
RESPONSE to the Consultation on the White Paper on Artificial Intelligence - A European Approach

The Human Error in AI and question about Children's Rights

Introduction

One of the key dangers of AI innovation lies in the question about human profiling. We are living at an historical time when every little detail of our lived experience is turned into a data point that is used by AI systems and algorithms to profile us, judge us and make decisions about us. These technologies are used everywhere. Health and education practitioners use them to 'track risk factors' or find 'personalized solutions'. Employers, banks and insurers use them to judge clients or potential candidates. Even governments, the police and immigration officials use these technologies to decide key issues about individual lives, from one's right to asylum to one's likelihood to commit a crime. In sum, as the White Paper *On Artificial Intelligence – A European Approach to Excellence and Trust*, acknowledges:

“The use of AI can affect the values on which the EU is founded and lead to breaches of fundamental rights, including the rights to freedom of expression, freedom of assembly, human dignity, non-discrimination based on sex, racial or ethnic origin, religion or belief, disability, age or sexual orientation, as applicable in certain domains, protection of personal data and private life, or the right to an effective judicial remedy and a fair trial, as well as consumer protection (2020:10-11)”.

The White Paper recognises the breaches in human rights that can emerge when AI systems are used for human profiling and automated decision making. Yet, in the current form the paper omits to take into account the issues of the fallacy of algorithms when it comes to human profiling, and its implications for human rights and society. In this report I would like to introduce the commission to my definition of '*the human error*' in algorithms (Barassi, 2019; Barassi, 2020 in press) and provide evidence of how the human error in algorithms is impacting a section of society: children. If the aim of the European Union is to support the development of AI systems that are grounded in fundamental rights such as human dignity and privacy, then we need regulations that acknowledge the fact that algorithms will always be inaccurate and biased when it comes to human profiling.

'The Human Error' in Algorithmic Profiling

If we want to tackle questions about human rights and AI innovation, if we really want to build trust and excellence in AI then we need to acknowledge the fact that AI systems can bring much positive outcomes if they are used to tackle diseases or climate change, but they will *always and inevitably be biased* when it comes to human profiling.

The issue of algorithmic bias is at the centre of current debates. More and more tech businesses and AI developers are trying to find solutions to 'fight algorithmic bias in their products and technologies. It is for this reason that they are funding research and establishing advisory boards that are meant to scrutinize the ethical and political impacts of their

technologies (e.g. AI ethics). At the heart of these strategies and practices adopted in the industry, lies the very understanding that algorithms are biased because they have been fed ‘bad data’ and hence in order to rectify algorithmic bias, companies need to train algorithms with ‘fair’ or ‘unbiased data’ (Gangadharan and Niklas, 2019).

Current strategies to ‘combat algorithmic bias’, in the industry are profoundly problematic because they push forward the belief that algorithms can be corrected, and be unbiased. Research has shown that algorithmic bias is not something that can be resolved. Dencik et al. (2016), for instance, have argued that a practice such as predictive policing faces three different (and unresolvable) challenges: the inclusion of pre-existing biases and agendas in algorithms, the prominence of marketing-driven software, and the inability of interpreting and dealing with unpredictability. What the research, is showing is that not only algorithms are profoundly racially biased (Noble 2018) but also that citizens are constantly profiled and subjected to automated systems that often reproduce social inequalities (Eubanks 2018).

Rather than trying to fix algorithmic bias or believe in the dream of developing ‘fair’ AI systems, we need companies, politicians and legislators to recognise that there is a fundamental **‘human error’** (Barassi, 2019; Barassi, 2020 in press) in algorithms, because *algorithms cannot, at any time, profile humans in just and fair ways*. Recognizing the human error in algorithms presupposes that we take into account three different yet interconnected dimensions (Barassi, 2020, in press):

1. **Algorithmic bias** - Algorithms are always, at any-time biased in one way or another. An algorithm is by definition a set of rules or steps that need to be followed to achieve a specific result. These set of rules or steps are never ‘objective’ because they’re designed by human beings and are the product of specific cultural values. This finding is of course not new. In 1996, Friedman and Nissenbaum (1996) identified three types of bias in computer systems: *pre-existing bias* (the bias of the humans that design computer systems and the bias produced by the cultural context that influences the design) *technical bias* (often there is a lack of resources in the development of computer systems, and engineers work with technical limitations) *emergent bias* (society is always changing and thus the technologies designed at one given time or cultural context might become biased in a different time and context). Algorithms and AI systems are human made, and will always be shaped by the cultural values and beliefs of the humans and societies that created them. Hence rather than trying to defeat the bias we need to co-exist with it. Anthropologists have long been trying to grapple with the fact that individuals necessarily interpret real life phenomena according to their cultural beliefs and embodied experience (Clifford and Marcus, 1997), and that cultural bias necessarily translates into the systems that we build, including scientific systems (Latour and Woolgar, 1986). From an anthropological perspective, there is nothing that we can really do to ‘correct’ or combat our bias, because it will always be there. The only thing we can do, is to acknowledge the existence of bias through self-reflexive practice and admit that the systems, representations and artefacts that we build will never really be ‘objective’. This same understanding should be applied to our understanding of AI systems and automated decision making.

2. **Algorithmic Inaccuracy** – AI systems that are used to profile individuals are often trained on a multiple variety of data points taken from different sources (consumer habits, social media, apps etc.). The basic idea behind these systems is that data holds the key to human nature and behaviour and can be used to ‘profile’ individuals, find ‘personalized solutions’ or ‘mitigate future risks’. Yet the data processed by algorithms is often the product of everyday human practices, which are messy, incomplete and contradictory, hence algorithmic predictions are filled with inaccuracies, partial truths and misrepresentations. We need to always bear in mind the fact that even if we can trace connections and patterns this does not necessarily mean that the knowledge we acquire from these connections and patterns is accurate. Boyd and Crawford (2013) explained this well when they suggested that: “big data enables the practice of apophenia: seeing patterns where none actually exist, simply because enormous quantities of data can offer connections that radiate in all directions.” (boyd and Crawford, 2013: 668).
3. **Algorithmic unaccountability** - One of the biggest challenges that we face today is represented by the fact that – in the majority of cases - the algorithmic predictions that are used to profile us and to make data driven decisions about our lives cannot be explained (and hence are cannot be held accountable). According to the U.S. mathematician O’Neil (2016) algorithmic models that are used in a variety of fields such as insurance, policing, education and advertising are opaque, unregulated, and incontestable, even when they are wrong. Current debates on AI accountability and explainability often focus on the fact that these processes are to ‘big and complex to be explained’ and that striving for explainability (and hence accountability) would stall AI innovation. When we think about the industry-led claim that algorithmic and AI explainability stall innovation, we have to constantly remind ourselves that these technologies are used in a variety of areas of social life and that their ‘predictions’ – as we have seen above - can cause real-life harms to people’s livelihood. If we have to choose between AI innovation and defending our rights I personally would choose the latter.

Looking at the human error in algorithm, enables us to appreciate that human experience and psychology cannot be translated into a mathematical deduction. Those who sell the promise of algorithmic accuracy or objectivity in human profiling are doing just that: selling a promise. What they are doing in actual fact is building systems that stereotype people, and present reductionist and simplified interpretations of their needs. This promise can have a fundamental impact on our democratic futures. What we need is regulations that recognize that there is a fundamental ‘human error’ in algorithms, and that realize **that any AI that makes data-driven decisions about individual rights (private or public) needs to be considered High-Risk.**

The Human Error in Algorithms and Children’s Rights

Children today are the very first generation of European citizens that are datafied from birth, and are critically more exposed to the human error in algorithms. Looking at children enables

us to really appreciate the human-rights implications of an AI-driven world. Today, in the life of a child, data is everywhere. From the moment in which a child is conceived, important personal data is uploaded on social media, pregnancy apps or virtual assistants. As children grow up, most of their health and educational data is stored and processed by data brokers and artificial intelligence technologies. What is becoming obvious is that children's personal information is now being collected, archived, sold and profiled in ways that was not possible before.

In the last three years I have been working on the Child | Data | Citizen research project and have been examining the construction of children into data subjects, describing how their personal information is collected, archived, and sold. The Child | Data | Citizen project explored this transformation through a multi-method approach which involved auto-ethnographic research; 50 semi-structured in-depth interviews with parents with children from 0 to 13 years of age (whose personal information is regulated by Child Online Privacy Protection Act); 8 months of digital ethnography of parents 'sharenting' practices on the social media accounts of 8 families; and the platform analysis of 4 social media platforms; 10 apps (baby apps and pregnancy apps); 4 home hubs; 4 education platforms. By focusing on different types of children's data (social media data, home life data, education data and health data) the project was able to shed light on the messy and complex experience of data technologies in family life and highlight the human emotions, beliefs, business models and technologies that are making the datafication of children possible.

The main findings of the Child | Data | Citizen project led me to conclude that current data protection laws are failing children especially if we consider the issue of the human error in algorithms. This is how the human error in algorithms may be affecting children's human rights, and the rights of future citizens in the EU:

1. Algorithmic Inaccuracy – When we talk about children's data traces we cannot only talk about 'personal data' as umbrella term, because the data of children is constantly collected and processed through the profiles of their parents. In 2018, in order to reflect on the complexity of home data, I came up with the term '**home life data**' in a report on AI systems in the home, which has been used as evidence by the Information Commissioner Office in the UK and was signed by Gus Hosein, the Executive director of *Privacy International* and supported by Jeff Chester the director of the *Centre for Digital Democracy* in the US (Barassi, 2018). In the report I argue one fundamental problem with AI systems in the home is that they collect children's data but they do not have to comply to COPPA or the GDPR special provisions for children's rights, because they are not designed or targeted at children. I show that most of the data collected is collected through aggregated child-adult profiles, and argue that this data cannot be understood merely as personal data but rather as 'home life data'. Home life data includes the following data categories (which I revised and updated following the report):

- *Household data* – AI systems and data technologies in the home collect a wide variety of household data from shopping lists to energy consumption and gather key information on families' behaviours, choices and routines (including the ones of children).

- *Family data* – AI systems in the home gather a lot of family data which refers to family socio-economic background, family history, ethnicity, religion, social and political values, medical conditions etc.
- *Biometric data* – Most Virtual Assistants and smart technologies rely on the gathering of biometric data (voice recognition or facial recognition) that can be mapped to unique users, including children.
- *Situational data* – AI technologies to function need to gather situational data of the individual and the family. They need to be able to answer questions such as what room belongs to whom? They need to be able to register changes in family members or changes in circumstances etc.? Conflicts and tensions? Etc.

The fact that companies can gather (and sell) all these different forms of data implies not only that they have the potential to harness *highly contextual* data from children (through adult profiles) but also that they can integrate this data with *biometric information*. The privacy implications of technologies that can integrate context and biometrics are immense. Yet what we also need to realise is that these data technologies are extraordinarily open to inaccuracy and fallacy. Families often do not use technologies as they are designed to be used. This is not only because on an average family day technologies and profiles always overlap and this confuses algorithms, but also because families often input inaccurate data in their technologies, or use them tactically to try and protect their privacy. This human messiness of home life data inevitably confuses algorithms and leads to inaccurate profiling.

2. Algorithmic Bias - Children are being affected by the practice of household profiling, which is established amongst data brokers and tech companies. For example, education data brokers in the US, sell not only the data of individual students, but also the data about their parents (job, ethnicity, financial situation, lifestyle factors, marital status, etc.) (Russell et al., 2018). It is for the same reason that tech-companies are trying to gather as much data as possible about the household. As we shall see in the next chapter, in 2018, Facebook, for instance, filed a patent request for the development of a technology titled *Predicting Household Demographics Based on Image Data* (Bullock et al, 2018). The technology would enable Facebook to ‘build more information about the user and his/her household’ and to provide improved and targeted content delivery to the user and the user's household’ (Bullock et al, 2018).

Household profiling is particularly exposed to algorithmic bias, because it relies heavily on sensitive data. In 2016, ProPublica, for instance revealed that Facebook was enabling discriminatory targeted advertising, allowing advertisers to target only ‘white households’. At the time, allegedly, Facebook replied that they were going to address the issue. Yet in 2017, ProPublica carried out a further investigation to see whether measures had been taken. To do so, ProPublica bought dozens of rental housing ads on Facebook, but asked that they *not be shown to certain categories of users, such as African Americans, mothers of high school kids, people interested in wheelchair ramps, Jews, expats from Argentina and Spanish speakers* (Angwin et al. 2017). In March, 2019, following the ad scandal Facebook had to implement measures to avoid such forms of racial and ethnic discrimination, yet the debate is still going on at the time of writing especially in relation to other sensitive data such as gender and politics (Gillum and Tobin, 2019). The Facebook ad scandal is just the most public example so far that

highlights how sensitive data is gathered through household profiling, and how open to algorithmic bias these technologies actually are.

3. Algorithmic Unaccountability - Individuals have historically been profiled on the basis of the families and the social groups they belong to. Yet, today these classifications, are made possible through algorithmic decisions that are difficult to understand or to challenge. AI systems that are trained on home life data and children's data in general, would be particularly exposed to inaccurate and biased automated decisions, and given the complexity of homelife data, these decisions may be impossible to explain leading to further AI unaccountability. The training of AI systems on home life data (or children's data in general) can have serious implications for children's rights today, and for future citizen rights in the EU. It not only reproduces existing inequalities and stalls social mobility but it can also impact on individuals right to self-definition, and non-discrimination. As our worlds become more and more AI-driven we should protect children from being profiled and judged for the values and choices of their family of origin and we should defend their right to choose what type of persons they want to become or which values they want to make their own

Concluding Remarks

Citizens today are governed through data in ways that were not possible before and thanks to AI innovation. Predictive analytics is used by policing and courts; biometric monitoring is a common practice by border patrolling; data driven decision making is used by governments to decide fundamental matters such as welfare provision or child protection. The governing of citizens through data implies not only that citizens are being datafied from birth, but also that they are exposed to all sorts of algorithmic error and inaccuracies. For this reason, we need to find political solutions that recognize the 'human error' in algorithms, and the fact that when it comes to human profiling *algorithms are always going to be inevitably inaccurate and biased*. We also need political solutions that discourage the use of children's data and the profiling of children. In this regard, I would personally encourage the EU Commission to add the following points to their White Paper *On Artificial Intelligence – A European Approach to Excellence and Trust*:

1. Make sure that ALL AI system that are trained to profile human beings and are used for data-driven decision making are considered high-risk.
2. Make it a legal requirement for those private and public sector actors who wish to use AI technologies for automated decision making on individual rights to : a) inform citizens that decisions have been made by an AI system b) offer the possibility to appeal the decision and request human oversight.
3. Make sure that AI systems aimed at human profiling and decision making, do not train technologies on the data of individuals gathered before the age of 18.
4. Make sure that all the children's data, which is collected through adult profiles - which are not designed and are not aimed at children and hence which are not required to abide to COPPA or the GDPR special children's protections – is deleted, not used and not sold or shared.

References:

- Angwin, Julia, Ariana Tobin, and Marlene Varner. 2017. "Facebook (Still) Letting Housing Advertisers Exclude Users by Race." *ProPublica*, November 21, 2017. <https://www.propublica.org/article/facebook-advertising-discrimination-housing-race-sex-national-origin>.
- Barassi, Veronica. 2018. "Home Life Data and Children's Privacy." Call for Evidence Submission Information Commissioner's Office. London UK: Goldsmiths University of London.
- . 2019. "TedTalk: What Tech Companies Know about Your Kids." Presented at the TedxMileHigh, Denver, Colorado, November 16. https://www.ted.com/talks/veronica_barassi_what_tech_companies_know_about_your_kids.
- . 2020. *Child Data Citizen: How Tech Companies Are Profiling Us from Before Birth*. Cambridge, Massachusetts: The MIT Press.
- boyd, danah, and Kate Crawford. 2012. "Critical Questions for Big Data." *Information, Communication & Society* 15 (5): 662–79. <https://doi.org/10.1080/1369118X.2012.678878>.
- Bullock, William, Liang Xu, and Li Zhou. n.d. Predicting Household Demographics based on Image Data. United States Patent Application, filed November 15, 2018. Accessed July 29, 2019. <http://appft.uspto.gov/netacgi/nph-Parser?Sect1=PTO1&Sect2=HITOFF&d=PG01&p=1&u=%2Fmetahtml%2FPTO%2Fsrchnum.html&r=1&f=G&l=50&s1=%2220180332140%22.PGNR.&OS=DN/20180332140&RS=DN/20180332140>.
- Clifford, James, and George E. Marcus, eds. 2010. *Writing Culture: The Poetics and Politics of Ethnography*. Second Edition, 25th Anniversary edition. Berkeley Los Angeles London: University of California Press.
- Dencik, Lina, Arne Hintz, and Zoe Carey. 2018. "Prediction, Pre-Emption and Limits to Dissent: Social Media and Big Data Uses for Policing Protests in the United Kingdom." *New Media & Society* 20 (4): 1433–50. <https://doi.org/10.1177/1461444817697722>.
- Eubanks, Virginia. 2018. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. St. Martin's Press.
- Friedman, Batya, and Helen Nissenbaum. 1996. "Bias in Computer Systems." *ACM Trans. Inf. Syst.* 14 (3): 330–347. <https://doi.org/10.1145/230538.230561>.
- Gangadharan, Seeta Peña, and Jędrzej Niklas. 2019. "Decentering Technology in Discourse on Discrimination." *Information, Communication & Society* 22 (7): 882–99. <https://doi.org/10.1080/1369118X.2019.1593484>.
- Gillum, Jack, and Ariana Tobin. 2019. "Facebook Won't Let Employers, Landlords or Lenders Discriminate in Ads Anymore." *ProPublica*, March 19, 2019. <https://www.propublica.org/article/facebook-ads-discrimination-settlement-housing-employment-credit>.
- O'Neil, Cathy. 2016. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown/Archetype.

Russell, N. Cameron, Joel R. Reidenberg, Elizabeth Martin, and Thomas Norton. 2018.

“Transparency and the Marketplace for Student Data.” *Center on L A w a Nd Information Policy* 4 (June 2018): 2–33.

Woolgar, Steve. 2002. *Virtual Society?: Technology, Cyberbole, Reality*. Oxford ; New York: OUP Oxford.

Prof. Veronica Barassi campaigns and writes about the impact of data technologies and artificial intelligence on human rights and democracy. She is an anthropologist, and Professor in Media and Communication Studies in the School of Humanities and Social Sciences (SHSS-HSG) at the University of St. Gallen, as well as the Chair of Media and Culture in the Institute of Media and Communication Management.

Barassi believes in the importance of publicly engaged social research and consults regularly for companies working on ‘privacy by design’ and non-profit organizations worldwide. In 2018, the Information Commissioner’s Office of the UK Government used her research as evidence for the development of age appropriate design code -- and in 2019 the Irish Government invited Veronica Barassi to discuss AI Ethics at their Digital Summit. Her Ted Talk [What Tech Companies know about your Children](#) has attracted hundreds of thousands of views.

Barassi's previous research focused on social media and political campaigning. She authored the [Activism on the Web: Everyday Struggles against Digital Capitalism](#). Over the last three years, she investigated the impact of children’s data traces on their civic rights. Her most recent book [Child | Data | Citizen: How Tech Companies are Profiling Us from before Birth](#) will be published by MIT Press in December 2020. This report combines the finding of her *Child | Data | Citizen: Data Traces, Family Life and the Digital Profiling of Children* (2016 – 2019) with the preliminary observations resulting from her new project *The Human Error Project: AI, Algorithmic Bias and the Failure of Digital Profiling*. (launch September, 2020).

Disclaimer: The views in this response are my own and do not reflect the opinion of my Institution or the colleagues at the School of Humanities and Social Sciences and in the Institute for Media and Communication Management at the University of St. Gallen (Veronica Barassi, 14th of June, 2020).