



## Episode 4: Kids and Emotional AI

As smart toys, virtual assistants, and machine learning apps spread across homes and schools, an increasing number of children are now living, learning, and growing up around artificial intelligence or “AI”. In this episode, Dr. Sara Grimes (Director of the KMDI) chats with Dr. Andrew McStay, Professor of Digital Life at Bangor University (Wales, UK) and Director of the Emotional AI Lab about the ethics and impacts of AI technologies designed to read and respond to our emotions, and their growing presence in children’s lives.

Dr. Sara Grimes 0:00

In late December 2020, just before New Year’s. News erupted about a pretty scary interaction that had taken place between a 10-year-old girl and Amazon’s virtual assistant, Alexa. The girl had asked Alexa via her family smart speaker for a challenge to do. Alexa replied by suggesting that the girl should try plugging a phone charger halfway into a wall outlet and touching the exposed prongs with a penny. The dangerous and potentially lethal penny challenge had been circulating around the web for some time, which is where Alexa’s algorithm found it. And identified it as a popular and therefore appropriate response to the child’s request. Every so often, a story like this makes the headlines, and we’re reminded that increasing numbers of children are now living, learning, and growing up around artificial intelligence or AI. The stories aren’t always as disturbing as that penny challenge example. A few years ago, there was a heartwarming incident in which a four-year-old in the UK saved his mother’s life, by asking Siri to call the police. And Siri responded by connecting him to an emergency dispatcher who then sent an ambulance to their home.

What these events all show us, however, is just how little we really know about children’s relationships with AI. What do children and think about these seemingly knowledgeable and empathetic voices coming out of a growing number of their devices and toys, both at home and at school?

According to Juniper Research, the number of virtual assistants in use worldwide reached 4.2 billion in 2020. Ad Tech magazine has published a number of articles about teachers using AI in the classroom. Turning virtual assistants into assistive technologies for kids with disabilities, or just asking them to keep track of the time. Meanwhile, we've seen a range of AI connected children's toys hit the shelves in recent years, from smart robots and stuffed animals, to the short-lived hello Barbie. The privacy and safety implications are obviously staggering, but what about the emotional or affective implications? Do children develop feelings and attachments to the AI entities that inhabit their everyday lives? And are these AI technologies designed to treat children ethically to respond to their feelings, their diverse needs, and their inexperience in responsible ways?

Enter the Emotional AI Lab, an international research group led by Professor Andrew McStay, that's dedicated to understanding the social and cultural impacts of AI technologies designed to respond to our emotions and moods. Dr. McStay is Professor of Digital Life at Bangor university in Wales. And the author of *Emotional AI: The Rise of Empathic Media*, published by Sage in 2018. Along with the rest of his team at the emotional AI lab, he's conducting critical research on the growing array of technologies that use effective computing and AI techniques to sense, learn about, and interact with human emotional life. Their projects include explorations of how various industries from advertising to car manufacturing have started using AI to read and react to people's emotions and emotional states.

One of their projects focuses on children and smart toys. And they've become a leading voice on children's rights when it comes to AI. Dr. McStay cutting edge research is published in numerous places, including two notable articles in the journal, *Big Data & Society*. First entitled *Emotional artificial intelligence in children's toys and devices, ethics, governance, and practical remedies*. Co-authored with Gilad Rosner in 2021. As well as *Emotional AI soft biometrics and the surveillance of emotional life*. An unusual consensus on privacy in 2020. The emotional AI lab submitted a really important and timely response to the UN committee on the rights of the child's general comment, 25.

And Dr. McStay currently working on a new book on the topic called *Automated Empathy*. 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 speaking with Dr.

Andrew McStay about his two recent articles in Big Data & Society. To find out more about his examination of children and AI toys, and his larger body of research into the emotional applications and implications of AI. So let's just jump right in. What is emotional AI?

Dr. Andrew McStay 4:56

Yeah. So I think it's probably useful to kick off with a little bit of understanding of what AI is. And I think, in its simplest terms, it's the idea that technologies are able to do what minds do. And that involves sorting, classifying, judging. So I think, this idea of emotional AI it refers to technologies that use effective computing and AI to sense, learn about, and interact with human emotional life. And I think you know that, that's what emotional AI is about. It's about sensing, it's about learning and it's about interacting with human emotions. So that doesn't mean understanding what emotions are, it doesn't in that emotional AI fields, but rather it's that ability to sense, learn, and interact in some way, shape, or form. So, in terms of classic AI, it tends to be talked about in relation to either strong AI or weak AI.

So strong AI is the idea that event AI will going to have consciousness and it can be fully fledged. Weak AI is something else. It's much more specific. So in relation to emotional AI, it's the idea these technologies will read and react to emotion. And that will happen through a range of modalities, such as text, such as on Twitter, the voice such as through a smart assistant, computer vision, such as through cameras, biometric sensing such as on wearables.

And potentially direct from the person themselves, but maybe also in relation to where that happens to be, and what they happen to be doing. So I think for me this idea of emotional AI itself. We're going to get into issues of toys. We're going to get into issues of children. But I think perhaps the first thing for listeners to be aware of, it's that this idea of emotional AI, this premise of emotional AI, we begin to see this in a range of different life context, in a range of different situations.

So that involves things such as smart assistance. It involves things such as cars, games, mobile phones, wearables, toys, marketing, insurance. They even have such as policing, education, and even at national borders. So when we think about this overall interest, and emotions. In many ways it's about optimizing emotion. In many ways it's about profiling

emotion. But again, yeah this has taken place in all different domain of social life. And I get it as I think we'll probably get into. This raises all sorts of ethical questions.

Dr. Sara Grimes 7:46

A term you use in your work to describe what these underlying technologies are doing through these processes that I really like is, "Feeling into" where does that term come from? And what does it mean?

Dr. Andrew McStay 8:01

Yeah, so it draws from longstanding interest and empathy. And yeah, I think empathy is, it's a good way of looking at these types of technologies because when we think about what technologies are increasingly doing. They're reading, they're sensing bodies, they're reacting too. And for me, this shows qualities of cognitive empathy. So what I mean by that is something quite different from sympathy. So sympathy is where we're really with somebody and we really ... Our heart is in what somebody's going through. We feel that heartfelt connection. But empathy and sympathy are quite different things. A slightly and bleak example, but a torturer can empathize in that they know what buttons and what triggers to push. So empathy rather has a cognitive dimension. It's about reading. representations of sports, too, and how they're shaped into different symbols within these conversations, too.

It's about sensing. It's about reacting. It's about placing in context of what is known about that person. And I think that that notion of empathy starts to characterize what these emergent technologies are doing. So the reading, sensing, and reacting. And that certainly includes emotion, but includes other elements such as fatigue, stress and things which certainly overlap with emotion, but are not reducible to emotion. So yes, one of reasons I use this term empathy. Is it slightly somewhat broader reach. I think as well, in terms of almost think of from a design point of view as well. You know that when we think about kind what empathy is, we could think about designers of systems that will feel into people, but also people on... As I'm sure will get into children, can also feel inter systems to understand how they work, what their functioning is, and what synthetic psychology buttons to press in the AI itself.

So I think this idea of feeling into is quite useful in terms of systems that feel into worse

via cognitive means, but also how we feel inter-system. Just to give you a slightly tangible example of that. For example, in the workplace in tele cell you'll have a system that is monitoring a call worker with monitoring the emotional tone of that call worker, the monitoring enthusiasm levels of that coworker. But when that coworker begins to work out what the system is looking for, they will perform, they will behave in a way to gamify what that emotional AI system is looking for. So it affects their feeling into system. We kind have a two way process taken place, a mutual feeling out if you will.

Dr. Sara Grimes 11:02

Your recent article in Big Data & Society coauthored with Rosner focuses on emotional AI in children's toys and devices. This work expands on the small, but emerging literature on the internet of toys and shifts the focus to the incredibly important role that emotion and effective relationships play or could potentially play in children's experiences with AI toys, smart toys, or what you refer to as emo toys. Can you tell us a bit more about emo toys? What are they? And what do they introduce to children's culture?

Dr. Andrew McStay 11:39

I can. I feel indulged with slightly, it'd be useful just to have a slightly retrospective look at the history of toys. So clearly things such as dolls and obviously toys that have been around since time immemorial, but the idea of dolls and toys that are able to simulate life processes have actually been around for some time. So, you know it doesn't just begin with this overlap between AI doors. It goes back 1920s, if not before. But for example, in the 1920s, the Dolly record was able to speak nursery rhymes through a use of phonograph. By 1950s, Mattel's Chatty Cathy was able to utter 11 phrases, including I love you. From 1970s onwards began to see more electronically or into toys built with transistors and built with circuit boards.

And these were able to play prerecorded messages. And by 1980s, interactive toys were pretty well established. But I think one of the pivotal decades was the 1990s. And 1990s really saw new development. So I think, whereas in the past toys were something which a child projected their feelings, projected their emotions onto. I think, small toys in the 1990s involved new forms of mutual engagement and critically their involvement in the toys emotional state itself. So these were semi autonomous toys. So for us the most famous is [inaudible 00:13:20] , but you know it include Furby as well.

By 2010, we saw the emergency connected toys which involve internet, Wi-Fi, Bluetooth, and so on. But I think for us, for myself and Gilad. Gilad Rosner, we're really interested in this overlapping between connected and smart toys that use connected protocols. Use different types of sensors, use voices, use image recognition and have early self-learning algorithms.

I really scope to interact in new in different ways. We were interested, kind it well with the overlap here with the trends going on in the emotional AI sector. And there were a few toys. That been I've doing this for a while. So Sony's probably one of the big ones. Sony's AIBO. So they had one of their robot dogs out in the late 1990s. And another one out in 2010s. That was using emotion profiling. Plus one more well known, certainly in the west is anchors Cosmo. And again, that used facial expressions as a means by which the toy would react to a child's facial expressions.

It would also recognize the Charles is not just about reacting to the expression, but also the identity of the child as well. And I think, you know given where we do see technologies heading through biometrics, and through emotion sensing, and this wider development in synthetic personalities. And you know it's really complex. In terms of how you study this. We thought well, we could do a child focused study. But we were mindful that people and fellow academics who were better are kind working with children, then we. We were new to this.

So we thought, well, why not? Let's do adults. So we did quite a large body of work. We interviewed a range of experts. People with expert in the toys' industry. People with the bulk policy interest in toys and NGOs working in this area. We also did a big survey in the UK. I think it was around a thousand pounds to gauge their attitudes towards toys that function in relation to human emotion. Well, I think there are a number of issues which came up for us. And I think one is this issue of ambivalence. And ambivalence in that parents were really keen that children are exposed to, and immersed. And UNCRC language have access to new and emerging digital services.

So there's almost a for more thing. A fear of their children missing out. So that's on one hand, there's almost kind of slightly pro-technology approach to this. And then the ambivalence pushed the other direction as well. And that there were inevitable concerns about privacy, but also as well. And this was really interesting. Parents were really

concerned about being replaced somehow. And that means, for example, in relation to home assistance. So a picture Amazon Alexa or any other voice assistant. You know if you have toying child oriented version sat in the corner of room. It's something that could read stories without being fatigued. It could be a confidant without judging. And I think from a parental point of view, they were really concerned. Surprisingly concerned I think. This idea that they might be replaced in some way, shape, or form. So that's not to be read of the AI's taking over and that kind of thing. In short, again I underlined that word ambivalent. That they were pulled in separate directions.

Dr. Sara Grimes 17:19

In the article. You describe another recurring theme that came out of your interviews. Which you call generational unfairness. What sorts of issues or concerns were your interviewees raising here? What are some examples?

Dr. Andrew McStay 17:23

Yeah. In terms of our assessment of the technologies themselves and what came out of the expert interviewing. It really was this issue of generational unfairness. And I think, we see this particular government in a climate related issue. And perhaps we can just expand the metaphor slightly in relation to digital ecologies as well.

In that these radical built environments, they're adult processes, that adult regulation laws. And I think a real concern from the experts that we spoke to and from our own assessment of how the technologies function, this issue of generational unfairness was a real problem that has roots in the critical literature around children's. So I'm thinking in relation to sharing, renting, exploitation of children's digital footprints.

But I think for us, this is growing further. This issue identified in sharing team is growing further to encompass questions about biometric privacy. And I guess, this kind of dates fight commercialization of childhood and possibly as well, the installation of questionable relationships between children and synthetic personalities. You know if we take a company, such as Mattel, these toys are imbued with values that reflect the brand, certain gender identities, normative identities, rank out of sexuality, jobs and so on. So I think that when we think about the idea of synthetic personalities being imbued with these values, this is really deeply questionable stuff. And I think once we start emotion

based profiling into the mix, natural language processing and more natural and enhanced interaction between children and synthetic agent. I think that really raises issues. I think quite profound issues.

And I think, you know we're just at the start of all of this. I think toys such as Cosmo which are based on analyzing facial expressions. They're quite gimmicky. But I think, look at where AI is heading in terms of the abilities to react in ways that are appropriate to the question being passed. I think we're in the start of something really big here. Then one of the things that we found in that paper. It's really important at regulatory level in terms of laws that are made regionally and internationally. But I think in terms of self-regulation as well, that could be within company or could be just larger industry standards. Yeah. There's a real important need I think, to take on board these ambivalent issues.

Dr. Sara Grimes 20:22

So I know the children in AI studies, just part of a much broader research agenda of your lab, the Emotional AI lab. Which explores the social and cultural impacts, AI technologies. Specifically, AI that feeds on data about human emotions. What are some of the other trends and technologies you and your team are investigating right now?

Dr. Andrew McStay 20:41

Yeah, so we've got a number of projects taking place, but one of the larger projects is cross-cultural analysis between the UK and Japan. So working with a range of Japanese universities and working with a range of UK universities as well. We're looking at this question, regard of emotion, biometric and sentiment analysis for that matter in a very international context. But really one of the reasons why I mentioned the UK and Japan, is that the philosophical context is very different. The ethical context is quite different, and the social context is quite different as well. I think sometimes that's something that's really missing when we have these tech ethics based conversations that largely come up [inaudible 00:21:31] Western liberal standpoint. And of course there are other regions around the world and some of those regions are much bigger than north America and certainly the UK.

So I think, it's really important I think to have. At least be sensitized to other frameworks and other the modes of understanding as they relate to technology. And then of course that expands down into the different instances of where emotional empathic systems

have been deployed. So that involves things such as smart ads that we don't just look at but they look at back at us. To gauge our facial reactions, possibly even body temperature as well. And then kind of change the ad themselves in relation to the responses that they're getting from the personal group of people, looking at the ad. But one of the other areas that we're looking quite closely up is the overlap between emotion, politics, and disinformation and misinformation. But particularly disinformation so to give some obvious examples of that, social media can get a bit emotional at times. People can get a bit heated.

We've all had these expressions such as filter bubbles. But these kind of discussions about news and the online environment. These are very much about sentiment analysis. It's about use of words and it's about clustering word types to label a particular emotional or reaction in particular set of responses. But then what next? What comes next after all of? We've had kind a various kind of US elections, which have involved issues of emotion, some quite famous ones. In the UK, we've had Brexit, which is the UK leaving the European union. And these were really fine margins. So you're talking to margin of the 2%-4%. So, in terms of the role of profiling technologies to micro target ads and to filter what kind of things that people are exposed to online. It clearly played role. But yeah, in terms of what NEX for is there, well, there's all sorts of data type. And we're really interested kind of what happens when we input biometrics into all of this.

So for example, in relation to voice systems, our voice systems use third party apps, these collect data as well. And these are absent systems that are in the home. Kind of quite intimate conversations. And that's not to say that Alexa is listening into words, but I think over time these ambient systems will begin to feed into data ecologies, and ways that people are segmented on profile. And as soon as we get into that issue of biometrics and human physiology, which of course is an extricable from the motion. We're moving into a new world of profiling. So yeah ensure one of the big questions for us is what happens when we cut of our biometrics and emotion to the question of disinformation?

Dr. Sara Grimes 24:38

A lot of the work you do has a clear policy focus. Your research advances the idea that now is an opportune time to introduce new regulation in this area. I'm thinking specifically about your other article in *Big Data & Society*, published in 2020. In that article, one of your key findings was a "Weak consensus" among stakeholders when

it came to emotional AI and privacy concerns. Can you talk a bit about this, and its implications for AI regulation?

Dr. Andrew McStay 25:09

Yeah. So it's an interesting paper that one. In essence, what I was interested in was, again, the use of emotion based systems and a range of different life contexts. And I was interested in talking with people from industry, people from a policy background, and people from the NGO sector. As you know, I did the interviews. I became increasingly surprised. So for example, you know interviewing in San Francisco and south tech startups, I expected it to be friendly, but kind of quite libertarian. You know pre-market, no rules, unaffected innovation, all that stuff. And that wasn't what they were saying what they wanted. They did actually want rules. They didn't necessarily want excessive rules, but they wanted rules that would apply across the sector.

And I think for startups they wanted to be on a level playing field with the larger technology BMOF. But that for me was really interesting that they wanted rules. They wanted that level playing field. And they were okay with those rules being relatively tough, as long as everybody abided by them. In Europe as well, talking with companies here, they were saying something similar that they were okay with rules as long as they were equal and fair. That was one part of all of this. So in interesting rules from industry. And then talking with the NGO sector. So for example, I talked with AFF in the US. And I talked with open rights group and Privacy International here in the UK and Europe. And so you might think an NGO, this idea of profiling of human emotion.

You might just think there'd be a ban all total aversion. And they were like, well, actually, they could see some interesting use cases. And with strict limitations in relation to where the data goes, who gets access to it, no third party usage, all that kind of thing. But I think, you know NGOs and people kind of work for these organizations. Typically, pretty interested in technology, and what it can do, and social benefits. And so from the policy area, obviously got a real interest and ethics searches, privacy and got related topics. But again, they were quite open to both the industry standpoint, but also I got a more critical NGO standpoint. So where seem to me is that there is a consensus of sorts between each these stakeholder groups.

Now the motivations for each of these actors are a little bit different. But it seems to be that there's almost a temporary alignment between these areas, which basically means that now is a really, really good time to start regulating in the area of emotion, recognition, emotional AI. And I guess wider automated empathy systems. And we're doing that in Europe. We've had the proposed AI regulations. So that's an addition to the general data protection regulation which is the large kind of gauge protection implement that spans Euro. That's going to be extended with an AI regulation. We've seen the proposed version of that right at the top is article one discussing emotion recognition. So it seems like that there is a move to actually kind of focusing on emotion recognition now. And I think, yeah possibly because of this weak consensus and this temporary alignment.

Dr. Sara Grimes 29:13

I'm going to circle back to kids now. The United Nations recently adopted a new general comment, confirming and outlining how children's rights apply in the digital environment. I know you are very familiar with this development as your lab submitted a response to the UNCRC's public consultation. Can you please tell us about some of the issues outlined in that response?

Dr. Andrew McStay 29:36

I think, yeah. For our own submission we focused in particular on the question of emotion. We thought others are much better equipped to offer kind a wider conversation. Yeah. I was focused on emotion. You know, there were current number of areas that we've really foregrounded. And one was about relationships. And we were really wanted to flag back products. And of course the people and organizations that build these products shouldn't abuse the trust of children and their willingness to engage with objects or non-human personalities.

And I think that issue of trust with the obvious offshoots of manipulation and so on. There seems to be such risk for us. That was something that we really wanted to flag, but, you know there are really thoughty issues as well. Which we highlighted such as, what actually happens if manufacturers do detect something. Let's just fast forward kind 5-10 years down the line. That, you know these systems do exist in a child's bedroom.

What if they do detect ill mental health? For example. What if they do detect child abuse? That raises really, really, really difficult question. But more broadly we're interested in

kind issues of stereotypes. And again, this is the synthetic personalities point. And that stereotypes are assumptions about emotion. You know they really need to be considered in the design of these objects. So, for example, girls as passive, boys as active. You know these are just assumptions that have long been built into toys. Again, when we're thinking about systems that children are feeling into and in reverse feeling into children, these stereotyping issues become quite acute, I think. In relation to issues of play we hope that if these systems are to be developed, we've got the strict restrictive controls.

You know this should be used to amplify child's imagination. And a present of kind of the toys and services that we've surveyed so far. We don't really see that the toys seem to be doing that. So for example, we talked about anki, A-N-K-I, which is a company and their toy Cosmo. It's just a little robot digger. It has a cute little phase on the front. It reads the facial expressions of children and it edit behaves acutely in relation to child behavior. But I'm not sure what that does for imagination. You know if these systems are to be developed, it would be interesting. And perhaps useful to kind of positively interact with a child imagination. I think, when GC25 came out. Once, everybody's input had been digested and they wrote up them the findings and the original final GC 25 came out really pleased to say that emotion was end there.

Dr. Sara Grimes 32:41

Now that the general comment 25 has been adopted. Do you think it'll have any impact on the issues you've raised around kids in emotional AI?

Dr. Andrew McStay 32:50

Yeah. I think it will have an impact. It's going to take a while to percolate through in to the conversations where this stuff needs to be addressed. You know, it's already fed into, in a significant way into the European proposed AI regulations. I don't see GC 25 being taken out of the eventual error regulations that see light of day. So already is inputting. I think, you know it's worth mentioning for any listener. Who doesn't kind follow. Got a law policy and regulation that closely basically where can a European policy making goes to an extent the world certainly takes notice if they don't follow it. So north American companies, for example, typically and often design around GPR standards. So, you know the fact that GC 25 is already in an influential piece of European regulation is really significant. And is really quite heartening. In other areas as well.

The existence of GC 25 again, has really [inaudible 00:34:00] front and center of any industry facing discussion about yeah, how to design new products that involve AI, but also, for the stuff that I'm working on. Particularly on emotion and empathy based stuff. And I would hope as well that within industry itself, although, they might not be talking about this quite so publicly. It's GC25, it's the UNCRC. It's kind of the biggest human rights framework for children. Now that there is something explicitly addressed to technology to medium and so on and so forth. Internally, this will be being discussed. So I think, although, it hasn't quite got the fanfare that some might have wanted. I don't think that should downplay either A, the significance, but B the impact it will have. Because again, it's work on progress. It's going to take a few years to filter through evolution rather than revolution and all of that sort of thing.

Dr. Sara Grimes 35:11

A big thanks to professor McStay for joining us today. Please follow the links in the podcast description to find out more about Dr. McStay research. The Emotional AI lab. The 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 [inaudible 00:35:38] music by Nicholas Manolo. Theme song by Taikon 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.