Who knows what about me?

A Children's Commissioner report into the collection and sharing of children's data

NOVEMBER 2018

Contents

Foreword by the Children's Commissioner, Anne Longfield

More information is collected and shared about us as we go about our daily lives than ever before. It's in the screens we watch, the websites and apps we use and the latest must-have toys and gadgets. And it's not just about technology – information is captured by public services too. Our data footprints are getting bigger and bigger.

This is true for all of us. But the difference for children today is that their data footprints begin from the very moment when their parents proudly upload that first baby photo to social media. On average, by the age of 13, parents have posted 1300 photos and videos of their child to social media. The amount of information explodes when children themselves start engaging on these platforms: on average children post to social media 26 times per day – a total of nearly 70,000 posts by age 18.

We need to stop and think about what this means for children's lives now and how it may impact on their future lives as adults. We simply do not know what the consequences of all this information about our children will be. In the light of this uncertainty, should we be happy to continue forever collecting and sharing children's data?

I don't think we should. We all need to pause and think. At the very least, schools need to start educating their pupils about the importance of guarding personal information. Children and parents need to be much more aware of what they share and consider the consequences. Companies that make apps, toys and other products used by children need to stop filling them with trackers, and put their terms and conditions in language that children understand. And crucially, the Government needs to monitor the situation and refine data protection legislation if needed, so that children are genuinely protected – especially as technology develops.

This is only going to get bigger – so let's take action now to understand and control who knows what about our children.

Anne Longfield OBE
Children's Commissioner for England

Introduction

This report draws attention to the vast amounts of data collected about children growing up today and the ways in which it might shape their lives – not just in the short term, but also in the future, as adults.

Concerns about privacy, and especially children's privacy, are nothing new. For many years now children have been taught that it is very important not to share personal information with people they do not know – whether that be a stranger in the street or when chatting to people online. The issue is framed very much in terms of immediate threats – what if you give someone your address and they wait for you outside your house? What if you give someone a photo and they use it in ways you do not like?

However, the way data is collected and used is changing – rapidly. There are numerous benefits to this, from more evidence-informed policy to services that are more responsive to individual needs. But there continue to be risks. Our old understanding of the risks involved in sharing personal information does not capture the full extent to which it may impact on children's lives in the future.

The Children's Commissioner's Office (CCO) began this project in response to two important observations:

1. More data about children is collected than ever before

It is very difficult to navigate today's world without developing a sizeable data footprint. An immense amount of data is harvested about people as they go about their lives, regardless of age, gender or background.

However, something that sets current and future generations of children apart from the rest of us is that their digital footprints extend right from the moment of birth and then grow exponentially throughout childhood. In fact, some children might find that their digital footprint begins pre-birth, with many parents posting ultrasound photos to social media as a means of announcing pregnancy.

This is not just about parents and children sharing information on social media, even though that is part of the issue. It is also increasingly about smart toys, speakers and other connected devices which are being brought into more and more homes. It is about the proliferation of monitoring equipment that parents can buy, from pedometers to location tracking watches. And it is about information that is given away when children use essential public services such as schools and GPs – something which they might have very little control over. Children are being "datafied" – not just via social media, but in many aspects of their lives.

2. The availability of this data might have significant consequences for children when they become adults.

We all know that in the wrong hands personal information can threaten a child's immediate safety. An obvious example is stranger danger – the risk that a stranger might use knowledge of a child's whereabouts or home address to cause harm to the child. But there is much less understanding of how personal data gathered in childhood might be used to shape an individual's experiences and prospects in the long term – for better or for worse. Could data about a child's language development and early

educational performance at age four play some role in their university application outcomes? Could their parents' shopping habits impact upon the products and services they are targeted with through advertising? Could personal health data affect their ability to take out insurance in future?

The potential for a person's data profile to impact upon their daily experience of life becomes more likely with continued developments in analytical techniques. Natural language processing and machine learning enable us to analyse large swathes of unstructured text that would have previously been unusable. There are methods for identifying individuals in disparate sources of data and linking the information. Algorithms can be used to make predictions about an individual's characteristics on the basis of other data about them. In essence, data can be used to learn, deduce or infer much more about individuals than ever before – and these techniques will continue to become ever more advanced. The rapid pace of development adds to the essential uncertainty in the question: how will data gathered about children today affect their lives in the future?

It is important that there is better understanding and awareness of the volume of children's data that is collected. Only then can policymakers consider whether there need to be greater protections put in place. But it is just as important that children and parents themselves are made aware of the data being collected and what they can do if they are concerned - something that is reflected in the government's revamped guidance on relationships and sex education, which is currently under consultation:

> "Pupils should have a strong understanding of how data is generated, collected, shared and used online, for example, how personal data is captured on social media or understanding the way that businesses may exploit the data available to them."[1]

Educating children early and comprehensively about the many ways in which their data might be used is an important way to foster digital resilience and to help rebalance the power between children and those that gather or use their personal information. Our Life in Likes research[2] showed that although staying safe online is a priority for many children, this is largely limited to protecting themselves from strangers, online predators, cyberbullying and harmful content shared by others. These findings are supported by evidence from ongoing research by the London School of Economics and Political Science which suggests that children see data privacy in a specific way, grasping the significance of information they share directly with others, but not the broader picture involving commercial organisations and public services.[3] Our aim in this project was to draw attention to the fuller picture, and provide some simple, practical steps that can be taken to minimise a child's data footprint if children and parents are concerned.

In July 2018, CCO convened a roundtable with representatives from industry, academia and government agencies to begin a conversation on this topic.[4] This briefing summarises the findings from the roundtable and subsequent research. CCO has also produced two outputs aimed at children,

[1] https://consult.education.gov.uk/pshe/relationships-education-rse-health-education/supporting_documents/20170718_%20Draft%20guidance%20for%20consultation.pdf

[2] https://www.childrenscommissioner.gov.uk/wp-content/uploads/2018/01/Childrens-Commissioner-for-England-Life-in-Likes.pdf

[3] http://www.lse.ac.uk/media-and-communications/research/research-projects/childprivacyonline

[4] The roundtable was attended by representatives from the following: Information Commissioner's Office, Oxford Internet Institute, UCL Institute of Education, Department of Media and Communications at LSE, Wellcome Trust, Snap Inc., Facebook, Google, LEGO Group, Barclays, Mind Of My Own, ASI Data Science and Schillings.

parents and schools: an interactive infographic showing points at which children's data is collected, and a list of ten top tips for minimising children's data footprints. Both of these can be accessed on the CCO website.[5]

What data about children is collected?

The key points at which data about children is routinely collected as they grow up can be divided into three broad categories:

- Data shared online
- Data shared in the home
- Data shared outside of the home

Data shared online

This includes:

- Social media updates on parents' profiles
- Smartphones and tablets
- Web browsing and search engines

Out of the three categories, there is greatest awareness among children and parents of the privacy risks posed by the online world. Children are getting online at younger ages and they are spending more and more of their day online: on average, children aged 5-15 spend 2 hours online on a weekday and 3 hours per day at the weekend.[6] Children aged 11-16 post on social media 26 times a day – if they continue at the same rate, that is a total of nearly 70,000 posts by age 18.[7] The effects of this are wide-ranging and not limited to data privacy – it also impacts on children's sleep, mental and physical health, and social lives (for example, bullying is no longer something that stops at the school gates). And it's not just children – parents of children aged up to 13 share an average of 100 photos and videos of their child each year.[8]

Internet safety was made a compulsory part of the school curriculum in 2014. Many schools participate in the annual Safer Internet Day and organisations such as Internet Matters have been established to support parents to deal with any issues faced by their children when using the internet – including questions of how they share personal information.

There are clear signs that messages around how to engage online positively are reaching children. Indeed we now hear cases of children teaching their parents about the basics of staying safe online rather than the other way around – for example, children explaining the dangers of updating social

[5] For infographic see: www.childrenscommissioner.gov.uk/childrens-data
For top tips see: https://www.childrenscommissioner.gov.uk/2018/10/31/ten-top-tips-for-minimising-childrens-data-footprints
[6] https://www.ofcom.org.uk/__data/assets/pdf_file/0020/108182/children-parents-media-use-attitudes-2017.pdf
[7] https://www.internetmatters.org/wp-content/uploads/2016/05/IM-social-media-A4-V3-1.pdf
[8] https://www.nominet.uk/2-7m-parents-share-family-photos-complete-strangers-online/

media profiles with “first day at school” photos, which often unintentionally reveal the child’s location or identity through details such as school logos and street signs.

But challenges remain. Our Life in Likes report showed that children engage with the online world in a different way to their parents. They use different platforms to do different things – whether it’s Roblox to create their own games, or Snapchat to send quick images and messages that are automatically erased. Although teachers and parents can – and do – give general advice about how to stay safe online, their lack of familiarity with platforms popular with children means that they cannot give specific advice, and parental controls tend to be underutilised. For example, many parents are unaware of Snapchat’s live location sharing feature which children may use without understanding the risks.

Furthermore, there is much greater awareness of some types of data given out online than others. As set out by the London School of Economics and Political Science,[9] there is an important distinction between the following:

- Data that is given directly – e.g. a date of birth posted on someone’s personal information section of their social media profile.
- Data that is “given off” – this is data that is given unknowingly when people go online, captured through technology such as web cookies. It includes metadata, e.g. someone’s location when they posted something or used an app, the time spent using a certain platform, etc.
- Inferred data – when the two previous types of data are analysed, it gives rise to inferred data. This is data based on algorithms and predictions. For example, when someone gives their age, gender and likes certain things on Facebook (all forms of direct data giving), this information might be used to predict which products they might buy – a type of inferred data.

Messages to parents and children about data privacy tend to focus on the first type of data – data that is given directly. They focus much less on raising awareness of data given off and inferred data. Yet these types of data might have real, long-lasting implications on children’s lives, as set out later in this report. Indeed, the amount of data inferred about children was of real concern to many of those who attended our roundtable.

Data shared in the home

This includes:

- Smart speakers
- Connected toys
- Connected baby cameras

Gone are the days when the internet could only be accessed through laptops, tablets and smartphones. An increasing number of personal and household items can now connect to the web, giving rise to what is known as the Internet of Things (IoT). It is an area of huge market growth, being

[9] http://blogs.lse.ac.uk/mediapolicyproject/2018/09/07/conceptualising-privacy-online-what-do-and-what-should-children-understand/

led from the US but with the UK and Europe quickly catching up: for example, in just 6 months – from Autumn 2017 to Spring 2018 – smart speaker ownership in the UK doubled (although remains modest, with 10% of the population owning one).

Many products are targeted for use by children who are too young to use the internet in other ways. For example, connected toys (toys that connect to the internet, e.g. CloudPets or Hello Barbie) are aimed at children as young as three, and location tracking watches are targeted at children who are not old enough to have a smartphone. As various high profile cases have shown, there are many ways in which a child's data (or their family's data) gathered through these devices can fall into the wrong hands. For example, unsecured Bluetooth connections mean that hackers can gain control of some devices, viewing a sleeping child on a baby camera[10] or talking to them through their toy.[11] Furthermore, the data collected and stored in the cloud might not be properly secured. Last year, 2 million CloudPets voice messages shared between children and their family members were found being stored unprotected online.

How connected devices work
CloudPets is an example of a connected toy. CloudPets are cuddly toys with in-built speakers and microphones. They connect to the internet via an app on a nearby smartphone or tablet. Someone away from the child (e.g. a parent working away) can record a message, which is then played through the toy through its connection with the app. It can also be used in reverse, with the child recording messages to be heard by the parent.

Data shared outside the home

This includes:

- Location tracking watches
- School databases
- Study and behaviour apps
- Biometric data in schools
- Retail loyalty schemes
- The Red Book (or Personal Child Health Record)
- Medical records
- Travel passes

Finally data is shared outside the home. Included in this category is information captured by location tracking watches (another type of connected device) and retail loyalty schemes.

[10] https://www.independent.co.uk/life-style/gadgets-and-tech/news/baby-monitors-hacked-parents-warned-to-be-vigilant-after-voices-heard-coming-from-speakers-a6843346.html
[11] https://www.theguardian.com/technology/2017/nov/14/retailers-urged-to-withdraw-toys-that-allow-hackers-to-talk-to-children

Lots of data in this category is collected and shared by parents and children when accessing key public services, such as education and health services. This is the type of data sharing that parents and children are least mindful of: the public accepts that schools, GPs and other services need to know things about children in order to provide them with high quality healthcare and education. Schools and GPs are often perceived to be more trustworthy than commercial organisations for whom there is perhaps a greater incentive to use children's data in ways people would object to. As a result, most parents and children share data with public services without giving it much thought at all.

There are many advantages to public sector bodies having more information about children. Children's data enables better planning at national and local levels, and helps make services tailored to the needs of the individual child. Furthermore, new technologies are being integrated into public service delivery as a result of collaboration between the public and private sectors, producing a diverse range of benefits. For example, many teachers now use apps such as Class Dojo to support their pupils' learning and behaviour. The growth of affordable biometric technology means that finger print scanners have become a common feature of many school canteens and libraries. In health, the Red Book known to generations of parents in its paper form is being digitised so that data can be accessed in real time by professionals.

And yet there are growing concerns in the academic and policy communities that our trust in public services with respect to children's data is misplaced – that there is no necessary reason to believe that public sector bodies are any better or worse than commercial organisations in terms of the standards they adhere to when handling children's data. Public bodies "do not always observe robust standards of privacy, transparency, security or redress".[12] Furthermore, despite the benefits, a clear implication of there being more public-private partnerships when delivering services to children is that more data about children is shared - often very sensitive data concerning their health or educational performance, and often without parents and children being fully aware. As Livingstone argues, it is more difficult to determine whether children's privacy and identity rights are protected in this context[13] as there are a greater number of actors involved and the relationships between children and families and those using the data less direct.

[12] https://www.childrenscommissioner.gov.uk/wp-content/uploads/2017/06/Case-for-general-comment-on-digital-media.pdf

[13] https://www.childrenscommissioner.gov.uk/wp-content/uploads/2017/06/Case-for-general-comment-on-digital-media.pdf

An example of a classroom app: Class Dojo

Class Dojo is a classroom app that is reported to have been used in more than 70% of schools in the UK. Using the app, teachers award positive Dojo points to children behaving well and negative points to those who misbehave. The app can also be used to communicate with parents – both teachers and parents can share written messages and photos, e.g. of the child or their schoolwork.

Teachers tell us that Class Dojo is an extremely valuable tool in the classroom. It can help them to engage children who are otherwise disruptive or disinterested, and makes the classroom a more fun learning environment for many children.

However, some concerns have been expressed about the implications of Class Dojo for data protection:

- Class Dojo does not require sensitive information to function. For example, children can be identified using nicknames rather than their real names. However some teachers use sensitive information anyway.
- Data is shared with 31 other organisations, each with their own privacy policies.
- If the company were sold, all ClassDojo data would come under its new owner's privacy policy. If concerned, parents would be responsible for deleting their child's data within 30 days.

Furthermore, there are worries that Class Dojo contributes to a practice where children are increasingly being monitored and tracked around the clock, which may impact upon their development and experience of childhood.

How might children's data be used – in the short term and long term? Why should we be concerned?

The fact that increasing volumes of data are being collected about children is clear. Less clear, however, is the impact of this: who is seeing the data, what are they doing with it, and to what effect on children's lives.

The benefits

Firstly, it is important to recognise the huge benefits there are to collecting greater volumes of data about people, making it more accessible and analysing it in new ways. Some of the benefits are described in a recent policy paper by the Treasury exploring the economic value of data:[14]

> "Data-driven innovation holds the keys to addressing some of the most significant challenges confronting modern Britain… data-driven innovation can have a significant impact on well-being, as well as productivity growth. Data can be used to personalise services and improve the consumer experience in areas like mapping, retail and video/music streaming. And it can form the basis of brand new products across a range of sectors – from unlocking new healthcare

[14] https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/731349/20180730_HMT_Discussion_Paper_-_The_Economic_Value_of_Data.pdf

treatments, to enabling smart devices. In the public sector, data is playing an increasing role in transforming public services."

Recent research by Reform which focuses on data sharing within the public sector produced similar conclusions. Given the general complexity of social issues, understanding individual needs can be complex. Data enables government to see beyond this complexity and understand individual needs, making services more personalised and joined up. For example, sharing data digitally between GPs and hospitals can enable early identification of patients most at risk of hospital admission, which has reduced admissions by up to 30 per cent in Somerset.[15]

Participants in our roundtable gave various examples of how increased volumes of data, data sharing and innovations in data processing could improve outcomes for children in particular. For example, inspections of services for children could focus on areas where the data suggests there are problems, ensuring greater accountability. Datasets such as the NSPCC's national case review repository can be analysed more quickly and in many different ways (e.g. through natural language processing which can be used to find common themes in large volumes of unstructured text), improving our understanding of how to prevent harm and promote positive outcomes. Last month it emerged that at least five local authorities now use predictive analytics involving data from children and adults to flag potential child safeguarding risks to social workers. Coverage emphasised the possibility of these systems to help councils target resources more efficiently at a time when local budgets are extremely stretched.[16]

An example of digitisation: the Red Book

The Personal Child Health Record (PCHR), more commonly known as the Red Book, is a national standard health and development record given to parents at the time of their child's birth. It is used by parents to record things like when their child reaches developmental milestones, any accidents they have, etc from age 0-5. Healthcare professionals also update it when they see the child.

The Red Book is often the fullest summary of a child's health and development journey there is. However, it is a paper record kept by the child's parents and therefore cannot be accessed independently by healthcare professionals. This means that GPs, Health Visitors and other professionals might not always have all the information needed to provide children with the best possible care.

To overcome this limitation, a new digital Red Book has been developed and is being trialled in some areas. It will enable professionals to access the record without relying on the parent. It offers benefits to parents by making it easier to update (e.g. when away from home), and offering a more personalised experience. It is also hoped that it might increase engagement among certain groups of parents – however, it is possible that it could reduce engagement if parents feel less ownership over the record.

Much of this data collection has clear advantages. Teachers have told CCO that apps such as Class Dojo make a real difference in the classroom, engaging pupils who might not otherwise want to learn.

[15] http://www.reform.uk/wp-content/uploads/2018/08/Sharing-the-benefits-how-to-use-data-effectively-in-the-public-sector.pdf

[16] https://www.theguardian.com/society/2018/sep/16/councils-use-377000-peoples-data-in-efforts-to-predict-child-abuse

Similarly, the digitisation of the Red Book could represent a real step forward in ensuring that all health professionals have the information they need to provide the right care to children at the right time, and engage parents as partners in the process.

The risks

The increasing availability of data offers enormous advantages, but it is crucial that we are mindful of the risks and mitigate them.

This is particularly the case in relation to children, who are typically less aware of the risks and consequences involved in the processing of personal data. This fact is cited in the GDPR (Recital 38) as grounds for children meriting special protection with regard to their data. Additionally, a research team led by Sonia Livingstone at the London School of Economics and Political Science[17] notes that there is a particular reason to be concerned by children's data privacy as they are often the first to adopt new digital devices, services and contents. In effect they are the "canary in the coal mine for wider society, encountering the risks before many adults become aware of them or are able to develop strategies to mitigate them."[18]

A fundamental challenge to exploiting the benefits of data while managing the risks is that we simply do not know what all of the risks are. The amount of data captured about all of us grows each day, and the rate of that growth becomes faster as technology becomes more developed. The OECD estimates that between 2010 and 2015 there was an eight-fold increase in the global volume of data. By 2020 it is expected that the proliferation of connected devices and other new technologies will increase the volume 40 times over.[19] And it's not just the volume of data which contributes to this rapidly changing picture: it is also developments in processing techniques, which enables analysts to read more and more into the data that already exists. We are all now datafied – but children growing up today are among the first to be datafied *from birth*. Children are following an untrodden path, and we cannot fully understand what the implications of this are going to be many years down the line.

Nevertheless, there is some recent and ongoing research into the risks involved in collecting children's data. While this does not give a full picture, it points to some potential challenges which might emerge. The fact that more and more researchers are beginning to look into the issue should be cause for concern in itself, as it suggests a general feeling of anxiety about the possible implications of children's data profiles – something which was reflected in our conversations with experts and roundtable.

Our conversations also revealed that a further challenge with understanding the possible risks in collecting data from and about children is a lack of transparency by those who handle it. If we better understood what happens to children's data after it is given – who collects it, who it is shared with and how it is aggregated – then we would have a better understanding of what the likely implications might be in the future, but this transparency is lacking. This is despite transparency being mentioned in the first key principle set out in the GDPR (Article 5), which states that data must be "processed lawfully, fairly and in a transparent manner in relation to individuals".

[17] http://www.lse.ac.uk/media-and-communications/research/research-projects/childprivacyonline
[18] http://www.lse.ac.uk/media-and-communications/research/research-projects/childprivacyonline
[19] https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/731349/20180730_HMT_Discussion_Paper_-_The_Economic_Value_of_Data.pdf

Short term risks – as children

Safety and wellbeing

Children and their parents are most aware of the risks posed by data misuse to a child's immediate (or short-term) safety and wellbeing, which may arise for instance through bullying, identity theft, information being seen by strangers or contact by people who wish them harm.

Qualitative research exploring children's perceptions of mobile media suggests that risks related to personal data are among children's major concerns.[20] Their concerns are well-founded: a 2011 survey by the EU Kids Online network found that 12% of children aged 11-16 in the UK had experienced personal data misuse in the previous 12 months. 10% said that somebody had used their password to access their information or pretend to be them. 4% said that someone had used their personal information in a way they didn't like, and 1% said that they had lost money by being cheated on the internet.[21]

Bullying and impersonation by other children can have a significant impact on a child, but even more worrying is the possibility of children's data being used by people who intend them more serious harm. In 2017 schools and police forces issued warnings about Snapchat's new Snap Maps feature,[22] which shows a map of user locations as collected by smartphone GPS sensors. Although only those on a child's friend list can see their location, it was warned that children often befriend people online who they do not know in real life, and that some might target children through the Snap Maps feature.

Child development and social dynamics

A less tangible impact of collecting data about children concerns their experience of childhood. Some experts have warned that a child's awareness of being monitored by their parents and teachers through connected devices and apps may have an impact on their development and family dynamics.[23] For example, pushing boundaries is a normal part of growing up. This may manifest itself in a number of ways: going further away from home than allowed by their parents, for instance, or looking at inappropriate content. But for children who are aware of being monitored – whether by having their location tracked with a tracker watch, their browsing history read or a 'spy' app – pushing boundaries is less possible. Collecting so much data about children raises important questions about their freedom and independence, even if it is collected with good intentions.

Concerns have also been raised that collecting personal information from children so regularly normalises the act of surveillance.[24] The risk is that children hand over their data so often that they become accustomed to doing so thoughtlessly and without hesitation, failing to question if or why it is

[20] http://netchildrengomobile.eu/ncgm/wp-content/uploads/2013/07/DEF_NCGM_SecondEdition_Report.pdf
[21] http://eprints.lse.ac.uk/33730/
[22] https://www.telegraph.co.uk/technology/2017/06/23/police-issue-child-safety-warning-snapchat-maps-update-reveals/; https://www.bbc.co.uk/news/technology-40509281
[23] For example see: https://academic.oup.com/iwc/article-abstract/25/3/204/874906?redirectedFrom=fulltext
[24] https://www.theguardian.com/sustainable-business/2016/feb/19/surveillance-state-fingerprinting-pupils-safety-privacy-biometrics

needed or how it might be used. It has been argued that a lax attitude towards sharing information can begin to develop in children as young as six.[25] It is a perilous habit for children to develop at an early age, especially given that it might persist into adulthood, when people are afforded fewer protections in data protection legislation than children. Essentially, collecting so much data from children sends the wrong message – it does not convey how valuable and sensitive personal information is and how important it is to guard it.

Furthermore, there are concerns that having so much data at their fingertips may increase parents' and teachers' expectations of their children, at a time when children already face enormous pressures growing up. Commenting on the growth of the Internet of Things (IoT), Dr Renee Singh (co-director of the Tavistock and UEL Family Therapy and Systemic Research Centre) says:

> "When the quantified self comes into the home, schools would have the potential to track children's reading speeds, sleep levels, bowel movements and other usually private data…The result may be more envy and competitiveness between siblings and higher expectations from parents…Schools may be hypercompetitive, but at the moment parents can be protective of the home as a sanctuary and keep the home free from this atmosphere. If the school is tracking how long it takes a child to complete homework, and who is doing it, this separation looks less possible."[26]

Campaigning organisation 5Rights has called for the Government to fund new research into the developmental implications of living digitally from infancy,[27] to understand more about the consequences of the growth of data and widespread use of new technology for children's development.
Technology also facilitates changes in parent behaviour which could impact negatively on children. The use of GPS tracker watches and other connected devices or apps, for instance, can lull parents into a false sense of security – after all, children can take tracker watches off. And there have been warnings that connected devices such as talking toys are no replacement for quality parent-child interaction.[28]

Long term – as young people and adults

Identity theft and fraud

With so much data being collected about today's children, they will be at an increased risk of identity theft and fraud as they grow up.

There is particular concern about "sharenting" – social media updates by parents – which might reveal more information about children than intended. According to Barclays, there are three key pieces of information used in identity theft: a person's name, date of birth and home address. These are often

[25] https://5rightsframework.com/static/Digital_Childhood_report_-_EMBARGOED.pdf
[26] https://www.designcouncil.org.uk/news-opinion/will-internet-things-set-family-life-back-100-years
[27] https://5rightsframework.com/static/Digital_Childhood_report_-_EMBARGOED.pdf
[28] https://www.theguardian.com/sustainable-business/2016/mar/29/smart-toys-lazy-parents-internet-of-things-hello-barbie

given directly by parents, or can be deduced from photos or updates on social media accounts – for example, a photograph of a child on their birthday with a location tagged might give all this personal information away.

With this information, criminals can make a start on accessing bank accounts or making credit applications. At our roundtable CCO heard reports of children's data being stored until they turn 18, at which point fraudulent loans and credit card applications were made. Further information such as a mother's maiden name, names of pets and names of schools might also be gathered through a parent's social media account, making it even easier to commit fraud given that these details are often used as security questions. Barclays has forecast that by 2030 "sharenting" will account for two-thirds of identity fraud facing young people over 18 and will cost £667 million per year.[29]

Impact on opportunities and life chances

Perhaps the most disconcerting risk associated with growing volumes of data about children, and new ways of processing it, is that it could shape children's long-term opportunities and life chances in ways that are unjust.

Data gathered in childhood could form part of the evidence used to profile people. Profiling is a process in which data about a person is analysed using algorithms and machine learning "to analyse or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements." The profile might be used by organisations in three ways:[30]

- To find out something about an individual's preferences. For example, advertisers might use profiling to target their products at certain people.
- To predict their behaviour. For example, employers might use profiling to predict how likely someone is to perform well at a job.
- To make decisions about them. For example, banks might use profiling to decide whether to allow a person to take out a loan or not.

Profiling is already used widely. One of its most significant applications is in advertising. Data about personal characteristics (e.g. age, gender) and browsing history are analysed to infer the products an individual might be more likely to buy, which are then promoted to the individual through online advertising.

Advertising might be thought of as a low stakes application of profiling. It affects which products a person sees when they go online.[31] But profiling is beginning to be used in "high stakes domains" – to

[29] https://www.bbc.co.uk/news/education-44153754

[30] https://ico.org.uk/for-organisations/guide-to-the-general-data-protection-regulation-gdpr/automated-decision-making-and-profiling/what-is-automated-individual-decision-making-and-profiling/

[31] Although note that we might think of advertising as a high stakes domain with respect to children, given that they are more susceptible to advertising messages and may make poor decisions as a result, e.g. purchase products that they cannot afford. See: https://ico.org.uk/for-organisations/guide-to-the-general-data-protection-regulation-gdpr/children-and-the-gdpr/what-if-we-want-to-target-children-with-marketing/

make decisions about whether a person is granted bail, whether they will be offered a job and whether they will get credit.[32] Sometimes a person is also involved in making these high stakes decisions, using the data profile to assist them – a process known as "semi-automated decision-making". In contrast, "solely automated decision-making" is a process whereby the decision is made without any human involvement, and there are examples of this happening in high stakes domains.[33] Predictive analytics are now also being used by children's services departments in an attempt to identify potential child safeguarding risks, albeit with oversight by a social worker who reviews the case when a risk is flagged. In these areas, profiling can have a significant impact on someone's life.

For children growing up today, and the generations that follow them, the impact of profiling will be even greater – simply because there is more data available about them. Profiling relies upon data, and data about children is now routinely collected by a wide range of organisations – and can be sold on – from birth (sometimes pre-birth). Some of that data could find its way into an individual's data profile, and be used to make highly significant decisions about them. For example, a child talking about their mental health problems on social media might find that this information hinders them from getting health insurance or perhaps even other types of credit. Colleges and universities might use not only exam results and personal statements to award places, but data collected from educational apps or connected devices. Essentially, children's digital identities, being created now, could have a long-lasting impact on the shape of their lives for many years to come.

The process is being fuelled by greater sharing, combining and linking of distinct datasets. The result is increasingly rich profiles that give insight into many different aspects of a person's life, and consequently their future behaviour and preferences. As profiles become more detailed, so too the pressure increases for them to be used to make high stakes decisions.

Aside from the basic injustice of actions and events in childhood determining life chances as adults, there is a further injustice in that some of these processing techniques are blunt instruments which cannot capture the full picture of who a person is and their potential. The possible risks, along with the opportunities, of automated decision-making, profiling and related developments are recognised widely, including by Government. For example, in a consultation[34] on the new Centre for Data Ethics and Innovation, it says:

> "enhanced decision-making through artificial intelligence can radically improve outcomes for society, including through more effective targeting of public resources and commercial products and services. However, automated decision-making can be opaque and, in certain contexts, may lead to unfair outcomes or overly restrict the level of control we have over the decisions that shape our lives. For example, job applications may be rejected without clear explanation or automated tools might exacerbate or reproduce inequities within the criminal justice system."

Profiling relies upon algorithms. An algorithm takes an input (i.e. the data), follows a series of steps and produces an output. When profiling is used in solely automated decision-making, the output is the

[32] Carl Miller, *The Death of the Gods: The New Global Power Grab* (2018)
[33] As above
[34] https://www.gov.uk/government/consultations/consultation-on-the-centre-for-data-ethics-and-innovation/centre-for-data-ethics-and-innovation-consultation

decision, e.g. whether an individual gets a job, credit, a university place, bail, etc. Making decisions in this way offers significant advantages: it can deal with vast volumes of data quickly and has the potential to eliminate human bias, producing more consistent results.

Compared to human decision-making processes, algorithms can be "unfairly reductive".[35] For example, a person might miss paying a bill or a fine because they were in hospital, but an algorithm would simply record the missed payment.

> "And therein lies the urgent challenge facing all of us in the digital world… If life-determining algorithms are here to stay, and it certainly looks that way, we need to figure out how they can embrace the nuances, inconsistencies and contradictions inherent in human beings. We need to work out how they can reflect real life".[36]

And algorithms can have their own biases. Algorithms are created by people and are trained using data selected by people. Algorithmic bias occurs when algorithms are created and trained in such a way that their results reinforce human biases.[37] For example, it has been reported that Amazon recently scrapped a recruitment tool that used machine learning because it was found to be discriminating against women – a result of it having been trained to vet applicants by observing patterns in applications received by the company over a ten-year period. Most applications came from men, reflecting the male dominance within the industry.[38]

The lack of transparency over algorithms means that decisions or outcomes resulting from profiling and automated decision-making cannot be understood or, very importantly, challenged. The problem will only become worse as algorithms become more complex. Algorithms increasingly exist within webs[39] – the output of one feeding into the inputs of another – meaning that it is not clear, even to those who have designed the algorithms, what processes are unfolding within them.

One participant in our roundtable proposed the introduction of algorithm safety audits to scrutinise their performance and legitimacy.

What is being done?

Data privacy has climbed up the political agenda in recent years – partly in response to the vast volumes of data now collected and shared. Yet policy is struggling to keep up with the pace of developments, and with uncertainty over future regulation, analytical techniques and technological advances, our understanding of the implications for today's children remains limited. Early findings from a project being led by Professor Sonia Livingstone suggest that children themselves see privacy overwhelmingly in terms of interpersonal privacy – i.e. the data they share directly with other people.

---

[35] Rachel Botsman, *Who Can You Trust?: How Technology Brought Us Together – and Why It Could Drive Us Apart* (2017)

[36] Rachel Botsman, *Who Can You Trust?: How Technology Brought Us Together – and Why It Could Drive Us Apart* (2017)

[37] The report published following the Science and Technology Committee's inquiry into algorithms in decision-making sets out four key sources of algorithmic bias: inappropriate training data, insufficient data, confusion of correlation with causation, and lack of representation in the algorithm development community. https://publications.parliament.uk/pa/cm201719/cmselect/cmsctech/351/351.pdf

[38] https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G

[39] Carl Miller, *The Death of the Gods: The New Global Power Grab* (2018)

They are less aware of institutional dimensions of privacy, such as the data that schools have, and how it might be used, and commercial privacy such as the data stored and used by social media companies There are several current research projects examining these issues. The Oxford Internet Institute is mapping the range of organisations which engage with children digitally, beyond big technology companies – e.g. manufacturers of connected devices and public services such as schools. The project is exploring whether the actions of these organisations might be exposing children to online risks, including data theft.

A project led by Professor Sonia Livingstone at the London School of Economics and Political Science is exploring children's conceptions of privacy online, supported by the Information Commissioner's Office. It aims to address questions and evidence gaps concerning children's capacity to consent to their data being used or shared, their functional skills (e.g. in understanding terms and conditions or managing privacy settings online), and their deeper critical understanding of the online environment, including both its interpersonal and, especially, its commercial dimensions. The commercial angle will explore the business models of online platforms and the nature of algorithms.

Research such as this will help fill the gaps in our understanding of how children's data might be used in the future, and how we can exploit the benefits while protecting against the risks.

Not all data is equal. It is clear that some of the data sources highlighted in this report are of much less concern than others – perhaps because we have greater trust in those who collect it and their reasons for doing so, such as medical records being created by health professionals. But for some there is clear cause for concern. Not just because of the short term risks to a child's safety, but also for the possibility of much deeper, long-term impacts on a child's life and opportunities. The key message from our research is one of uncertainty: we do not understand how data collected now might be used in the future.

It is therefore unsurprising that policymakers are struggling to keep up with developments. Broadly speaking, GDPR is a step in the right direction – it at least recognises that children have specific vulnerabilities and merit special protection with regard to their data privacy, even if it does not – and cannot, at this stage – say what those protections need to be, nor how they will need to change as technology evolves. But it does not address the most fundamental, long-term challenges that might be posed by increasing volumes of data. Government, industry, regulators and others will need to be ready to respond quickly as our understanding of these challenges develops.

Policy

GDPR

The most significant recent development in data privacy has been the implementation of the General Data Protection Regulation (GDPR) this year. In the UK its provisions were incorporated into law and further developed by the Data Protection Act 2018.

Specific protection for children

GDPR makes some notable improvements to children's privacy rights and protections. Most fundamentally, unlike its predecessor (the Data Protection Act 1998) it requires that there are specific protections in place for children. Recital 38 of the GDPR makes clear that because children are likely to be "less aware of the risks, consequences and safeguards concerned and their rights in relation to the processing of personal data" they need to be protected by additional safeguards.

Provisions related to profiling and automated decision-making

The long-term risks posed by automated decision-making, facilitated by user profiles and algorithms, were discussed in the previous section. GDPR demands particular protections for the use of children's data "for the purposes of marketing or creating personality or user profiles and the collection of personal data with regard to children when using services offered directly to a child."[40] Furthermore general protections which apply to all data subjects apply to children as they do to adults.

- Article 22 states that children have the right not to be subject to decisions based solely on automated processing (including profiling) if they have legal or similarly significant effects on them.[41]
- Where a solely automated decision-making process *is* used, and it produces legal or similarly significant effects on the individual (whether adult or child), the individual must be told that it is happening, about the logic involved and its significance.[42] They have the right to ask for human intervention, to express their point of view and challenge the decision.[43]
- Article 21 gives data subjects (both adults and children) the right to object to profiling that is related to direct marketing.
- Recital 71 indicates that profiling with respect to children should not be the norm.[44]

However, this protection is limited in at least the following ways:

- Although Recital 71 suggests that profiling is not the norm with respect to children, the practice is not prohibited outright.
- The provisions only apply to solely automated decisions (when no human is involved in making the decision whatsoever). They do not apply to decision-making where humans play some role, however minimal that role is.
- Determining whether an automated decision-making process will have "similarly significant effects" is difficult to gauge given that we do not yet understand the full implications of these processes – and perhaps even more difficult to judge in the case of children.
- There is still great uncertainty about how Article 22 and profiling will work in respect of

[40] Recital 38, GDPR
[41] Article 22, GDPR
[42] Articles 13 and 14, GDPR
[43] Article 22, GDPR
[44] https://ico.org.uk/for-organisations/guide-to-the-general-data-protection-regulation-gdpr/children-and-the-gdpr/what-if-we-want-to-profile-children-or-make-automated-decisions-about-them/

children. The key area of concern will be in respect of any limitations in relation to advertising products and services and associated data protection practices.

Additional important provisions

Other important provisions which go beyond the 1998 Act include:

- The right to be forgotten (article 17) - adults and children now have the right to have their personal data erased in some specified circumstances where, although the original collection and processing may have been compliant with the GDPR, they no longer with the personal data to be held. There is a particular expectation that data given by children is erased, especially if it seems they did not fully understand the implications of giving it – however, the right to erasure is not an absolute right and can be overridden.

- Age appropriate privacy notices (recital 58) – there is now a requirement that privacy notices must be provided to children in "clear and plain language that the child can easily understand."

- Consent (article 8) – outside of the UK in several countries it is necessary for an online enterprise to obtain parental consent before they can allow a child to be a member of their site or service. In the UK, children who are 13 or above can give consent themselves to the processing of their personal data in the context of an information society service (ISS) offered directly to them, without the need for the ISS to engage with a parent to secure their consent. Examples of ISSs include online shops and marketplaces, social media platforms and streaming services. Note, however, that consent is not the only lawful basis for processing children's personal data and while these provisions only apply to online services offered *directly* to children, it is clear that any service which permits, allows or encourages anyone under the age of 18 to be a member of their site or service is likely to be considered to be offering their service directly to a child.

  Some services may specify a minimum age of consent which is greater than 13. For example, WhatsApp specifies a minimum age of 16, but here the company may proceed without obtaining the consent of the child's parent if they utilise "legitimate interests" as their basis for processing the child's data, as per Article 6(f).

The age-appropriate design code

GDPR recognises that children warrant special protection in relation to their data but does not describe in detail what that entails in practice. The age-appropriate design code is designed to fill this gap. A product of an amendment to the Data Protection Act, the code will be produced by the Information Commissioner Office, who recently consulted on its contents. The code will set standards on various aspects of design including data minimisation, default privacy settings, language of privacy notices, sharing and resale of data and automated and semi-automated profiling.

The code will be statutory. Organisations who ignore it risk fines of up to £18 million or 4% of global turnover.

DCMS Security by Design project

There are increasing concerns among children's rights organisations, the Information Commissioner's Office and DCMS about the privacy and safety risks associated with connected devices. In March 2018 DCMS published a report setting out its vision for Security by Design, "a fundamental shift in approach

to moving the burden away from consumers having to secure their internet connected devices and instead ensure strong cyber security is built into consumer IoT products and associated services by design." The report set out a draft Code of Practice for manufacturers of consumer IoT products and associated services, with the final Code launched in October 2018. The Code is currently voluntary but DCMS has stated that it will be made compulsory by law if it is not adhered to. Although the development of a Code has been welcomed by security experts, it has been argued that that the measures suggested would not have prevented many of the recently reported security breaches of smart devices.[45] Furthermore the Code makes no specific reference to children.

The government has stated in its Internet Safety Strategy green paper that it will consider where the lessons from the review might inform future work on connected toys specifically.

Internet Safety Strategy

The Government's central response to online harms is the Internet Safety Strategy. Currently under development, a green paper was published in October 2017 with a white paper to follow this winter. The white paper is being developed jointly by DCMS and the Home Office and will cover the full spectrum of legal and illegal harms.

The strategy is wide-ranging, addressing a diverse range of challenges including data privacy but also exposure to inappropriate content, cyberbullying and mental health risks. In relation to data, key provisions include the social media code of practice, providing greater support to parents and building children's digital literacy. The strategy will also address technological solutions to internet safety problems.

The Science and Technology Committee's inquiry into algorithms in decision-making

In 2017 the Science and Technology Committee launched an inquiry into the use of algorithmic decision-making in the public sector and business. The aim of the inquiry was to explore the extent of current and future use of algorithms in decision-making, to identify examples of good practice in terms of eliminating algorithmic bias and promoting transparency, and to explore methods for providing regulatory oversight of algorithmic decision-making. The report from the inquiry was published in May 2018. While the report recognises the many benefits of incorporating algorithms into decision-making (including in health, criminal justice, social media and government data-sharing), it makes a strong case for the need to reduce bias and boost accountability and transparency, identifying a series of recommendations aimed at achieving this. Particular emphasis is placed on the role of the forthcoming Centre for Data Ethics and Innovation in providing regulatory oversight, along with the Information Commissioner's Office.

Centre for Data Ethics and Innovation

In the 2017 Autumn Budget the Chancellor announced funding to support the creation of a new Centre for Data Ethics and Innovation. The Centre will bring together representatives from regulating bodies, academia, business and the public to identify the measures needed to strengthen and improve the way data and AI are used and regulated. It will articulate best practice and advise the Government on specific policy or regulatory action.

[45] https://www.bbc.co.uk/news/technology-43305346

It has been proposed that the Centre focuses its work in six areas, including targeting, fairness and transparency – all issues raised in our discussion of the possible risks associated with children's data. For example, discussing fairness, the Government states:

> "Algorithms make use of data about past behaviour, which means biases embedded in the data can be reinforced and strengthened over time."

And considering transparency, it notes:

> "Data technologies have the potential to significantly augment human cognition. However, the decisions and recommendations they offer may not be easily interpretable or explainable. This raises questions about the extent to which we need to be able to explain decisions in different contexts and, ultimately, when and to what extent we should retain human control over decision-making."[46]

What else needs to happen?

Recommendations for policy and practice

- There is increasing recognition within digital policy that children have particular needs and therefore warrant special consideration and protection. In keeping with this welcome development, the new **Centre for Data Ethics and Innovation** should undertake a programme of work specifically focused on children.

- CCO supports the Science and Technology Committee in their recommendation that the **Centre for Data Ethics and Innovation** and **ICO** review the operation of GDPR by May 2019. The rapid pace of technological change means that swift regulatory action may be needed in order to protect children from being disadvantaged by the way their data is used, especially with regard to profiling and automated (and semi-automated) decision-making. The **Government** must respond quickly to this review, refining data protection legislation if necessary.

- **The Government** should consider introducing an obligation on those using automated decision-making to be more transparent about the algorithms they use and the data fed into these algorithms, where data collected from under 18s is used.

- **Companies** producing apps, toys and other products aimed at children should be more transparent about any trackers capturing information about children. In particular where a toy collects any video or audio generated by a child this should be made explicit in a prominent part of the packaging or its accompanying information. It should be clearly stated if any video or audio content is stored on the toy or elsewhere and whether or not it is transmitted over the internet. If it is transmitted, parents should also be told whether or not it will be encrypted during transmission or when stored, who might analyse or process it and for what purposes. **Parents** should ask if information is not given or unclear (see top tips for parents and children below).

- **Companies** should state their terms and conditions using language children can understand,

[46] https://www.gov.uk/government/consultations/consultation-on-the-centre-for-data-ethics-and-innovation/centre-for-data-ethics-and-innovation-consultation

explaining clearly what data is collected and how it will be used.

- **Schools** should teach children about how their data is collected and used, and what they can do to take control of their data footprints. These lessons should cover information shared online but also information gathered in the home (e.g. through connected devices) and outside the home (including through public services). CCO encourages schools to use our infographic and top tips to help.

Finally, the Children's Commissioner believes there should be a **statutory duty of care** governing relationships between social media companies and the audiences they target and will be working with the law firm Schillings to draft one.

Ten top tips for children and parents

While much more needs to be done by Government and industry, our message to children and parents is that they are not powerless in this situation. There are steps they can take, now, which will significantly reduce a child's data footprint – some of which are very simple. CCO has brought the key steps together below.

For children

1. **Stop and think** when you're about to share some personal information. Ask yourself, "Do I *need* to share this"? If you can't do what you want (e.g. play a game) without giving away this information, ask yourself, "Is it worth it?" – sometimes it is, but lots of times it isn't.

2. **Read our Digital 5 A Day guide** if you spend lots of time online and on social media, to help you think about other ways you can spend your time: connect, be active, get creative, give to others and be mindful.

3. **Look through terms and conditions** to understand what data is collected when you use social media, websites and gadgets. We've simplified some here.

4. **Mute smart speakers** when you don't want them to listen to you.

5. **Talk to an adult you trust** if you are worried about someone else knowing something about you, or if you want to learn more about your data rights.

For parents/carers

1. **Don't post photos and videos which reveal personal information about your children online.** Sometimes it isn't obvious – for example, tagging a child at home on their birthday gives away their date of birth and home address.

2. **Change the default passwords on all the gadgets your children use** – whether it's a smart speaker, internet-connected toy or location-tracking watch. Don't forget the router!

3. **Make sure the gadgets you buy your children are genuine.** Counterfeit versions can be less secure than the originals.

4. **Watch out for security updates** and install them as soon as you are prompted.

5. **Talk to organisations that hold information about your child** about what information they collect and why, including schools, online services and retail loyalty schemes. Raise any concerns you have.

Children's COMMISSIONER