Reclaiming Expertise: Facing Challenges to Knowledge, Practice, Authenticity, & Trust

5th annual Research & Inquiry Conference

Keynote Lecture: Child | Data | Citizen

Datafication, Algorithmic Inaccuracies, and the Profiling of Future Citizens.

Dr. Veronica Barassi

Department of Media and Communications Goldsmiths, University of London

Slide One Introduction and Thanks.

So today we are here to talk about knowledge, expertise and the what happens when this knowledge and expertise are being challenged, questioned and belittled. What is our role as researchers, educators and information experts?

 I personally believe that the contemporary challenge to knowledge and expertise lies within two different albeit interconnected cultural/societal changes that have taken place in the last 10 years.

- Firstly, after 2008, we have seen a steep decline in public trust in institutions. The financial crisis, its deep economic, social and of course personal and emotional repercussions have created the basis for a rise in popular frustration and anger. Working and middle classes were the ones who had been affected most across the world and of course, as the rise in populism within the E.U. and U.S. has shown, they transformed their anger into votes, votes that favoured those parties who challenged the status quo, who worked on emotions over facts, and whose political agenda was linked to the spreading of mistrust and the demonising of experts knowledge, journalists and facts.

- The second societal transformation that we have witness in the last decade and that has seriously impacted on the ways in which we think and talk about knowledge is represented by our societies changing attitude to data. The 'data revolution' if we want to use a catchy phrase, was triggered by three different yet interconnected transformations:

- a) a change in technologies – the development of supercomputers able to integrated large data sets, a rise in machine learning, artificial intelligence and facial recognition technologies
- b) a change in cultural narratives - the growing societal belief that data matters because it leads to profit and that and that algorithmic logic and larger datasets offer us a precise and accurate type of knowledge, which enables us to frame individual and social patterns and use this knowledge for different purposes. (boyd and Crawford, 2012:663)
- c) a change in social practices as governments, businesses, educational institutions, healthcare providers and multiple other agents have started to turn all aspects of everyday life into data (Cukier & Mayor-Schoenberger, 2013).

In this talk I want to concentrate on my latest project (Child \ Data | Citizen) and I want to explore the relationship between data and knowledge.

Especially if we look at the example of children, it is now becoming obvious that 'datafication of everything' is leading to the creation of large datasets of very intimate, private and sensitive information about citizens' lives from the moment in which they are conceived. Yet as the talk will show, we need to critically question the accuracy of our datafied knowledge, as we need to carefully consider that there is a profound human disconnect between technological structures and human practices.

Slide Two

Before I start talking about the datafication of childhood, I want to dedicate a couple of minutes to reflect on my own knowledge practices as a researcher.

I am a trained anthropologist but I have been writing and teaching at the intersection of a plurality of disciplines, which include communication studies, political science, sociology, critical internet studies and research methods. Despite working and writing across different disciplines, throughout my career have always been committed and passionate about anthropological knowledge and the way in which anthropological knowledge is produced.

I believe anthropological knowledge differs itself from other types of knowledge production in social research. Anthropologist are committed to the ethnographic method. And what is fascinating about the ethnographic method in anthropology, is that in contrast to what often happens within other disciplines (sociology or communication research) in anthropology the enthographic method is not used to validate hypothesis or to explore specific theories. The basic conception is that as a researcher, with your bias, your own personal history and emotions, you put yourself out there and you analyse through self-reflective knowledge practices the different dimensions of human experience.

In this context, as Pink has argued ethnography is not a 'method' but rather as a research perspective (Pink et al, 2016), one that is aimed at opening the field ahead of us. In fact as Tim Ingold suggests when within anthropology, ethnography is about research design, 'designing is about imagining the future. But far from seeking finality and closure, it is an imagining that is open-ended' (Ingold in Pink et al, 2016:11).

As anthropologists, we seek this open-endedness through our commitment to to thick description and cultural relativism. We like to be surprised, challenged, we like to look for the contradictions, disconnections and unpredictabilities. In this context as Greaber (2006) has argued, anthropologists believe that human possibilities are always greater than what we imagine, and they are there to look for these possibilities.

Slide 3
One of the strengths of anthropological knowledge is thus its ability to start from obervations, and thick descriptions and devise the theory. So I want to start from three ethnographic anedoctes which all relate to education.

1. One evening, last year I was having dinner with Ana (fictional name, all names are fictional), with whom I had become friends after I interviewed her in 2017, my daughters, and her two boys. Los Angeles. She was discussing how good the new school of her eldest was. It was her first choice, very highly ranked and few blocks away from her home. She

was excited and honoured that her son was studying there. Then she took out the phone and showed me the app that the school used for internal communications. She was excited that through the app she could track what happened at school, she could monitor absences and performance, and she could upload and download pictures. I asked her if she knew the company that outsourced the app, whether she read the terms and conditions or knew how all that data was going to be used. She looked at me laughed and said "of course not, and also what choice would I have?".

2. Few months later I was chatting to Jay, a grandfather of two, entrepreneur who also sat on the trustee board of the school. We were talking about the softwares and apps that the school outsources and uses. We were discussing what could be done to protect students' educational data. He demonstrated a real concern for the issue. But also told me that as a school they had little choice. At the heart of the issue was the matter of budget and funding. The school – he told me - was forced to choose the outsourcing of technologies not on the basis of the fairness of their data policies but on the basis of quality/price ratio.

3. One day in Spring 2017 I was interviewing a mother of two girls who lived in South London and who was very concerned about the online privacy of her children. Her and her husband had chosen not to post or share any photo or information on social media. Yet their position was being challenged by the other parents at the school. The issue of contention was the school play. Liz and her Husband together with another couple of parents were the only ones who refused to sign the permission to allow members of the audience to take photos, because they rightly thought that the photos would land on social media.

As the experience of Ana, Jay and Liz suggests, something is changing and its profoundly transforming the lived experience of education, and this something has a lot to do with the production and sharing of data traces. But how are we understand this transformation? And how are we to appreciate the tension, fears, insecurities and problems that arise as the education sector becomes increasingly more datafied?

Slide

With the Child | Data | Citizen Project my aim is to engage with these questions and I argue that we cannot understand the datafication of education without looking at the broader datafication of childhood. The project is funded by the British Academy and explores the everyday datafication of childhood by focusing on families in London and Los Angeles, with children between 0 and 13 years of age, whose personal information online is ruled by the Children's Online Privacy Protection Act (1998). The research relies on a multi-method approach, which includes, participant observation, interviews, digital ethnography and the qualitative platform analysis of 10 early infancy apps, 4 home hubs and 4 AI ToysThe Child.

So if we analyse an average child, we see that there are a plurality of technologies that are gathering large amounts of data. We can start by considering facial recognition (used by social media platforms or other services such as Google Photos) and then proceed to consider the role of voice recognition (such as AI Toys or home hubs, and here the situation is more complex, as these technologies gather large ammounts of very personal and contextual forms of data), we can then go on and consider educational data, health records, social media etc.

Slide 5

How can we understand this datafication of childhood in theoretical terms? In an article Lupton and Williamson (2017) attempt to map some of the multiple technologies that collect, share and process the personal information of children and they argue that we are seeing the emergence of the datafied child, where children become data assemblages they become 'calculable persons', they are not only the subject of calculations performed by others (and by other digital things) but are also enabled to think about, calculate about, predict and judge their own activities and those of others (Williamson and Lupton, 2017:787).

My own ethnographic research on the datafication of childhood reveals that whilst it is tempting to refer to concepts such as datafied child or quantified selves when we map - like Lupton and Williamson (2017) successfully do - the multiple digital technologies that collect the data of children, we must acknowledge that the datafication of children is not a linear,

cohesive or even a rational process that is transforming them into quantified selves. It is a rather complex and messy process defined by an incredible and almost untraceable plurality of digital technologies and practices that lead to the construction of multiple, messy, inaccurate and contradictory predictions.

During the research it became evident that parents most of the times did not use the technologies as they were supposed to and that the narratives that children's data traces were often the result of imprecise behavior or carefully employed *tactics* (Barassi and Trere, 2012) to protect their privacy and the ones of their children. When we think about datafication we need to realize that there is a clear *human disconnect* between technological discourses and structures (e.g. the promotional culture of self-tracking apps or their design) and everyday practices.

The ethnography of data therefore is there to highlight the messiness and contraddictions but also to explore the feel of data, how families are understanding and experiencing datafication. A second finding of the research reveals that datafication of childhood is not only happening because parents are voluntarily using technologies (such as sharenting, in other words sharing information on social media) but it is happening because families found themselves 'coerced' into complez scoio-technical environments where as Ana suggested they don't have a choice.

In the last few years I have been writing a lot about the 'coercion of digital participation' . Parents use the online platforms that are used by health providers or educational institutions of their children , they download a wide variety of apps from health trackers to educational data, they join facebook groups with other parents, they set up the latest AI or internet connected toys that were given to their children for their birthdays, Christmas and other special occasions. In doing so they constantly sign off the terms of conditions that decide how their children's data is used and passed on, believing like Ana that they don't have a choice. In doing so they often feel unease and powerless and this situation is made worse by the fact that they find themselves dealing with communication infrastructures and business models, that make it extraordinarily difficult to opt out, use ambiguous data policies and keep on changing terms and conditions.

Slide 6

But why does it matter to look at what is happening within our homes, and to the datafication of childhood as a broader phenomena. When I talk to parents they often tell me what do I care if the apps I use? The social media I join know what I do and use this info for targeted advertising.

Well it matters, it matters because as The Cambridge Analytica scandal of 2018 is perhaps the most vivid example of the complex ways in which digital profiling is impacting on democratic choices and civic rights., Yet this is certainly not a new phenomenon, Facebook for many years has been sharing information with data brokers (and of course this is a bit different because there wasn't a breach, but still data brokers have been profiling individuals for long now). Yet Facebook, in the last few years has also been buying large amounts of offline data and is now becoming a fierce competitor of data broking.

But what are data brokers? According to a report by the Federal Trade Commission (2014), already four years ago it was clear that data collected by data brokers related to numerous different dimensions of a citizen's life

"Data brokers collect data from commercial, government, and other publicly available sources. Data collected could include bankruptcy information, voting registration, consumer purchase data, web browsing activities, warranty registrations, and other details of consumers' everyday interactions. Data brokers do not obtain this data directly from consumers, and consumers are thus largely unaware that data brokers are collecting and using this information. While each data broker source may provide only a few data elements about a consumer's activities, data brokers can put all of these data elements together to form a more detailed composite of the consumer's life." (FTC Report, 2014: IV)

"marketers could even use the seemingly innocuous inferences about consumers in ways that raise concerns. For example, while a data broker could infer that a consumer belongs in a data segment for "Biker Enthusiasts," which would allow a motorcycle dealership to offer the consumer coupons, an insurance company using that same segment might infer that the consumer engages in risky behavior. Similarly, while data brokers have a data category for "Diabetes Interest" that manufacturer of sugar-free products could use to offer product

discounts, an insurance company could use that same category to classify a consumer as higher risk". (FTC Report, 2014:v)

What is interesting about the FTC report is the fact that individuals are identified as 'consumers' yet when the data collected and profiled - as the report suggests - is about voting registration or details about one's own religion, ethnicity etc, then we are not simply talking about consumers and consumers' rights, but citizens' and citizens' rights.

Slide 5

Through my research, by talking to parents, by seeing how they are coerced into sharing the data of their children and by considering the fact that messy data traces are often used to profile families, I came to the conclusion that we cannot understand the datafication of childhood without asking critical questions about the emergence of new forms of digital citizenship.

To date the concept of 'digital citizenship' has been used by very different scholarly traditions from the education sector to communication sciences, to explore the way in which through digital technologies we participate in society (Mossberger et al. 2007) .Within the education literature digital citizenship is understood as something that can be taught and learned; it's a toolkit, a set of civic values to be passed on to students.

The understanding of digital citizenship merely as a form of 'participation' has been crticicised by Isin and Ruppert (2015) who have argued that this definition leaves out too much in the analysis of how the internet has impacted on the figure of the citizen. The scholars believe that digital citizens are brought into being through speech acts that are 'digital acts, which enable them to make rights claim on the internet (2015:69). Their definition of digital citizenship draws heavily on the work of those scholars that see 'citizenship' as performance (Isin, 2008) and the one of Couldry et al. (2014) who explored how citizens' digital acts on the internet are tightly linked to digital storytelling.

One important aspect of digital technologies is that they enable the civic engagement of those who do not enjoy full citizens' rights, such as children. In this regard the work of Third and Collin's (2016) is particularly insightful as they re-think the notion of digital citizenship by looking at children's digital practices. The scholars place a special emphasis on the concept of *performance*, and show that children/youth's digital acts are often directed at confronting, contesting and challenging the adult world in a public and performative way. According to Third and Collin's (2016), however, children's ambiguous position in society, as not-yet-citizen, makes the performance of their digital citizenship more creative and radical than the adult one.

All these works on the performance and enactment of digital citizenship are of central importance as they add a new dimension to the very notion of digital participation. Yet in the last few years, the notion of digital citizenship has been at the centre of an interdisciplinary debate. The question at heart is: WHAT HAPPENS WHEN DATA TRACES SPEAK FOR AND ABOUT CITIZENS. In fact, According to Hintz et al., (2016) "at an historical time where both state agencies and companies surveille every aspect of citizen's life, we are not just digital citizens because of our actions but also because we increasingly live and operate in a datafied environment in which everything we do leaves data traces" (2016:732).

This latter point emerges particularly well if we look at children's data traces, children are not only digital citizens because their digital practices enable them to enact and perform their public persona (Third and Collin, 2016), they are datafied citizens because they are coerced into digitally participating to society through their data traces (Barassi, 2017a, 2017b).

Deconstructing the Datafied Child

When we question the datafication of citizens from birth, the understanding of the messiness of datafication is essential because it enables us to move to a further question, which relates to what happens when these data traces are collected with the purpose of using them for predictive analytics and the profiling of consumers/citizens. What type of knowledge can produced from these types of inaccurate data traces of citizen lives?

The truth is that so far we don't know the answer. What we know is that when we sign off terms and conditions or agree to use specific services and platforms, we are making a step towards their profiling not only as consumer subjects but also as citizen subjects.

It is precisely by looking at what is happing in the industry, at how data is bought, sold and profiled that we need to start question the role of children data traces. What we know so far is that 1) children's data traces are innacurate 2) that these traces are being collected, sold and used for predictive analytics in ways that escape our control 3) we also know that not only algorithms have biases, but that as Crawford and Shultz have shown there can be profound real life harms (they call these privacy harms) that emerge form predictive analytics.

Final remarks
So what can we do? Some concluding remarks.
What is scaring me at the moment as I work with families that whilst I find that reveals families are actively making sense and negotiating with digital surveillance and privacy. I also find that very few are aware of issues such as data broking and profiling. They are also often unaware of technological developments such as the developments in facial recognition (Gates, 2011), genetic discrimination (Mittlestadt and Floridi, 2015) or predictive policing (Dencik et al, 2017).

So I find myself coming to the conclusion of this talk in full circle. What we need now, I think is knowledge, we need to study, become experts, understand what is happening to our data traces. And we need to disseminate our findings as much as we can, not simply to raise awareness about issue and leaving individual frustrated about the fact that often they do not have a choice as the experiences of Ana, Jay and Liz have shown. We need to disseminate the knowledge to start asking questions as institutions, organisations and as a collective entity and and shift current debates about surveillance and privacy by turning the light towards freedom of expression, self-representation and data justice.