



Episode 2: Child Data Citizen

We all know that the global data economy relies on the ongoing collection, exchange and use of massive amounts of our data – from personal information, to what we do online, to algorithmic forecasts about what we might do in the future. But what about children’s data? In this episode, Dr. Sara Grimes chats with Dr. Veronica Barassi, about her research and theories of how childhood itself is being transformed by the production and manipulation of personally identifying digital data.

Dr. Sara Grimes 0:00

Let’s talk about data. These days everybody’s talking about data, it’s now common knowledge that the information economy relies heavily on the collection, exchange, and use of our data. Personal details, browsing histories, behavioral traces, photos, and posted comments, along with the enumerable associations, inferences, predictions, and classifications that are generated out of that data.

Data flows from and around us continuously, streaming in and out of our various apps and networks across our devices, blurring our public and private lives emerging out of our intentional and unintentional interactions with other people, but also out of an increasing number of our dealings with institutions, companies, and governments. What happens to that data, how it’s used, by whom, and to what ends is an issue of growing urgency and mounting public concern. In some parts of the world, like the EU it’s even led to the introduction of significant new privacy regulations aimed at curbing the access and power of the data economy and protecting the rights of users.

If there’s one group of users that’s seen as especially in need of privacy protection online, it’s children. Images of the child at risk are often evoked in policy debates and in media coverage of issues relating to data privacy. Indeed, many countries have enacted special protections and limits on how service providers can gather and display children’s personally identifiable information and other sensitive data. And yet children’s data is still

being collected, in massive quantities, across all areas of their lives and of those of their parents and caregivers. It is harvested, profiled and mobilized by the children's industries by big data brokers and by governments for various reasons, both known and unknown.

What's the impact of this on children's identities? On their wellbeing? On their rights? Dr. Veronica Barassi, professor in Media and Communication Studies at the University of St.Gallen in Switzerland is an anthropologist who spent the past several years researching these very questions.

She's the author of *Child Data Citizen: How Tech Companies Are Profiling Us from Before Birth*, a groundbreaking new book published by The MIT Press in 2020. In it, Dr. Barassi shows how children today are the first generation of citizens to be datafied. A process that increasingly starts before a child is even born. Think pregnancy apps and digitize medical records and all the ways a data is gathered from and about expected mothers. She argues that this has critical implications, not only for children as they're profiled and sorted, but for our shared democratic future.

Dr. Barassi has presented this important and eye opening work at talks and symposia around the world, including a 2019 TEDx Talk, that's attracted nearly two million views to date. It builds on her larger body scholarship, which examines digital participation, civic engagement, data rights, and data justice. Her previous publications include the book *Activism on the Web: Everyday Struggles Against Digital Capitalism* published by Routledge in 2015, as well as the journal article *Datafied Citizens in the Age of Coerced Digital Participation*, published in the journal *Sociological Research Online* in 2019.

Dr. Barassi does a lot of public outreach, working with companies to help them implement privacy by design, into their products and services and sharing her findings with policymakers in the UK and Ireland among others.

I'm Sarah Grimes, director of the Knowledge Media Design Institute at the University of Toronto and host of *The Critical Technology Podcast*. Today, I'll be talking to Dr. Veronica Barassi about her recent book, *Child Data Citizen*, and her thoughts on children's privacy, the datafication of childhood and the potential for data justice for children in the digital environment.

Let's jump right in. Who is the child data citizen?

Dr. Veronica Barassi 4:29

Well, The Child Data Citizen, I think it's my daughters. It's all the children that were born after 2011, those children that for the very first time history have become datafied from before they were born. I decided to write the book and to launch the project precisely for this reason, because I was studying and writing about surveillance and data collection in the context of activists. I got pregnant with my first daughter and while I was writing my first book, I actually realized that a lot of parents around me and a lot of people around me were producing so many data traces of children. Sometimes these data traces were also very sensitive and also political. I started questioning what did that mean for the children of that generation to be datafied from before birth? That's who the child data citizen is.

Dr. Sara Grimes 5:38

In the book you talk quite a bit about big data leading to a process called datafication. Would you mind telling us a little bit more about what this term means?

Dr. Veronica Barassi 5:49

The rise of big data and the term big data was describing a big technical transformation that was happening at the time. What we were seeing around the 2011, 2012, was the creation of big computers and powerful computers, which were able to integrate and analyze larger and larger data sets. So we saw a growth in the computer power and the ability to process data, but also we saw a growth in terms of databases.

The idea of big data was that it precisely linked to the idea of hugeness of big [inaudible] but then something else happened at the same time, the rise of data coincided with the rise of surveillance capitalism or came just after the rise of surveillance capitalism. A lot of society was concerned with idea that data was value and that data was capital. Around us, at social level, what we actually had was, different institutions, different businesses, different organizations who started to invest massively in the accumulation and the analysis of personal data. That's why we talk about datafication.

The first one to mention the term datafication were Mayer-Schönberger and Cukier in

their book *Big Data: A Revolution*, but really datafication came to be a term that was used mostly within social research to explore how our society was being datafied around us.

Dr. Sara Grimes 7:47

You mentioned surveillance capitalism, this is another concept that comes up a lot in the book. What role does surveillance capitalism play in the datafication of children and how does your work contribute to the larger discussion around this idea?

Dr. Veronica Barassi 8:02

The very idea of surveillance capitalism started to emerge 10 years before, what we currently understand, as the rise of the data, so around 2002. This was a time in which digital companies needed to create a new understanding of the online economy, because the online economy had just crashed. It was just after 2000, just after the crash of the .com bubble and we had different companies, digital companies in the Silicon Valley that had to reinvent themselves. To reinvent themselves, they started to realize that there was a lot to gain if they started to capitalize on data. Zubo believes that the first to notice this was Google.

In 2002 Google discovered, what she understands as, this behavioral surplus. Basically the idea that all of the data that companies were already collecting on people's behaviors, they were collecting that data to improve their technologies. That data could be used to create new value.

The idea that we needed to gather a lot of people's data, that we could cross reference this data, then once we did that, we had an insight into people's behaviors and we could use this insight to make decisions, to mitigate risks or to target them with specific content and so on. All of these ideas, that to us, are very common today, like the idea of predictive analytics, the ideas that they value really emerged during that historical period, which was actually led to a fundamental economic transformation. Now, obviously the economic transformation that surveillance capitalism enabled was precisely that of datafication because everywhere around us, people started to make money off of our data and wanted collect our data.

So, how does this affect children?

First of all, a lot of the data on children, at the moment, is being gathered from many different sources. This data can be integrated into unique ID profiles. What is happening at the moment is that companies have the technological means and the data means to follow individuals across a lifetime from the moment in which are born until they die.

Another fundamental issue that emerges when we talk about children's data, of course, is how this data is being used. Not only is it used to create comprehensive unique ID profiles, but it's to classify children and close them and lock them into profiles that then are sold and repackaged to different companies for different means. This data then is going to impact their life because we don't know how they're being profiled. The real problem is the fact that they often are going to be profiled on the basis of the groups that they belong to and the basis of the family. There's a real scope for discrimination for locking children in stereotypes and evade social mobility. This were my main concerns when I launched the Child Data Citizen Project.

Dr. Sara Grimes 11:51

Your research explores how data is collected about children, both actively and passively from the children themselves, but also through parents, teachers, grandparents, and so on. Can you explain the relationships between these different modes of data collection?

Dr. Veronica Barassi 12:06

This is something that I'm very much interested in because, the example I like is the one for kids, for example, a messenger for kids, YouTube for Kids, you can see that the big tech are actively and directly aiming at creating platforms that gather a lot of children's data. TikTok as well. There's a lot of different companies that are doing the same. If you look into the children's privacy policy of the companies, they're not subtle in the amount of data that they're collecting, Facebook messenger for Kids collects all of the videos, all of the content, all of the interactions, all of the communications that the children have. That's obviously a key example of the fact that these companies are directly trying to gather the data of children, with parental consent, and so legally. They obviously accumulate this data and they share this data. That's one way in which children's data is being produced. Whenever children download the new app and they have online ID profiles and things like that, that's direct ways in which children's data is being produced and collected.

There are other forms that are indirect ways in which children's data is being produced

and collected. This is through adult profiles. For instance, usually parents share a lot of information, a lot of photos on their adult profiles on Facebook, Instagram or Twitter. Another example can be, if you think about home technologies. Home technologies like Alexa, Amazon Echo, or Google Assistant, they are not directed at children, they're not designed for children, but children interact with this technology on a daily basis. Now, one fundamental problem with this is the fact that these technologies are not designed and targeted at children so they don't have to comply with children data protection regulations.

The other big problem that is emerging is the problem of biometrics. Photographs can be turned into face prints and face prints are like fingerprints, the same can be said about voice data. Both face data and voice data are actually biometrical data that can be turned into biometrical data. Now, a lot of companies say that they don't do it, especially Facebook for instance says that it doesn't take the face presence of people under the age of 18. Having said that there is a fundamental problem there and debates are emerging around Facebook and the collection of face grids. The fundamental problem that we have is that companies say that they're not doing it, but they have the technologies to do it. They have a lot of data and a lot of data that can be aggregated under unique ID profiles, because it can be connected to biometrical data.

Dr. Sara Grimes 15:29

So what role does intersectionality, race, ethnicity, gender class, and sexual orientation play in these relationships?

Dr. Veronica Barassi 15:37

Well, I think that it plays a fundamental role. When we're thinking about stereotypes, and when we're thinking about how algorithms reads human beings, we need to understand that we are not being profiled on the basis of our personal and individual data. I know that this sounds weird to say, because everybody said that the case, but that's not technically the case. We are often profiled on the basis of the groups that we belong to. If we belong to a specific family, a specific class background, a specific ethnicity, and so on, we are being profiled for that.

Okay.

Now, historically, people have always profiled others on the basis of this sensitive data. We've always been judged on the basis of our religion, our ethnicity and we've always encountered human bias. What's actually new is that, because AI systems at the moment, are cross referencing different types of data, they are making decisions on the basis of the process of algorithm profiling. We don't have any control over these decisions, we don't have any understanding of these decisions.

What really happens is that I'm being profiled in a particular way, or my child has become profiled in a particular way. Then it's automatically not going to target with a specific content, he's going to be automatically excluded from some form of information. This is the real problem that emerges when we're talking about data justice and equality and how social inequality intersects in algorithm profiling, this is the key idea.

When you think about children, for instance, one example that I make in my book, which I find that technically problematic in this regard, is the use of personalized learning. Now, the use of personalized learning emerged because people thought that it was a great idea, especially if you could gather data of students and identify risks, or identify children that were struggling, and then you would target them with specific conduct. There was a good idea behind it but when you think about personalized learning, there is something extremely problematic that emerged. This is the fact that, at that age are being profiled on the basis of very different types of people. In the book, I talk about the case of Summit Learning, which is founded by the Mark Zuckerberg and Chan, his wife.

When you go to Summit Learning it's quite interesting because they talk about personalized learning but, when you actually look at the data that they're collecting, they're collecting a wide variety of data. They're collecting data about ethnicity, they're collecting data about the class background, they're collecting data about the kind of notes, and teachers observations, anything that they could actually integrate together. Then on the basis of this data, they create targeted education. Children are being stereotyped and they're being locked into stereotypes at that early age. On the basis of their data they're being classified and then specific content is decided, the best content for them. This is the logic of education and it's particularly problematic because this stalls social mobility, this creates a problem in the making.

Dr. Sara Grimes 19:56

Your response just now, again points to the wide range of [00:20:00] ways and places that children's data is collected. In the book you divide this up into four categories of data flows that impact children's lives and futures. Health data, educational data, as in the example, you just described home life data, and social media data. Why did you focus on these four categories?

Dr. Veronica Barassi 20:20

I suppose it's, when you do graphic research and you do a bunch logical research, you just write on what you find out, those were the main categories that came out during my research. One aspect that really surprised me was that, usually the parents that I talked to were normally quite relaxed with the idea of data privacy. They were like, "Oh, I don't mind posting photos of my children on social media" but they would be very worried when it came to, for instance, health profiling, especially in the US where everything is based on health insurance, they would be very private about health. I could see that people had different attitudes towards different data, especially health and education data, that was more sensitive data to the families.

Parents don't want their children to have a bad record at school that will follow them across their lifetime. What I want to show with the book, is that if you start breaking down the key data flows in family lives, social media data, home life data, health data, and educational data, what you realize is that the different data flows have different impacts because being profiled for a health problem across your lifetime, that could have massive implication on their life. This is beyond the people that say, "Oh, you're going against the children's privacy because you posted a photo of your child eating an ice cream on social media." What I want to show is how people are producing these types of data, how this type of data is being collected and the implications of this type of data on children's lives.

Dr. Sara Grimes 22:15

A common question that comes up in discussions of children's privacy is "Why do the parents agree to all of this in the first place?" I'm thinking here about examples, like the sharing teens debates, which you touched on in your response just now. In much of the literature, describes it as the result of a "Privacy trade off" or a "Digital resignation that

parents end up engaging in,” but one of your key contributions is the idea that parents are actually often participating under coercion, not as the result of active decision making, or even decision fatigue in the book you call this, “The coercion of digital participation.” Can you talk a bit more about this?

Dr. Veronica Barassi 22:54

Yeah, sure. Can I tell you a story?

One day, at the beginning of 2020, as a special treat for my daughter, I decided to take her, just me and her, to Disneyland. I go to Disneyland and it was her special thing, it was her present. I arrive at Disneyland, I buy the tickets, which are not cheap, and we are queuing to get in after we bought the tickets. The lady at the door took my face scan and then my daughter’s, I immediately realized that they’re using facial recognition. I asked the lady “What’s going on, why are you using facial recognition?”

She explained to me that it’s a security measure, if somebody takes the child out of the park, or if I want to park hop, because there are two different parks at Disneyland in Los Angeles, that way didn’t have to show our tickets at both places. I was like, “Is there a way that I can I opt out? Can I see a data privacy policy?” They were very unsure about all this and she called the manager. By that time, you have to imagine that I’m standing there, I’m asking these questions, but I feel that it’s so unfair that they haven’t even asked me in the first place, if it was okay. [inaudible] showed me a data privacy policy.

My daughter is literally dying to get in and she’s getting really frustrated and rightfully upset because I am lingering at the door and everybody’s looking at us. That was a really awful experience. I decided to just go for it and they took her face and then they took mine as well. They promised me that they would delete it in within the day. Again, I didn’t see any data privacy policy just to know for sure. Even if I did, would I have had a real choice? This is the thing that I’m asking in my book, the majority of parents are often confronted with these situations where they don’t have a choice.

I moved to Zurich, where my daughter was starting her new school at the beginning of the pandemic and they were using Google classroom and I had to create a Google account. It had been five years that I’ve been studying this topic and I was not really happy and

I asked questions. The principal was like, "Well, this is what we use, that's what we can afford, it's free." I draw on experience of parents and experiences of myself and I noted a lot of acknowledgments of this. I rely on the theory around citizens and the idea that coercion is not necessarily the use of force, but caution is understood as force compliance. We are constantly forced to comply and to agree to terms and conditions, that's why I talk about the coercion of digital participation.

Dr. Sara Grimes 26:28

This brings up another point that you make, which is that the centrality of parental consent in existing regulation is in itself problematic.

Dr. Veronica Barassi 26:37

Once you understand that most of the parents are coerced into digitally participating then the other thing that comes quite natural to understand is the fact that when parents agree to terms and conditions in the majority of the cases, their consent is not informed at all. That's a real problem. That's something that I am particularly worried about, because up until now, when we were thinking about privacy regulation, and a good example of this is, the General Data Protection Regulation in Europe, we often talked about the importance of companies to be transparent and the importance of always requesting consent. The problem of consent, at the moment is enormous. We live in an environment that is completely datafied and if we were to try and achieve informed consent for every single data policy that we agreed to, it would be practically impossible.

I'll make you an example, an example that I also talk about in the book. One night, I was reading an article on the Me Too movement on my [inaudible]. It had nothing to do with children or the Child Data Citizen project. I was on my own, just reading online, like many of us do and I clicked on an article of an Italian magazine of the website. To comply with the General Data Protection Regulation, the website asked me if I wanted to opt out personalization and find out about the third party with whom they would share my reading that website. Right. I clicked "Yes" and the next thing I know is that I am presented with a list of companies that will be gathering my data. I counted the list of companies, and there were 439 companies.

The website was explicit by saying, "This is how we process your data, but actually, if you want to understand how they process your data, these are their privacy policies." There were 439 for one single click.

Parents don't read the terms and conditions, not only because they can't be hassled or because they decide not to read them, but because they literally don't have the time. To give you an idea of how much time it would be, the only research that I could find on this was research by McDonald and Crandall in 2008. They calculated that reading all of the privacy policies of websites that the users encountered in a year would amount to around 201 hours per year of the user's life, 201, just to read the privacy policy.

Their calculation was based on the fact that they estimated, that to read one privacy policy, would usually take around eight to 10 minutes. Also, their calculation was based on the fact that an average user in the US, at the time would encounter 119 websites. For me, with one click, I encountered 439 privacy policy, in one click, that's a big difference.

There's no way that we're ever going to be able to provide that informed consent as parents. The other problem is that, even if we do provide informed consent, is that enough to protect children's rights? Is that enough? For instance, not all children are going to have the same values of the parents. Maybe parents are going to give their consent to share specific data of their children and then that data can follow their children across their lifetime and the children can't do anything about that. I think that the very notion of parental consent is really problematic because it doesn't capture what's happening globally.

Dr. Sara Grimes 30:50

Towards the end of the book, the discussion shifts to data justice. What does data justice look like in the specific context of children and children's data flows?

Dr. Veronica Barassi 30:59

Generally speaking, if we want data justice for children, then we need, not only a policy change, so a change in data regulations, but we also need a political change, a change

in institutional models and organizations. Also, consult change, a change in current perspectives about data privacy. Specifically, I think that if institutions and governments and academics and anyone interested in the field wants to pursue data justice for children and wants to fight for the data justice for children. I think we need three main transformations, which is one that we acknowledge that parental consent is often acquired through the portion of digital participation, so it's not informed consent and it's not transparent. The other aspect that I think we need to address is the fact that a lot of the data that is being collected of children today is collected through adult profiles. It's all aggregated profiles with adults.

I think new regulations should make sure that children are not judged or profiled on the basis of their families and collectives, so that household profiling, something that is happening at the moment should be illegal. Also, that all of the data that is being collected through technologies that are not designed or targeted at them should be deleted. I think that's something that we should be demanding at the moment.

Another aspect that we should be demanding is that more action is taken over how data brokers are using and sharing children's data. I think in that sense, regulation could be very important. When we think about children's data justice, I think that there's something really practical that we can do, but also very theoretical and that we have to keep in mind, which is the fact that all the data of children that is being collected at the moment could be in the future, used to determine their rights.

Dr. Sara Grimes 33:14

That was a great lead in to my next and final question, which is one that I'll be asked all of my guests this season. The United Nations recently adopted a new general comment confirming and outlining how children's rights apply in the digital environment. Do you think this will have any impact on the issues and relationships that you describe in your book?

Dr. Meryl Alpher 33:35

Well I think to be honest, I'm glad that you're asking me this question, because this is something that's very important. We're seeing so many different, important steps that are being taken. We are seeing is a greater understanding and awareness of aspects of the

identification of children that were not known five years ago. I think that all of this is very important because it's going to set the future. What I believe is happening is the fact that we have created technologies that are very powerful, that could bring lots of benefits as well. But we have to really recognize their problem, especially when it comes to algorithm profiling. I think that, at the moment we are seeing policy makers and we are seeing researchers working together to set out a series of steps to make this technologies a little bit safer.

An example that I always use is the fact that I would never fly in an airplane of the 1950s, 1950 or 1920, I would never fly on an airplane [inaudible], because it was a new technology and it was really not that safe, to our eyes today, and I certainly wouldn't put my children on them. Right.

This is what we're facing at the moment with AI, we're facing a very complex process of negotiation. We're fighting a war between different interests, between different political and visions about what these technologies do and how can these technologies can be used. All of the work that is being done between policy makers and academia and the proposal for regulations. Also, for instance, the proposal to regulate AI in Europe, which I find they very interesting. All of these are very important steps forward if we want to try to ensure that we live more demographic futures.

Dr. Sara Grimes 36:05

A big thanks to professor Barassi for joining us today, please follow the links in the podcast description to find out more about Dr. Barassi's research, her book, Child Data Citizen, and the other publications mentioned in today's episode, as well as information on where to send your questions or comments. The Critical Technology Podcast is produced by me, Sarah Grimes, with support from the KMDI. Audio mix and sound design by Mika Sustar. Music by Nicholas Manalo. Theme song by Taekun Park. Our logo was designed by JP King and the artwork for today's episode was created by Kenji Toyoka. Please subscribe to stay up to date on new episodes and posts as they become available. And thank you for listening.