Issue brief no. 4 | October 2020
Good Governance of Children's Data project
Office of Global Insight and Policy

unicef
for every child

Responsible group data for children

Andrew Young, The GovLab

Introduction

In 2007, the Eyes on Darfur campaign collected satellite imagery and the associated metadata to help the humanitarian response to the War in Darfur. This data informed response efforts but at an unexpected cost. A subsequent analysis of the campaign showed that villages monitored by Eyes on Darfur were more likely to be attacked by militants.[1] The data released to the public provided insight into the impacts of the conflict on different regions, but did not feature any personal information about individuals in targeted villages.

This story is an example of the under-scrutinized risks of group data. While the data protection field largely focuses on individual data harms, it is a focus that obfuscates and exacerbates the risks of data that could put groups of people at risk, such as the residents of a particular village, rather than individuals. Though not well-represented in the current responsible data literature and policy domains writ large, the challenges group data poses are immense. Moreover, the unique and amplified group data risks facing children are even less scrutinized and understood. To achieve Responsible Data for Children (RD4C)[2] and ensure effective and legitimate governance of children's data, government policymakers, data practitioners, and institutional decision makers need to ensure children's group data are a core consideration in all relevant policies, procedures, and practices.

What is group data?

Groups can take many forms. Individuals can be grouped together based on characteristics including:

1. common demographic traits such as ethnic background, eye colour, or genetic makeup;[3]
2. associations between people, such as members of a certain religion or political party;
3. shared geo-location; or
4. in humanitarian settings, groups can be formed based on a common threat of harm or a similar type of privacy interest.[4]

In addition to these groups, which are based on individual traits, groups can also be formed as the result of the way technologies work. Data-driven technologies, such as artificial intelligence (AI) and machine learning, do not identify individuals but types of individuals.[5] This group, the designed group, includes all those people the algorithm has been trained to think of as meaningful according to criteria defined by its creators. The contours of these designed groups only exist as a result of certain decisions made during the design of the algorithm or the analysis of data.[6]

Online advertisers, for example, could use algorithms to create a group featuring males between the ages of 18 and 35 who have demonstrated inter-

est in video gaming, mixed martial arts, and a particular political candidate. Advertisers could then micro-target messaging to members of that group that are more likely to pique their interest. Despite their shared interests, the members of this group were never brought together prior to advertisers' creation of this marketing segment.

While privacy infringements are likely to take place at the group level… rights to redress and rectification are granted almost exclusively at the individual level

Groups can also be structured in several ways. Luciano Floridi outlines how groups subject to data analytics could be real or artificial, self-proclaimed or framed, self-aware or not, stable or fluid, and hierarchical or egalitarian.[7] The specific components of a group can have major implications on the particular sensitivities and types of risk they face.

Demographically identifiable information

This paper uses the umbrella term of "group data" also in relation to demographically identifiable information (DII). *The Signal Code*, the result of a six-month study at the Harvard Humanitarian Initiative, defines DII as "data points that enable the identification, classification, and tracking of individuals, groups, or multiple groups of individuals by demographically defining factors. These may include ethnicity, gender, age, occupation, and religion."[8]

Much of the current academic and policy literature on the risks of group data focus on algorithmically defined and generated groups as well as increasingly on the impacts of AI on such groups. The term DII, on the other hand, is more common among humanitarian and development actors where demographic categorization can support and enable service delivery, monitoring and evaluation.

Unlike many algorithmically generated group datasets, DII tends to correlate more closely with real-world groups — such as those with a common ethnic background. Still, decisions made during data analysis impact the characteristics of a group. Group composition and level of risk will differ if the analysis focuses on individuals with a certain medical condition or occupation living in a certain province compared to a similar analysis conducted at the neighbourhood or village level, for example.

Group data and the responsible data policy ecosystem

The emerging literature on group data highlights the mismatch between the current thinking on data responsibility as a largely individual-level concern, and the reality that groups or *types* of individual are often most at risk. This focus on the risks around individual-level data, to the detriment of group data responsibility, exists in data protection regimes, the recourse and avenues for redress afforded to parties harmed by data use, and the data rights provided to people represented in institutional datasets.

General Data Protection Regulation (GDPR), the most influential data protection regime at the time of writing, is, as Martin Tisné notes, "premised on a relationship between data controllers and data subjects".[9] Group data, however, makes it difficult to identify which parties occupy these roles and define the relationship between them, thus negatively impacting the effectiveness of data protection policies, including but not limited to GDPR.[10]

These problems are further exacerbated should a violation of data rights occur. While privacy infringements are likely to take place at the group level or impact an individual due to their group-based identity profile, rights to redress and rectification are granted almost exclusively at the individual level.[11] Although some thinkers and organizations are pushing toward a more multifaceted data protection landscape providing both collective and individual data rights,[12] current practice still lags. It will be difficult for groups to meaningfully establish and act upon such collective rights, particularly in cases of passive or ad hoc group formation through data analytics.[13]

Group data considerations are also absent from many, but not all, policy frameworks and guidance. A brief produced by the UN Office for the Coordination of Humanitarian Affairs (OCHA) in 2016, *Building Data Responsibility into Humanitarian Action*, includes the key message that the "disclosure of sensitive personal and demographic data in the humanitarian space can lead to already vulnerable individuals and communities being further harmed or exploited."[14] The focus on both personal and demographic data extends across the brief's

recommended policies and practices. The *510 Data Responsibility Policy* similarly calls for data holders to determine whether and how DII will be used at the initiation stage of any data project.[15]

The Signal Code explicitly defines individuals' right to agency to include not just their personally identifiable information, but also their DII. It also declares that "care be taken" and "additional protections" afforded to both "persons or groups" facing particular threats to their right to privacy and security.[16]

The Council of Europe provides recommendations for data-driven group profiling to ensure responsibility and trustworthy data use. Those recommendations include keeping humans "in the loop" when using algorithms to sort and profile individuals and to diagnose the ethical and social impacts of profiling in addition to more traditional data protection concerns, such as avoiding data breaches or unauthorized access to data.[17]

The Government of India has also taken steps toward establishing a regulatory framework to support responsible handling of "non-personal data", which includes aggregated or anonymized personal data.[18] However, questions remain regarding its impact on business competition,[19] workability in practice,[20] and its exclusive focus on de-anonymization or re-identification risks, with less consideration of risks to groups themselves.

Scholars such as Alessandro Mantelero make the connection between group data and privacy concerns and collective rights in international law.[21] Collective rights for indigenous peoples[22] and minorities (defined by "national or ethnic, cultural, religious and linguistic identity"[23]) are increasingly well established in international law, and could help to improve responsible handling of DII in certain contexts. But while the potential for enshrining collective rights in international law is evident, these approaches are only beginning to emerge as it relates to group data concerns.

Unknowns, persistent challenges and risks

Governments and institutional decision makers face a number of challenges in creating effective mechanisms for mitigating group data risks facing children. Some key challenges are outlined here.

Unclear risk profile for aggregated, statistical data or big datasets

Data holders and policymakers face significant barriers to anticipating group data risks and harms. This difficulty in capturing risk and making decisions to help mitigate them is a result of the shifting composition of groups and the shifting components of group data analysis. The risk profile is clear for highly personal information about children, such as case-level records or granular geolocation information.[24] The risks associated with group datasets, on the other hand, can be less evident and less likely to factor into responsible data decision-making within data-holding or governing institutions. The risk profile for group data is likely to remain opaque as long as greater documentation of critical incidents is lacking; or as Brent Mittelstadt puts it, until a group data system "fails spectacularly and in public".[25]

Governments and institutional decision makers face a number of challenges in creating effective mechanisms for mitigating group data risks facing children

Mosaic effect and challenge of defining "sensitive data" in a diffuse data ecosystem

The concept of a singular and influential data controller is especially misaligned with the current group data ecosystem given the so-called "mosaic effect". The mosaic effect refers to the compilation of disparate, often publicly accessible, datasets to create new and potentially sensitive insights. As is the case with much of the data responsibility and privacy literature and policy ecosystem, actors have primarily focused on the mosaic effect's capacity to identify individuals using apparently low-risk datasets, potentially exposing them to harm. Latanya Sweeney has demonstrated that this effect can cause risk to individuals in various contexts. In 2000, she found that 87 per cent of the US population could be uniquely identified with no more information than their zip code, gender, and date of birth.[26]

Uncoordinated actors can expose people accidentally through the release of seemingly innocuous data, but these same factors can be exploited by malicious actors intentionally. What's more, the

mosaic effect further obscures the risk profile of group datasets. The result is that data holders and policymakers face a near-impossible task of determining which types of group data are too sensitive to collect or make accessible absent a clear understanding of what risks could emerge by combining that information with other, unnamed datasets.

Perhaps most troublingly, mosaic effect risks are only likely to metastasize as seemingly benign data, as described by Linnet Taylor, continues to "spread and multiply, becoming ever more linkable, mergeable".[27] Every year, data on children and groups of children are being collected to an unprecedented degree, amplifying their risk of mosaic group data harm.

Emergent groups, challenges of pre-emptive protective action and meaningful redress

Groups can be algorithmically created by various actors for various purposes. Policymakers and the public often lack insight into these group formation processes. Given this large and complex ecosystem of data-driven group formation, policymakers face difficulty in codifying fit-for-purpose responsible group data policies and in creating redress and rectification mechanisms for group members subject to harm.

> Children's limited capacity for agency… serves to disempower them to an even greater degree

Moreover, as discussed above, groups are often established through data analytics and the segmentation choices made during analysis. As a result, many group data subjects are not aligned with established protected classes or demographic attributes. This complicates the use of traditional anti-discrimination provisions and policies in a group data context.[28]

Individual-level privacy harms or critical incidents, if identified in a transparent and accountable way, lend themselves to redress mechanisms. When an individual is the locus of harm, there is little doubt regarding who the beneficiary of any redress mechanism should be. In the case of group harm, it is far less clear who can and should speak for the group and who is eligible and empowered to seek redress. Algorithmically defined groups are particularly challenging in this area given the opaque and shifting criteria of group membership and sorting.[29]

Responsible group data for children

Group privacy issues are increasingly recognized as important and under-scrutinized components of the data responsibility ecosystem writ large. There is a small but growing body of research and practice in the realm of group data risk assessment and mitigation. Little of this emerging body of knowledge, however, focuses uniquely or primarily on the group privacy risks facing children. This evidence gap is important due to the unique ways in which group privacy risks and challenges affect young children, adolescents, and teenagers. This section introduces a number of these challenges or concerns related to the handling of children's group data.

Groups of children struggle to exert agency

In many cases, individuals included in a group struggle to exert agency in the context of the group. An individual receiving a government service in a particular geolocation, for instance, is unlikely to be aware that their data are included in an institutional analysis of service-delivery effectiveness by region, let alone possess any capacity to stop, influence, or seek redress from that analysis.[30] Not only does this individual lack the ability to influence the analysis, but their representation as a member of the group is also influenced by actions beyond their control. In his analysis of group privacy, Mittelstadt examines group data concerns in relation to "identity tokens", which are "distributed across members of a group".[31] An individual's identity token is likely to be created without their control, and the actions of other individuals assigned the same identity token can change the character of and subsequent engagement with that token. In other words, an individual's data profile can be altered by the actions of other individuals with the same data profile (for a given analytical use case). This leads to what Mittelstadt calls a "shared ownership of identity".[32]

Indeed, the process of aggregating individual, case-level information into a group dataset serves to remove the types of identifiers that would allow an individual to retain some level of control over their information.[33] Thus data aggregation techniques intended to minimize data privacy risks can negatively impact the ability of individuals to exert control over the use of their information or seek redress from inappropriate uses.

As Martin Tisné puts it, "we are prisoners of other people's consent".[34] Adults face significant impediments to exerting individual or collective agency. Meanwhile, children are subject to other people's consent and decision-making. Levels of agency differ between children of different ages, with teenagers, for example usually more able to exert agency in comparison to younger children. But every individual, regardless of age, is disempowered in a group data context. Children's limited capacity for agency, added to the fundamental barriers to agency in a group data context, serves to disempower them to an even greater degree.

Children are, by definition, part of a vulnerable group

Across contexts, data profiles can be grouped together in innumerable ways depending on the types of data available and the level of abstraction used in the sorting and analysis. Each of these groupings would be subject to some level or type of risk — from the vague and unlikely risks to the immediate and actionable. The level of vulnerability in a data grouping varies significantly according to the types of profile or "identity tokens" grouped together for analysis.

While the specific vulnerabilities of a group are context-dependent, all data groups comprising the profiles of children have persistent vulnerabilities. Groups of children are subject to the particular risks created by any data grouping or profiling exercise *and* the inherent vulnerabilities of childhood. Depending on their age and particular circumstance, children may have more limited cognitive or developmental capacity compared to older people represented in group datasets. Younger children in particular also rely on parents or caregivers to help meet their basic needs and protect them from malicious actors. Children also lack many important legal rights that could help them exert agency and seek redress in cases when they are wronged by others.

Researchers, practitioners and policymakers are only just beginning to support group privacy efforts and better address group data risks writ large. The unique considerations of group data have not, to date, been the subject of much focused research, practice, or policymaking. Children, as a group, are subject to unique risks and children's group data warrants additional duties of care. These additional, child-specific risks and duties are not well-represented in current policy and guidance, compounding the risks and challenges present in a group data context.

Certain groups of children have additional, specific vulnerabilities

The severity of group data risks can accumulate as the sensitivities or vulnerabilities facing individuals represented in the group compound. The analysis of aggregated data on individuals' movement patterns in a location, for example, will create some level of group data risk. If the individuals in question are predominately children, the risk of, for example, human traffickers pinpointing the location of groups of children can be amplified through the data analysis. If those children are on the move because they are unaccompanied refugees crossing into a new country, the risks they face continue to accumulate.

Today's children are the first generation growing up at a time of rapid datafication

To take another example, if a public agency conducts an analysis of all two-person households in a particular village, there is a risk that people might use the granularity of that dataset to identify a particular resident. This exposure only exacerbates if the data narrows to focus specifically on two-person households with a child at home. If the analysis is further narrowed to two-person households where a child is the head of household, malicious actors could use the data to identify areas where highly vulnerable children are clustered. The risks inherent to a generalized analysis of two-person households are not replaced by the risks of analyzing child-headed households, those risks amplify and metastasize.

Group data and group data-derived profiling of children can enable demographic-based discrimination,[35] lead to stigmatization of children represented by a particular profile, and, most troublingly, enable action by malicious actors.

Children's identities and experiences are quantified and sorted to an unprecedented degree

Today's children are the first generation growing up at a time of rapid datafication where almost all aspects of their lives, both on and offline, are turned into data points. The UK Children's Commissioner's "Who Knows What About Me" report, for example, highlights data-generating technologies and activities that children engage with:

1. In the home, such as connected toys or smart speakers;

2. Online, such as children's use of social media or web-browsing behaviours; and

3. Out and about, such as digital health records or school databases.[36]

Indeed, every year the average child will have more data collected about them in their lifetime than would a similar child born any previous year.[37] This supply of data allows an ever-increasing number of groups to form when it is compiled, mingled, or processed by an algorithm.

There is no panacea for addressing the challenges of children's group data, but good practices are beginning to come into focus

The potential uses of such large volumes of data and opaquely generated group datasets are unpredictable. When used responsibly and effectively, data can provide significant value for children by improving, for example, service delivery and needs assessment. At the same time, these assets can be used to profile or discriminate against groups of children, target them for malicious activity, or expose them to other risks that are nearly impossible to forecast comprehensively and mitigate effectively.

Recommendations and conclusion

The responsible handling of group data for or about children poses significant and complex challenges for government and institutional decision makers. The lack of available research means policymakers and practitioners will be disappointed in their search for rigorous, field-tested methodologies for using children's group data effectively and mitigating the risks such datasets create.

There is no panacea for addressing the challenges of children's group data, but good practices are beginning to come into focus. Based on research detailed above, we outline three recommendations for advancing responsible group data for children. These recommendations are presented in relation to the RD4C Principles: Participatory, People-Centric, Purpose-Driven, Proportional, Protective of Children's Rights, Professionally Accountable, and Prevention of Harms Across the Data Life Cycle.[38]

1. Participatory and people-centric — seek insight into perceptions, challenges, and contextual considerations through participatory engagement and learning exercises.

As discussed above, data subjects often have little control over how data about them is managed and used, if they are aware of it at all. To provide individuals with some input into data use, data holders and decision makers could establish citizen juries or mini-publics of community members, domain experts, caregivers, and children to deliberate on children's group data collection, analysis, or use. This engagement could identify context- or community-specific opportunities, risks, or challenges that practitioners working in isolation might not consider.

In effect, this work would enable informed participation, "the effort to inform populations about how group data, including DII that may include them, will be acquired and used".[39] As argued in *The Signal Code*, informed participation can ensure legitimate and ethical data collection and use, especially in circumstances that do not allow for informed consent. Actors using children's group data would benefit from a participatory approach that prioritizes not just information provision but also the insights and perceptions from caregivers, community leaders, and children themselves.

2. Purpose-driven, proportional, and protective of children's rights — ensure there is a clear and well-defined purpose for algorithmically generating a new child group as the basis for data analysis and use.

As discussed, it can be hard for organizations using data to control for risks when use cases are undefined and opaque. This problem can be addressed, in part, by ensuring some control over data, by focusing use toward a specific purpose.

The RD4C Principles call for the collection, processing, sharing, analysis, and use of children's data to be driven by a clearly defined and articulated purpose. Purpose-driven group data use is also subject to this need, even if there are additional complexities due to the capacity of group data to constitute new

identities or profiles of children represented in that data through segmentation and analysis.[40]

Establishing a new and potentially vulnerable identity for children necessitates a clear and compelling purpose. Actors creating new group identities for children should be able to articulate not just the intended value of the analysis of that group data, but also why the benefits of constituting this new group are likely to outweigh the risks.

3. Professionally accountable and prevention of harms across the data life cycle — establish clear policies, procedures and responsibilities for mitigating group data risks.

There is a lack of state policies and institutional procedures directed at responsible group data handling. While individual risks dominate institutional attention, the failure to consider group data risks makes protecting vulnerable groups difficult. Thus, as discussed, improvements in responsible data handling are unlikely to be realized unless institutions define roles and responsibilities to support these objectives. Responsible data for children relies on people trained and empowered to support data responsibility through codified practices that support their efforts and ensure their accountability.

Data holders and the partners they engage can help ensure responsible handling of children's group data by defining and communicating which parties are responsible and accountable for relevant activities and decisions across the data lifecycle. Tracking and monitoring Decision Provenance — the chain of actors, inputs, processes, procedures influencing choices made within a system[41] — can clarify roles and responsibilities and enable more rapid diagnoses of issues that may arise.[42] In cases of inter-institutional data collaboration, actors could enshrine responsible children's group data practices in contracts, memoranda of understanding, or other agreements. Such standardized language could help to set expectations and ensure that all parties handling children's group data are cognizant of the risks and accountable for any harms.

Institutional procedures could also include, as recommended in *The Signal Code*, codifying processes for documenting and disclosing information about critical group data incidents. This documentation could help crystallize the group data-risk profile in various contexts and inform evidence-based iteration and course correction in data handling policies and practices.

Finally, institutions should have procedures in place for regular evaluation of the risks and sensitivities of any children's group data they generate or use. Beyond evaluation, there is a need for codified procedures for ceasing group data manipulation if risks are determined to be untenable, or if the risk profile remains so opaque that actors cannot meaningfully assess their activities' potential for doing harm.[43]

This paper was developed by members of the Working Group on Good Governance of Children's Data. Learn more about the project ▶

Good Governance of Children's Data project

The Office of Global Insight and Policy is bringing together 17 global experts in a project to explore trends in the governance of children's data, including the tensions between different rules and norms, emerging concepts and practice, and implications for policy and regulation. Debate on the future of children's data affects a diverse range of issues, including data ownership and control, data fiduciaries, profiling for digital marketing purposes, child-friendly privacy notices, data erasure upon request, age verification, parental responsibility, data protection by design and default, algorithmic bias, and individual and group data.

The project aims to highlight the gap between the world we want for children and today's reality, developing a manifesto on how children's data could be optimally managed and what steps need to be taken. To help develop this manifesto, members of the working group will publish short analyses of different approaches to data governance.

Endnotes

1 Greenwood, Faine, Howarth, Caitlin, Escudero Poole, Danielle, Raymond, Nathaniel A., Scarnecchia, Daniel P. *The Signal Code: A Human Rights Approach to Information During Crisis*, Harvard Humanitarian Initiative. January 2017.

2 https://rd4c.org

3 Halliman, Dara and de Hert, Paul, "Genetic Classes and Genetic Categories: Protecting Genetic Groups through Data Protection Law", in L. Taylor, L. Floridi, B. van der Sloot (eds.), *Group Privacy: New Challenges of Data Technologies*. Dordrecht: Springer 2017.

4 Taylor, Linnet, Floridi, Luciano and van der Sloot, Bart, "Introduction: A New Perspective on Privacy", in L. Taylor, L. Floridi, B. van der Sloot (eds.), *Group Privacy: New Challenges of Data Technologies*. Dordrecht: Springer 2017.

5 Ibid.

6 Floridi, Luciano, "Group Privacy: A Defence and an Interpretation", in L. Taylor, L. Floridi, B. van der Sloot (eds.), *Group Privacy: New Challenges of Data Technologies*. Dordrecht: Springer 2017.

7 Taylor, Linnet, van der Sloot, Bart, Floridi, Luciano, "Conclusion: What Do We Know About Group Privacy?", in L. Taylor, L. Floridi, B. van der Sloot (eds.), *Group Privacy: New Challenges of Data Technologies*. Dordrecht: Springer 2017.

8 Greenwood et al. *The Signal Code*.

9 Tisné, Martin, "The Data Delusion: Protecting Individual Data Isn't Enough When the Harm Is Collective", Luminate, July 2020.

10 Zwitter, Andrej, "International Humanitarian and Development Aid and Big Data Governance", in B. Schippers (ed.), *The Routledge Handbook to Rethinking Ethics in International Relations*. Routledge 2020.

11 van der Sloot, Bart, "Do Groups Have a Right to Protect Their Group Interest in Privacy and Should They? Peeling the onion of rights and interests protected under Article 8 ECHR", in L. Taylor, L. Floridi, B. van der Sloot (eds.), *Group Privacy: New Challenges of Data Technologies*. Dordrecht: Springer 2017.

12 Tisné, Martin, "The Data Delusion: Protecting Individual Data Isn't Enough When the Harm Is Collective", Luminate, July 2020.

13 Mittelstadt, Brent, "From Individual to Group Privacy in Big Data Analytics", *Philosophy & Technology* 30: 475–494 (2017).

14 Raymond, Nathaniel A., Al Achkar, Ziad, Verhulst, Stefaan, Berens, Jos, "Building Data Responsibility into Humanitarian Action", UN OCHA Policy and Studies Series, May 2016. https://www.unocha.org/publication/policy-briefs-studies/building-data-responsibility-humanitarian-action

15 "510 Data Responsibility Policy Version 2.0"

16 Greenwood et al. *The Signal Code*.

17 Council of Europe, "Towards CoE NEW Recommendations on Profiling?" Namur Digital Institute, July 1, 2020. https://rm.coe.int/profiling3/16809eed1e

18 "Report by the Committee of Experts on Non-Personal Data Governance Framework", Ministry of Electronics and Information Technology, Government of India, 2020. https://static.mygov.in/rest/s3fs-public/mygov_159453381955063671.pdf

19 Reuters, "Panel of US Firms to Push Back Against India's Regulation of Non-Personal Data", *The Wire*, August 10, 2020. https://thewire.in/tech/panel-of-us-firms-to-push-back-against-indias-regulation-of-non-personal-data

20 Basu, Samraat and Sonkar, Siddharth, "Examining India's Quest to Regulate, Govern and Exploit Non-Personal Data", *The Wire*, July 18, 2020. https://thewire.in/tech/non-personal-data-protection-anonymisation

21 Mantelero, Alessandro, "From Group Privacy to Collective Privacy: Towards a New Dimension of Privacy and Data Protection in the Big Data Era", in L. Taylor, L. Floridi, B. van der Sloot (eds.), *Group Privacy: New Challenges of Data Technologies*. Dordrecht: Springer 2017.

22 Nuila H., Andrea, "FIAN International Briefing Note: Collective Rights in the United Nations Declaration on the Rights of Peasants and Other People Working in Rural Areas", FIAN International, March 2018. https://www.fian.org/fileadmin/media/publications_2018/Reports_and_guidelines/droits_collectifs_UK_web.pdf

23 United Nations Office of the High Commissioner on Human Rights, "Minority Rights: International Standards and Guidance for Implementation", 2010. https://www.ohchr.org/Documents/Publications/MinorityRights_en.pdf

24 Raymond, Nathaniel A., "Beyond 'Do No Harm' and Individual Consent: Reckoning with the Emerging Ethical Challenges of Civil Society's Use of Data", in L. Taylor, L. Floridi, B. van der Sloot (eds.), *Group Privacy: New Challenges of Data Technologies*. Dordrecht: Springer 2017.

25 Mittelstadt, Brent, "From Individual to Group Privacy in Big Data Analytics", *Philosophy & Technology* 30: 475–494 (2017).

26 Sweeney, Latanya, *Simple Demographics Often Identify People Uniquely*. Carnegie Mellon University, Data Privacy Working Paper 3. Pittsburgh 2000. (PDF).

27 Taylor, Linnet, "Safety in Numbers? Group Privacy and Big Data Analytics in the Developing World", in L. Taylor, L. Floridi, B. van der Sloot (eds.), Group Privacy: New Challenges of Data Technologies. Dordrecht: Springer 2017.

28 Mittelstadt, Brent "From Individual to Group Privacy in Big Data Analytics", *Philosophy & Technology* 30: 475–494 (2017).

29 van der Sloot, Bart, "Do Groups Have a Right to Protect Their Group Interest in Privacy and Should They? Peeling the onion of rights and interests protected under Article 8 ECHR", in L. Taylor, L. Floridi, B. van der Sloot (eds.), *Group Privacy: New Challenges of Data Technologies*. Dordrecht: Springer 2017.

30 Taylor, Linnet, "Safety in Numbers? Group Privacy and Big Data Analytics in the Developing World", in L. Taylor, L. Floridi, B. van der Sloot (eds.), *Group Privacy: New Challenges of Data Technologies*. Dordrecht: Springer 2017.

31 Brent Mittelstadt, "From Individual to Group Privacy in Big Data Analytics", *Philosophy & Technology* 30: 475–494 (2017).

32 Ibid.

33 Ibid.

34 Tisné, Martin, "The Data Delusion: Protecting Individual Data Isn't Enough When the Harm Is Collective", Luminate, July 2020.

35 Kammourieh, Lanah, Baar, Thomas, Berens, Jos, Letouze, Emmanuel, Manske, Julia, Palmer, John, Sangokoya, David, Vinck, Patrick, "Group Privacy in the Age of Big Data", in L. Taylor, L. Floridi, B. van der Sloot (eds.), *Group Privacy: New Challenges of Data Technologies*. Dordrecht: Springer 2017.

36 United Kingdom Children's Commissioner, "Who Knows What About Me". https://www.childrenscommissioner.gov.uk/digital/who-knows-what-about-me/

37 Young, Andrew and Verhulst, Stefaan, "Why We Need Responsible Data for Children", *The Conversation*, March 23, 2020. https://theconversation.com/why-we-need-responsible-data-for-children-134052/

38 Responsible Data for Children. https://rd4c.org

39 Greenwood et al. *The Signal Code*.

40 Floridi, Luciano, "Group Privacy: A Defence and an Interpretation", in L. Taylor, L. Floridi, B. van der Sloot (eds.), *Group Privacy: New Challenges of Data Technologies*. Dordrecht: Springer 2017.

41 Singh, Jatinder, Cobbe, Jennifer, Norval, Chris, "Decision Provenance: Harnessing Data Flow for Accountable Systems", *arXiv Computers and Society*, 16 April 2018. https://arxiv.org/abs/1804.05741

42 See the RD4C Decision Provenance Mapping tool. https://files.rd4c.org/RD4C_Decision_Provenance_Mapping.pdf

43 Raymond, Nathaniel A., "Beyond 'Do No Harm' and Individual Consent: Reckoning with the Emerging Ethical Challenges of Civil Society's Use of Data", in L. Taylor, L. Floridi, B. van der Sloot (eds.), *Group Privacy: New Challenges of Data Technologies*. Dordrecht: Springer 2017.

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

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

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

The author would like to thank the many contributors to the Responsible Data for Children (RD4C) Initiative, which formed the basis for this work: Stuart Campo, Stefaan G. Verhulst, Robert MacTavish and Karen Carter. Additional thanks to Kerry Albright, Atthawoot Angkharat, Lori Bell, Theirry Beniflah, Gabrielle Berman, Eduard Bonet Porqueras, Tamima Boutel, Jasmina Byrne, Sumaira Chowdhury, Benoit Conti, Gabriele Erba, Vidhya Ganesh, Corina Gugler, Miles Hastie, Shreyasi Jha, Hye Jung Han, Remy Mwamba, Louise Mwirigi, Kristina Rashid, Cecilia Sanchez Bodas, David Stewart, Peter de Vries, Mark Waltham, Toby Wicks and Cornelius Williams. Thanks also to Andrew J. Zahuranec of The GovLab for their review of and input on an early draft of this paper.

This document is interactive and designed for digital viewing.

Please consider the environment and refrain from printing.