MRI scans	☐
Task results (e.g. questionnaires, diary completion)	☒	IP addresses (refer to Best Practice Guidance 09: Data collection, protection and management for guidance)	☐
Data already in the public domain. Specify the source of the data:	☐	Other (please specify below)	☐
Previously collected (secondary) data	☐		
Bank details for payment	☐		

2. **How and where will each type of data be stored whilst the research is ongoing (until the end of all participant involvement)?**

List each type of data selected above, and explain how each will be physically transferred (including movement/sharing of audio files, paper records, electronic downloads etc.) from where it is collected to a suitable storage site (e.g. Nexus365 OneDrive for Business, SharePoint, University servers). State the storage location for each. Do not store unencrypted data in freely available cloud services or unprotected USB drives.

Refer to Best Practice Guidance on data collection, protection and management (BPG09).

Consent records, and contact details (for the purpose of this research & for future use) will be collected from participants online, and stored electronically in password-protected folders on the research team's OneDrive within the University network. Audio recordings will be transferred from the recording device or secure web server to be stored as password-protected files on folders on the research team's OneDrive within the University network. They will then be deleted from the original recording device (encrypted audio recorder with password protection). A researcher from the team will transcribe the audio recordings. The audio recording held by the researchers will then be deleted. The transcription will be stored as a Word file on encrypted OneDrive folders within the University network.
3. Will you use a unique participant number on research data instead of participant name? **If yes,** state whether or not you will retain a list of participant names against numbers (pseudonymisation via a linkage list). **Where will the list be stored, and when will it be destroyed?**
Yes, this linkage file will be kept in a separate encrypted folder on an encrypted hard drive, which will be kept in a locked cabinet in the PI's office.
4. Who will have access to the research data?
Researchers listed on this form will have access to the research data.
5. If research data is to be shared with another organisation, how will it be transferred / disclosed securely?
n/a
6. **Are there any risks associated with the collection or transfer of the research materials, including at border checks? If so, describe the steps that will be taken to address these risks.**
n/a
7. When and how will identifiable data (including audio/video recordings & photos) be destroyed or deleted? N.B. If any identifiable data will be retained beyond the end of the study and/or indefinitely, please state what data this is, and the reasons for retention (e.g. contact details for future studies; photos used in publication). This must be clearly stated on participant information, and specific consent obtained.
Audio recordings will be transferred from the recording device or secure web server to be stored as password-protected files on folders on the research team's OneDrive within the University network. They will then be deleted from the original recording device. A researcher from the team will transcribe the audio recordings. The audio recording held by the researchers will then be deleted as soon as possible (within a week). The transcription will be stored as a Word file on encrypted OneDrive folders within the University network.

8. Please confirm that you will store other (non-identifiable) research data safely for at least 3 years after final publication or public release and adhere to any additional research funder policies. For more information about the University policies, please see the University's webpages on research data management. **If 'Yes'**, please give details of who will store the data and on storage format, location and security. **If 'No'**, please provide further details.	Yes ☒	No ☐
Data will be stored electronically in password-protected files on encrypted computers within the University network		

Section J. Publication and dissemination of results

1. Will you preregister this research?		Yes ☐	No ☒
2. If yes, please state the platform where it will be preregistered	n/a		
3. How will you disseminate project outcomes at the end of the research?	Research results will be disseminated as public reports, technical reports, as well as academic publications. Public reports will be written in a language accessible to the general public. We will ask all participants for their permission to use direct quotes, and all quotes will have all identifiable names and places removed. All research outcomes will be made available through the project's website and open access wherever possible. Study participants can also opt-in to receive updates of research publications directly at the end of the study.		

Declaration and signatures

In providing signatures, the IDRECs accept either:

Option 1: Email confirmations sent from a University of Oxford email address. Separate emails should be sent by each of the relevant signatories as outlined below, indicating acceptance of their responsibilities.

Option 2: That the form be fully-signed with handwritten (wet-ink) signatures. Please scan these and the rest of the form pages to create a single PDF document and email to us.

The form should be sent with Word versions of all documents by email to:

ethics@medsci.ox.ac.uk (for applications from the Medical Sciences and MPLS divisions)

ethics@socsci.ox.ac.uk (for applications from the Social Sciences and Humanities divisions)

Applications from departments with a departmental research ethics committee (DREC) should first be sent for initial review to the relevant DREC.

Pasted images of signatures cannot be accepted

Principal Investigator (and student if applicable)

I/We, the researcher(s):

- Understand our responsibilities as outlined on this form and in the CUREC glossary and guidance
- Agree to start this research only after obtaining approval from the IDREC;
- Understand that the Principal Investigator must ensure that all researchers are suitably qualified and trained to conduct the research described, or are appropriately supervised until deemed qualified/trained;
- Agree to provide additional information as requested by the IDREC before approval is secured and as research progresses;
- Agree to maintain the confidentiality of all data collected from or about participants;
- Agree to notify the IDREC in writing immediately of any proposed change to the research, and await approval before proceeding with the proposed change;
- Agree to notify the IDREC if the Principal Investigator changes and supply the name of the successor;
- Will use the data collected only for the research for which approval has been given;
- Will grant access to data only to authorised persons; and
- Have made arrangements to ensure that personal data collected from participants will be held in compliance with the requirements of UK GDPR and the Data Protection Act 2018.

Principal Investigator (Name)	
Principal Investigator (Signature) **(**Wet-ink signature, not pasted electronic image)	

Date	
Student (Name)	Ge Wang
Student (Signature) **(**Wet-ink signature, not pasted electronic image)	
Date	27 September 2022

Acceptance by Head of Department/Faculty or Designated Nominee*

*Another senior member of the department may sign where the head of department is the Principal Investigator, or where the head of department has appointed a nominee. Example nominees include Deputy Head of Department, Director of Research, and Director of Graduate/ Undergraduate Studies.

On the basis of the information available to me, I confirm that:

- I am aware of the research proposed and have read this application;
- To the best of my knowledge, the proposed design and scientific methodology do not raise ethical concerns;
- I support this research in principle, subject to ethical and other necessary reviews.

Head of Department or designated nominee (Name)	
Head of Department or designated nominee (Signature) Wet-ink signature (not pasted electronic image) *or* The Head of Department/nominee can send an email (including PI name and study title) to ethics@medsci.ox.ac.uk or ethics@socsci.ox.ac.uk confirming the above	
Date	

SOCIAL SCIENCES & HUMANITIES
INTERDIVISIONAL RESEARCH ETHICS COMMITTEE

Research Governance, Ethics & Assurance Team, Research Services, University of Oxford,
Boundary Brook House, Churchill Drive, Headington, Oxford OX3 7GB, UK
Tel: +44(0)1865 289881 Email: ethics@socsci.ox.ac.uk

UNIVERSITY OF OXFORD

Tiffany Wang
Department of Computer Science
University of Oxford

30 November 2022

Dear Tiffany,

Research ethics approval

Research title: "A longitude study of KOALA Hero for assisting family data privacy literacy development"

Research ethics reference: R83638/RE001

The above application has been considered by the Social Sciences and Humanities Interdivisional Research Ethics Committee (SSH IDREC) in accordance with the University's procedures for ethical approval of all research involving human participants.

I am pleased to confirm that, on the basis of the information provided to the IDREC, ethics approval has now been granted for this study for the period 28 November 2022 to 27 May 2024.

Please note the following:
Personal data: It is the responsibility of the PI to ensure that all personal data collected during the project is managed in accordance with the University's guidance and legal requirements.
In-person activities: Any data collection involving in-person interactions with participants must have an up-to-date fieldwork risk assessment in place; further guidance is available from the Safety Office's website.
Amendments: Please notify the committee if you intend to make any amendments to the information in your ethics application as submitted at date of this approval, as all changes must receive ethical approval prior to implementation. The amendment form is available on the SSH IDREC webpage.
Adverse events: The SSH IDRECs must be notified within seven days of any unexpected adverse consequences to the research participants, researchers or other people involved in this research project.
Audit: The SSH IDREC audits a sample of projects each year to enable the Committee to monitor the ethical aspects of research in progress.

We welcome feedback on your experience of the ethical review process and suggestions for improvement. Please email any comments to ethics@socsci.ox.ac.uk.
Yours sincerely

DocuSigned by:
E Sparke
8213B0CC40BB4D7...

Mrs Bess Sparke
Research Ethics Administrator
For
Jennifer Blaikie
Research Ethics Manager (SSH IDREC)

cc: Dr Jun Zhao, ethics@cs.ox.ac.uk

A.4 CUREC forms and ethics approval letters for Chapter 7

Central University Research Ethics Committee (CUREC)

CUREC 2 Form

Higher-risk research involving human participants and/or their data

The University of Oxford places a high value on the knowledge, expertise, and integrity of its members and their ability to conduct research to high standards of scholarship and ethics. The research ethics clearance procedures have been established to ensure that the University is meeting its obligations as a responsible institution. They start from the presumption that all members of the University will take their responsibilities and obligations seriously, and will ensure that their research involving human participants is conducted according to the established principles and good practice in their field and in accordance, where appropriate, with legal requirements.

Before completing this application, please work through the guidance on our website to ensure that you do not need to submit a CUREC 1/1A (minimal-risk review) or CUREC 3 (full Committee review for certain applications to the MS IDREC) application instead. Only type-written forms will be accepted.

Section A. Research Details	
1. Full title of research	Supporting Children's Data Autonomy on Social Media Platforms
2. Short title of research	Supporting Children's Data Autonomy on Social Media Platforms
3. Principal Investigator (PI) / Student Supervisor	Ge Wang, Dphil in Computer Science
4. PI's training in research ethics and research integrity Information about online training	Research integrity online training - 28 September 2020
5. PI - date of completion of Information Security training	
6. Student name and degree programme (if applicable)	Ge Wang, Dphil in Computer Science
7. Department/Institute name	Department of Computer Science
8. University email address	Ge.wang@cs.ox.ac.uk
9. University telephone number	01865 273 875

10. Funding Source (required for ethics team use)	Ethical Web and Data Architecture (EWADA)
11. State any conflicts of interest and explain how these will be addressed	None

Section B. Researchers

1. Researcher title and name	Jun Zhao
2. Department / Institute name or affiliation	Department of Computer Science
3. Role in Research	DPhil co-supervisor
4. Training in Research Ethics and/or research integrity Information about online training	Research integrity online training - 28 September 2020
5. Date of completion of Information Security training	25/05/2023

Section C. Basic information

1. Provide a brief lay summary of the aims and objectives of the research. This should cover the questions it will answer, any potential benefits and what you will do to address the question. **(Maximum 300 words)**	We will conduct a series of lab studies investigating how different designs have impact on children's sense of data autonomy. They will be invited to interact with a newly developed system by our team, during which their interactions with the system will be recorded. The goal of this project is to look into how can we better support children's control of their own data online. More specifically, participants will be invited to experience how their data could be handled on TikTok. Supported by ALPODCA, they will have the opportunity to manage how data will be collected from, how their data will be processed by the algorithms, and finally what type of inference might be made on them.
2. List all places where research will be conducted (including any other countries and	The study will take place in the Department of Computer Science, through university approved protocols. The study participants will be residents of the UK.

online)	
3. Anticipated research start date	17 June 2023
4. Anticipated research end date (n.b. A maximum of 5 years approval can be granted)	31 December 2023
5. Please list any CUREC Approved Procedure(s) you will follow	AP25 Non-invasive research methods with children
6. Please list any CUREC Best Practice Guidance used to develop your research	BPG 05 Payments and incentives in research BPG 09 Data collection, protection and management BPG 10 Conducting research interviews
7. Please list any Professional Guidelines used	Association for Computing Machinery (ACM) Code of Ethics and Professional Conduct
8. Name of departmental / peer reviewer (if applicable)	N/A
9. Will you submit, or have you submitted, this research for ethical review or consideration elsewhere (e.g. local or collaborator's ethics committee, or other local approval)?	No

Section D. Participants

(n.b. where there is no contact with human participants (in person or virtual) and no observation of them, but only use of data about them, please omit this section, and complete section I instead)

1. Age range of participants	8-15 (we found that from 8 onwards, children will more capable of articulating their thoughts on more advanced topic such as data practices online, and we would like to see how children's perceptions change across different age groups)
2. Are research participants people who may not be able to give free and informed consent? e.g. those under 18, prisoners, or adults 'at risk'	Yes, children under 18.

3. Anticipated number of participants	10-20.
4. How was the number of participants decided?	Based on previous similar research, established practices and the nature of this type of research (user study with prototypes), we believe this is sufficient number.
5. Inclusion Criteria	Children between 8-15 who have experience with social media platforms
6. Exclusion criteria	Children outside the age range, or have limited experience with social media platforms
7. Please mark 'X' against all planned recruitment methods Provide copies of all recruitment material for review	Poster advert ☒ Flyer ☒ Email circulation ☒ In-person approach ☒ Website ☒ Social media (e.g. twitter, Facebook) ☒ Snowball sampling (recruiting through contacts of existing participants) ☒ Newspapers ☐ Research recruitment sites (e.g. Prolific Academic, Amazon Turk) ☒ Existing departmental contacts or volunteer database ☒ Other (please specify below) ☐
8. How will potential participants be identified and approached?	We will use Prolific Academic as our main recruitment channel, and we will share flyers and poster avert via our project website and Twitter account. We will also directly send emails to school head teachers. We will verify the suitability of the potential participants via email or over the phone. The criteria for recruitment is that children needs to be between 8 and 15, and have some experience of using social media.
9. Will informed consent be obtained from the research participants or their parents/ guardians? If not, please explain why not.	Yes, we will obtain consent from their parents and guardians, and assent from the children participants. Parents, guardians and children will be fully acknowledged that participation is voluntary and that they are free to withdraw without giving any reason. The deadline by which they can withdraw any information the child has contributed to the research is [*31 September 2023*].
10. For each activity or group of participants, explain how informed consent will be obtained from the participants themselves and/or their parents/guardians, if applicable. How will their consent be recorded?	Consent will be collected digitally, with parents/guardians and children reply in an email with the following sentence. The consent emails will be stored on an encrypted drive as detailed in section H2: I, _____, confirm that I have read and understood the details of the study, and have had the opportunity to ask questions and discuss the study with others. I have received satisfactory answers to my questions. I understand that the project has received ethics clearance through the University of Oxford's ethical approval process for research involving human participants, and I understand who will have access to the data, how it will be stored and what will happen to the data at the end of the study. I understand that participation is voluntary and that my child and I are free to withdraw at any time without giving any reason. I understand how to raise a concern or make a complaint.

	I, _____, I have read and understood the details of the study, and understand what this study is about. I have had the opportunity to ask questions that I want and my questions have been answered in a way that I understand. I understand that it's OK to stop taking part at any time. I am happy to take part and for my video and voice to be recorded.

Section E. Research Methodology

1. Please mark 'X' against the methods that will be used in your research

Ensure you address each method you will use in your informed consent documents and on this form

Use of casual or local workers (e.g. interpreters)	☐	Audio recording of participant	☒
Interview (refer to guidance in BPG 10: Conducting research interviews)	☒	Video recording of participant	☐
Focus group	☐	Photography of participant	☐
Participant completes questionnaire in hard copy	☐	Physiological recording from participant	☐
Participant completes online questionnaire or other online task (refer to guidance in BPG 06:	☐	Taking a sample of blood or other bodily fluid from a participant	☐
Use of social media to recruit or interact with participants (refer to guidance in BPG 06: Internet-mediated research)	☒	Participant observation	☐
Analysis of existing records	☐	Covert observation	☐
Participant performs verbal or aural task	☐	Systematic observation	☐
Participant performs paper and pencil task	☐	Observation of specific organisational practices	☐
Participant performs computer based task	☒	Other (please specify below)	☒
Measurement/recording of motor behaviour	☐	Screen recording	

2. **Provide a lay description of the research design and methods. In particular, describe clearly what participants in the research will be asked to do.**
Through a series of lab studies (90 minutes) held in-person in the department of Computer Science, we would like to learn more about how children's data autonomy on social media platforms can be supported. In the study, children participants interact with a developed prototype called 'PodTok" which enables them to freely manage their data being collected, how their data are processed and how things about them are inferred as they interact with social media. During the lab study (see attached study design), we will first have a brief chat with children regarding their existing experience and perception of how their data is treated online. We will then introduce our app prototype, and observe their interaction with the system. We will conclude the study with an interview on their perceptions on the prototype. Each session will be accompanied by at least two researchers, with at least 1 DBS checked. Parents and guardians are not required to be present in the study, as long as the children are comfortable with the arrangement. However, parents and guardians are more than welcome to stay if they feel necessary.
3. **Will the research include any audio, video or photographic recordings?**
Yes. Audio recordings of the interviews with children will be recorded for the sake of data collection. Audio recordings will be mandatory for participating in the study. Screen recordings might also be taken for children's interaction with the prototype if needed.
4. **Please detail any expenses or gifts that will be offered to participants** Guidance is available in Best Practice Guidance: 05 Payments and incentives in research.
A £15 Amazon gift card will go to each family, and stickers will be given to children.

Section F. Ethical Considerations

For guidance on ethical issues, please see http://researchsupport.admin.ox.ac.uk/governance/ethics/resources

(N.B. To complete, double click on the check boxes and select 'checked')

1. Will the research involve any participants considered vulnerable in the context of the research (e.g. children, elderly, prisoners, adults "at risk")? **If yes,** please describe how they are defined as vulnerable and detail any CUREC Approved Procedures or guidance that will be applied to the research (for current documents and templates see https://researchsupport.admin.ox.ac.uk/governance/ethics/resources). For research involving children, please state why either CUREC Approved Procedure 15 or 25 cannot be applied wholly to your research.	Yes ☒	No ☐
At least 2 researchers with DBS check will accompany the children during the studies, their parents are allowed to stay if they feel the need. The names of all participants (both children and their parents) in the lab will be identifiable, should any issue arise.		
2. Will unequal relationships exist between participants and those obtaining informed consent? **If yes,** describe the nature of the unequal relationship and how arising ethical issues will be addressed	Yes ☒	No ☐
We will make sure with both the parents and children that they are free to stop at anytime before and during the study, and they have till 31 September 2023 to inform us if they do not want their data to be stored and used by this study.		
3. Will the research involve questions and/or discussions of contentious and/or sensitive issues (e.g. information relating to ethnicity, political opinions, religious beliefs, physical/mental health or sexual life)? **If yes,** please justify why this is required and provide the questions (or an outline of them) raising the issues that will be used in your research.	Yes ☐	No ☒
4. Will taking part in the research put participants under any particular burden and/or risk (including risk of prosecution)? **If yes,** describe how risks will be mitigated. If there is a risk of prosecution to the participant, justify why incriminating data are sought. During the consent process, participants should be made aware of the risks of disclosing potentially illegal information and understand what the researchers would do if they were to receive that information.	Yes X	No
Participants wishes as to confidentiality will be respected unless legal and safeguarding obligations require otherwise, but that in certain circumstances it may be necessary to inform other agencies if a safeguarding or police issue arises		
5. Will the research involve deliberate deception of participants beyond that covered by CUREC Approved Procedure 07? **If yes**, justify why deception is used, describe deception and debriefing process, and include debriefing documents in the application	Yes ☐	No ☒

6. **Could the proposed research affect your own physical and/or psychological safety as a researcher?** **If yes,** describe how you will manage this. Explain what safety procedures, structured mentoring or other ongoing support will be in place during this research. Include details of lone working procedures, if applicable.	Yes ☐	No ☒

7. How will you ensure the research is conducted according to the details given in this form?

All members involved in the research have received extensive research ethics training and have extensive experience of conducting research with human/children participants. Furthermore, we will run weekly project meetings to discuss project progress and address any potential concerns that could be raised during the study.

Risks to researchers:

Researchers will carry out activities within the department of Computer Science, following pre-established protocols and risk assessments. This reduces risks to the researchers. Further, we will not study a risky or sensitive topic, and therefore we do not anticipate any expression of distress from the participants during researchers' interactions with them. To further minimise risks to the researchers, all research staff would read through the University's Safeguarding Code of Practice before conducting the research study. We will also carry out a risk assessment (including Covid-19 related risks), following the attached risk assessment form, prior to the study.

Dr Zhao will be appointed as the "Designated Safeguarding Lead" of the study and work together with the University Safeguarding Officers to ensure all safeguarding processes are in place and the University's Code of Practice is thoroughly followed through. The following procedures will be carried out:

- The Safeguarding Lead will carry out a risk assessment according to the attached risk assessment form for all members of staff involved in the study, and identify the need for any pre-employment or pre-activity checks in accordance with the University's guidance.

- The lead will also ensure that all members of staff will be under appropriate supervision at all time. The likelihood of researchers working alone with young children will be minimised.

- Both the contact details of the University's Safeguarding Officer and the nominated Safeguarding Lead will be made available to the children and parents/guardians at the start of the study, in case any complaints made against any member of the team arise. The procedure outlined in the University's Safeguarding Code of Practice (Section 4) will be followed through to ensure that issues are dealt with in a timely and thorough manner. Any allegations will be reported to the relevant University Safeguarding Officers ***without delay***. Any such allegations that may need onward referral to external agencies will be dealt with within one working day. The Designated Safeguarding Lead must not investigate the matter, and must refer as promptly as possible.

- The Designated Lead completed OSCB's (Oxfordshire Safeguarding Children Board) online training on Safeguarding Children in July 2017 (see certificate attached), and received an enhanced DBS check clearance in April 2018, which has been recently renewed.

- Ge Wang has been advised by the Departmental Research Ethics Committee to undertake generalist safeguarding training, and that this will be completed on 25 May, in time for the study to begin; The certificate will be sent to the Departmental Ethics Committee upon completion. She received a DBS check clearance in June 2023.

Risks to the participants:

We do not anticipate any ethical, emotional, physical (for Covid-19 related risk please see attached risk assessment form) or political risks to our participants, given that our topic focuses on understanding technology usage patterns and education of technology usage. As outlined above, we have ensured that all research staff have received appropriate training, including 1) the use of appropriate research methods, 2) conduct relating to engagement with children, 3) recognition of and dealing with ethical issues, and 4) recognition of and dealing with situations where abuse and/ or serious risk is identified.

The questions asked in the studies do not concern personal matters, and thus are unlikely to cause participants undue stress or carry any other potential psychological risks.

Participants will be reminded that they can withdraw from the study at any time, and studies will be stopped immediately when young children show any sign of stress or unwillingness to cooperate.The deadline by which they can withdraw any information the child has contributed to the research is [*31 September 2023*].
8. Please give details of any other ethical and/or safety considerations, including whether there might be any risks or benefits to the wider community.
N/A
9. How do you propose to deal with / handle any incidental findings?
There is a possibility that during the study we discovered that children participants have been exposed to inappropriate content. The implications of such incidence can vary on a case by case basis, depending on the age of the children and the content. If our researcher recognises that such an incident may pose serious risks to children's safety and well-being, we would stop the interview immediately and have a group discussion, with the presence of our Safeguarding Lead, to decide the most appropriate response to the situation. We would always obtain the children's assent before raising our concerns to their parents/guardians or police forces.
10.Will any data or information from this study be provided to individual participants?
No, but participants will be provided with an option if they would like to be kept informed of the study results.

Section G. Other considerations

1. Is any part of this research being conducted overseas? **If yes,** please give details below. Explain how you will address any ethical issues specific to the local context. Please provide details of the local review, approval or permission obtained or required. If there will be no local review, explain why not. You may find it helpful to refer to CUREC's BPG 16: Social science research conducted outside the UK and the Code of Conduct for Ethical Fieldwork. Ensure you complete and submit a travel risk assessment to your departmental safety officer, if your department requires this. (This is necessary to ensure the travel/ fieldwork is covered by the University's travel insurance – see http://www.admin.ox.ac.uk/finance/insurance/travel)Please also address any physical or psychological risks for Oxford researchers and local fieldworkers in the 'Ethical Considerations' section above and discuss these with your safety officer.	Yes ☐	No ☒
2. Please list any stakeholder or community engagement that has been, or will be, undertaken in relation to the research.		
3. Does your research raise issues relevant to the Counter-Terrorism and Security Act (the Prevent Duty), which seeks to prevent people from being drawn into terrorism? **If yes,** please say how you plan to address any related risks. Please see advice on this on our Best Practice Guidance Web Page.	Yes ☐	No ☒

Section H. Data management and handling

All information provided by participants is considered **research data** for the purpose of this form. Any research data from which participants can be identified is known as **personal data**; any personal data which is sensitive is considered **special category data**.

Management of personal data, either directly or via a third party, must comply with the requirements of the UK General Data Protection Regulation (UK GDPR) and the Data Protection Act 2018, as set out in the University's Guidance on Data Protection and Research. In answering the questions below, please also consider the points raised in the Data Protection Checklist and whether, for higher-risk data processing, a separate Data Protection Impact Assessment (DPIA) may also be required for the research. Advice on research data management and security is available from Research Data Oxford and your local IT department. Advice on data protection is available from the Information Compliance team.

1. Please mark 'X' against the data you will collect for your research

Screening documents	☐	Audio recordings	☒
Consent records including participant name or other identifiers (e.g. written consent forms, audio-recorded consent, assent forms)	☒	Video recordings	☐
Consent obtained anonymously (e.g. via online survey)	☐	Transcript of audio/video recordings	☒
Opt-out forms	☐	Photographs	☐
Contact details for the purpose of this research only	☒	Information about the health of the participant (including mental health)	☐
Contact details for future use	☒	Physiological test results / measurements	☐
Field notes	☒	MRI scans	☐
Task results (e.g. questionnaires, diary completion)	☐	IP addresses (refer to Best Practice Guidance 09: Data collection, protection and management for guidance)	☐
Data already in the public domain. Specify the source of the data:	☐	Other (please specify below)	☒
Previously collected (secondary) data	☐	**Screen recording of children's interaction with the platform.**	
Bank details for payment	☐		

2. How and where will each type of data be stored whilst the research is ongoing (until the end of all participant involvement)?

List each type of data selected above, and explain how each will be physically transferred (including movement/sharing of audio files, paper records, electronic downloads etc.) from where it is collected to a suitable storage site (e.g. Nexus365 OneDrive for Business, SharePoint, University servers). State the storage location for each. Do not store unencrypted data in freely available cloud services or unprotected USB drives.

Refer to Best Practice Guidance on data collection, protection and management (BPG09).

Consent records, and contact details (for the purpose of this research & for future use) will be collected from participants online, and stored electronically in password-protected folders on the research team's OneDrive within the University network.

Audio recordings, as well as screen recordings will be transferred from the recording device to be stored as password-protected files on folders on the research team's OneDrive within the University network. They will then be deleted from the original recording device (encrypted audio recorder with password protection, and encrypted phone). A researcher from the team will transcribe the audio recordings. The audio recording held by the researchers will then be deleted. The transcription will be stored as a Word file on encrypted OneDrive folders within the University network
3. Will you use a unique participant number on research data instead of participant name? **If yes,** state whether or not you will retain a list of participant names against numbers (pseudonymisation via a linkage list). **Where will the list be stored, and when will it be destroyed?**
Yes, this linkage file will be kept in a separate encrypted folder on an encrypted hard drive, which will be kept in a locked cabinet in the PI's office.
4. Who will have access to the research data?
Researchers listed on this form will have access to the research data.
5. If research data is to be shared with another organisation, how will it be transferred / disclosed securely?
N/A
6. Are there any risks associated with the collection or transfer of the research materials, including at border checks? If so, describe the steps that will be taken to address these risks.
All research data will be stored in digital format, stored as password-protected files on folders, and deleted once the transcription is done. There is no known risk during this process.
7. When and how will identifiable data (including audio/video recordings & photos) be destroyed or deleted? N.B. If any identifiable data will be retained beyond the end of the study and/or indefinitely, please state what data this is, and the reasons for retention (e.g. contact details for future studies; photos used in publication). This must be clearly stated on participant information, and specific consent obtained.
Screen and audio recordings will be transferred from the recording device to be stored as password-protected files on folders on the research team's OneDrive within the University network. They will then be deleted from the original recording device. A researcher from the team will transcribe the audio recordings. The audio recording held by the researchers will then be deleted as soon as possible (within a week). The transcription will be stored as a Word file on encrypted OneDrive folders within the University network

8. Please confirm that you will store other (non-identifiable) research data safely for at least 3 years after final publication or public release and adhere to any additional research funder policies. For more information about the University policies, please see the University's webpages on research data management. **If 'Yes'**, please give details of who will store the data and on storage format, location and security. **If 'No'**, please provide further details.	Yes ☒	No ☐
Data will be stored electronically in password-protected files on encrypted computers within the University network		

Section I. Research involving secondary use or disclosure of personal data or special category data

This section of the form is only to be completed for research activity (as part or all of the research) where there is no contact with human participants (in person or virtual) and no observation of them, only use of data about them.

Your research must meet the standards laid down in the Data Protection Act 2018 with respect to the collection, use, and storage of personal data about human participants.

1. Will you seek data access agreements for these data? Yes ☐ No ☐

If yes,

- List the individual(s) or organisation(s) from which the information will be sourced
- Attach a copy of the agreement with the individual(s) or organisations in question
- Provide details of any conditions imposed by the organisation(s) concerning the release of the information

If no, please explain how and when the agreement of the disclosing organisation(s) will be obtained

2. Could these data be linked back to an individual or individuals? Yes ☐ No ☐

If yes,

- Please explain why data cannot be collected in a way that prevents linkage with an individual/individuals
- Say how individual consent was obtained for the collection, use or disclosure of linkable data

If no, you do not need to complete the rest of this section

3. How will any personally identifiable data be transferred to you?

Please describe the arrangements for any physical transfer of personal data (including paper records and data captured electronically via portable media) from where you are obtaining it to local storage

4. Where, and for how long, will personally identifiable data be stored during and after the research?

Please outline procedures for ensuring confidentiality, e.g. security arrangements, pseudonymisation etc.

5. Who will have access to the personally identifiable data?

If data is to be shared with another organisation, other than the researchers listed, how will it be transferred / disclosed securely

6. When and how will personally identifiable data be destroyed?

Section J. Publication and dissemination of results

1. Will you preregister this research?	Yes ☐	No ☒

2. If yes, please state the platform where it will be preregistered	N/A
3. How will you disseminate project outcomes at the end of the research?	Research results will be disseminated as public reports, technical reports, as well as academic publications. Public reports will be written in a language accessible to the general public. We will ask all participants for their permission to use direct quotes, and all quotes will have all identifiable names and places removed. All research outcomes will be made available through the project's website and open access wherever possible. Study participants can also opt-in to receive updates of research publications directly at the end of the study.

Section K. Additional questions for applications to the Medical Sciences IDREC

1. List any standardised questionnaires that will be utilised (there is no need to send a copy)	
2. List any additional questionnaires designed by the researchers – a copy of these must be sent to the MS IDREC for review	
3. Give details of any biological sample(s) that will be taken (e.g. blood, urine, saliva, faeces)	State the volume of sample, and the frequency of sampling. Describe briefly how the sample will be processed and stored once taken, and confirm that it will be rendered into a form not relevant under the Human Tissue Act before use in the research. All stored samples must be fully anonymised (no means of identification by any member of the research team) or pseudonymised (samples may be identified via a linkage document securely held elsewhere). Please say which will apply to your research. Say who will have access (e.g. research team only), and whether it will be stored long-term for use in future ethically approved studies). Provide a brief overview of the laboratory analyses that will be performed and how the samples will be destroyed (if appropriate).

Declaration and signatures

In providing signatures, the IDRECs accept either:

Option 1: Email confirmations sent from a University of Oxford email address. Separate emails should be sent by each of the relevant signatories as outlined below, indicating acceptance of their responsibilities.

Option 2: That the form be fully-signed with handwritten (wet-ink) signatures. Please scan these and the rest of the form pages to create a single PDF document and email to us.

The form should be sent with Word versions of all documents by email to:

ethics@medsci.ox.ac.uk (for applications from the Medical Sciences and MPLS divisions)

ethics@socsci.ox.ac.uk (for applications from the Social Sciences and Humanities divisions)

Applications from departments with a departmental research ethics committee (DREC) should first be sent for initial review to the relevant DREC.

Pasted images of signatures cannot be accepted

Principal Investigator (and student if applicable)

I/We, the researcher(s):

- Understand our responsibilities as outlined on this form and in the CUREC glossary and guidance
- Agree to start this research only after obtaining approval from the IDREC;
- Understand that the Principal Investigator must ensure that all researchers are suitably qualified and trained to conduct the research described, or are appropriately supervised until deemed qualified/trained;
- Agree to provide additional information as requested by the IDREC before approval is secured and as research progresses;
- Agree to maintain the confidentiality of all data collected from or about participants;
- Agree to notify the IDREC in writing immediately of any proposed change to the research, and await approval before proceeding with the proposed change;
- Agree to notify the IDREC if the Principal Investigator changes and supply the name of the successor;
- Will use the data collected only for the research for which approval has been given;
- Will grant access to data only to authorised persons; and
- Have made arrangements to ensure that personal data collected from participants will be held in compliance with the requirements of UK GDPR and the Data Protection Act 2018.

Principal Investigator (Name)	
Principal Investigator (Signature) **(**Wet-ink signature, not pasted electronic image)	
Date	
Student (Name)	Ge Wang
Student (Signature) **(**Wet-ink signature, not pasted electronic image)	
Date	16-04-2023

Acceptance by Head of Department/Faculty or Designated Nominee*

*Another senior member of the department may sign where the head of department is the Principal Investigator, or where the head of department has appointed a nominee. Example nominees include Deputy Head of Department, Director of Research, and Director of Graduate/ Undergraduate Studies.

On the basis of the information available to me, I confirm that:

- I am aware of the research proposed and have read this application;
- To the best of my knowledge, the proposed design and scientific methodology do not raise ethical concerns;
- I support this research in principle, subject to ethical and other necessary reviews.

Head of Department or designated nominee (Name)	
Head of Department or designated nominee (Signature) Wet-ink signature (not pasted electronic image) *or* The Head of Department/nominee can send an email (including PI name and study title) to ethics@medsci.ox.ac.uk or ethics@socsci.ox.ac.uk confirming the above	
Date	

SOCIAL SCIENCES & HUMANITIES
INTERDIVISIONAL RESEARCH ETHICS COMMITTEE

Research Governance, Ethics & Assurance Team, Research Services, University of Oxford,
Boundary Brook House, Churchill Drive, Headington, Oxford OX3 7GB, UK
Tel: +44(0)1865 289881 Email: ethics@socsci.ox.ac.uk

UNIVERSITY OF OXFORD

Ge Wang
Department of Computer Science
Mathematical, Physical & Life Sciences Division
University of Oxford

15 June 2023

Dear Ge,

Research ethics approval

Research title: Supporting Children's Data Autonomy on Social Media Platforms

Research ethics reference: R86939/RE001

The above application has been considered by the Social Sciences and Humanities Interdivisional Research Ethics Committee (SSH IDREC) in accordance with the University's procedures for ethical approval of all research involving human participants.

I am pleased to confirm that, on the basis of the information provided to the IDREC, ethics approval has now been granted for this study for the period 12 June 2023 to 11 December 2024.

Please note the following:

Personal data: It is the responsibility of the Principal Investigator to ensure that all personal data collected during the project is managed in accordance with the University's guidance and legal requirements.

Fieldwork: The University's Safety in Fieldwork Policy must be followed.

Amendments: Please notify the committee if you intend to make any amendments to the information in your ethics application as submitted at date of this approval, as all changes must receive ethical approval prior to implementation. The **amendment form** is available on the SSH IDREC webpage.

Adverse events: The SSH IDRECs must be notified within seven days of any unexpected adverse consequences to the research participants, researchers or other people involved in this research project.

Audit: The SSH IDREC audits a sample of projects each year to enable the Committee to monitor the ethical aspects of research in progress.

We welcome feedback on your experience of the ethical review process and suggestions for improvement. Please email any comments to ethics@socsci.ox.ac.uk.

Yours sincerely,

DocuSigned by:
Szilvia Bajkan
D1FA2ECDAD8646E...

Dr Szilvia Bajkan, Research Ethics Administrator, *for*

Jennifer Blaikie, Research Ethics Manager (SSH IDREC)

cc: Dr Jun Zhao, Prof Marina Jirotka, Jordan Summers (ethics@cs.ox.ac.uk)