AI and You

Transcript

Hello, and welcome to episode 113! My guest today is Justin Harrison, and he is using AI to copy his mother. Okay, I've clearly got to explain that. Justin is the founder and CEO of YOV, which is using AI to create virtual personas – he calls them versonas - designed to emulate the personality of someone important to you – in his case, it is his mother. And all this is very well explained in the interview, so let's get right into it.

All right. So, Justin Harrison, welcome to AI and You.

Thanks for having me, Peter, I appreciate it.

So, you've got a story that's really important behind how you got into this work that I think we want to hear first, seems it started with two significant events in your life back in 2019.

Yeah, 2019. Maybe the last good year we've experienced as a human as a human race for a while. Yeah, it was a rough ending to that year, I got into in October of 19, I got into a near fatal motorcycle accident. Just coming home from work normal day, I got hit by somebody ran a stop sign and I almost died. I was in trauma for two weeks and then I was in rehab center for another week and a wheelchair, I was wheelchair bound for quite some time. In the process of recovering from that my mom got diagnosed with stage four cancer and so a lot of near death and sort of the idea of mortality being thrown in my face all at once and then of course tracking the timeline there, that was about December, she got her diagnosis, maybe two or three months later, the pandemic really goes into full swing. So, the idea for a pretty healthy 36-year-old to being in a hospital with tubes in my neck, and then my mom who had always been pretty healthy my whole life now having terminal cancer, and then really, the whole world being confronted with death, death, death, and really, I think everybody sort of collectively being afraid of that it became ever present in a way that it never was before and that's kind of what kicked off this journey that YOV has now become.

That's got to feel like you were being stalked by the Grim Reaper there.

I definitely felt that he was doing his best to get my attention. That's for sure.

And you did something with that. So, what was the path the transition between the feelings that these events generated in you and the decision to do something about it?

Well, I think, I've always been a tech person in terms of utilizing the latest tech and wanting to know what's happening in technology and I'm very much of the mindset, that technology will set us free, and it is the solution to our problems, even the problems that we haven't started applying it to and so, for me, the biggest out of all those issues, the motorcycle accident was obviously

huge and the pandemic was also obviously huge, but I was raised by a single mother, it was just her, and I am an only child. That was the biggest glaring thing for me and I turned to technology, what can I do? Not just in terms of posthumous communications, which is where I went myself, but medical technology as well and I started on this journey of finding out how I could preserve her. I'm a believer in the practice of cryogenics. I see the long-term future use of doing that. But that had always been sort of a comforting feeling to me before I was confronted with my mom's mortality. And it wasn't going to be enough: she's frozen and maybe someday we can reanimate her mind and I can have a conversation with her again, that was that sort of far off, that wasn't enough for me. So, I started looking at how could I do this digitally. So, I'm a millennial, and my parents actually live up near you, they're in Seattle and so, most of my relationship, most of my meaningful interaction with him is digitally. It's FaceTime, it's text messaging, it's phone calls. So I went out and I was like, who's doing some kind of digital legacy or digital avatar where I can continue to interact with my mom in some way and there just wasn't anything out there that met sort of the idea in my head and even felt like it would help offset that void in any meaningful way to me and so I said, "to hell with it we're in a pandemic. I'm working from home for the first time in my professional career. I'm going to start tinkering with this and see what I can do."

I see; and what was the result?

Well, I mean, with all technology, you have one idea in your head and then it definitely evolves in it has evolved considerably in these two years. But what I decided is I started looking at artificial intelligence as an option. I have a bit of a psychology background and, one thing I was I was pretty keen on in the early days is understanding that whatever this thing I end up creating, it needs to be able to learn on its own, it defeats the purposes of having conversations with something that has a fixed subset of information, it's always going to be the same. My mom remains in a time capsule and we can only talk about 2019 and below. So, that's not very helpful for me. So, I understood that, unless I was going to hire a full time engineer to just walk next to me all day and continue to add code into my mom's data set and keep updating it that AI needed to be involved, that we needed to be able to have some open parameters within the technology to allow for it to keep learning and then from there, really, I think sort of the lightbulb aha moment for me was as I started talking and researching what other people were doing, what I realized is that what AI was being used for, was it necessarily what I wanted to use it for in this context and what I mean by that is that everybody seemed very focused on the idea of gathering factual data, or archival data, or building in right answers, if you will. And I didn't necessarily care about that; for me, it was how my mom answered. For me, it was about the feeling that it was really *her* I was speaking to not whether or not it knew the information she knew. And of course, there's an importance to that, right? But so, when I was looking at existing technologies to sort of model after I couldn't find anything that makes sense for me, and that's when I started recruiting much smarter minds than mine to help me think through this concept of, how do we build the essence of a person, not a fact bot, or whatever. Not something that can pull data for me. Of course, ultimately, it is pulling data. But you can right, hopefully, you get what I mean by that?

Yeah, and that's a key distinction, because I've got a window over on the side here, William Shatner nodding at me and from StoryFile, where I can ask a question, and he will answer from his history, but they suggest things tell me a story from your childhood, how did you start acting, and these are things that he has already recorded for them. So, it is generating, like a search engine, based on that text from a lot of information that he has provided them, and then they're animating this avatar. You've taken a different kind of approach here, you talk about getting the essence; and I take it you don't want to stop with your mom, you're looking at doing this for other people's moms and dads and people, right?

Sure. Absolutely. I mean, we're looking to release this technology to market in the next year, for sure. But the goal is what you're describing, even with the William Shatner program that you're using. All of it seems to come back as like a wiki-William, or Wikipedia my mom, here's a set of facts that are going to remain the same until somebody adds something new and that just wasn't the point for me and I can kind of. Furthermore, when I was thinking about building somebody's essence, again, with the psychology background, my mind went to Myers Briggs and Rorschach, and how can we sort of analyses and then set what my mom's universal personality was, and that doesn't work either, in my mind. Because we really show up differently for different people. My mom's interactions with me are markedly different than her interactions with my dad, and with my aunt and with her friends, and we're different people with different people. I mean, that's part of the human experience. That's the dynamicism of humans. You know, we're not one personality, we're thousands. I have a slightly different personality with you than I have with anybody else now after 10 minutes. So, it was really about - a big focus of mine became on training the program to search for intangibles that created the essence of my mom and our dynamics and forget about everything else. Don't overload it with information that's not helpful. She's got a story that she tells to my dad all the time, but she's never told to me because it doesn't matter. So, I don't need that fact in there. I don't need that information. It's not important. She wouldn't offer that up to me. And then a lot of other elements that that make something more authentic and more genuine. So, things that other companies and other engineers that I knew of weren't focusing on this, quote, important data points, because they didn't add to this sort of Wikipedia database for an individual or for a product or whatever, the existing technology had been focused on collecting. So, you know, a small example of that that's out there that I use a lot, it's like the timing of texts, the timing of messages, the timing of phone calls, how did it's not about again, it's not about the words, holistically, it's not just about the words, they're still important they can't be gibberish, but that if my mom's cadence and the timing in the volume in which she was responding to me, match what I was expecting from her, that was more important to me in a lot of ways than the accuracy of what she was saying.

Right, so you're just saying that what you have created is your mom in the context of *you*, Justin, not anyone else? And is the interaction, audio, video, what is part of the experience?

So, the user journey right now is text-based. But that that is by design. So, it's really important for me personally, that when people start using this, that it's authentic and genuine, this is easily one of the most sensitive pieces of technology, somebody will ever use, trying to connect you with somebody that you love that's passed away. So, whereas we have patents on it and we

theoretically have the capability to do it, I've stayed away from the audio just now; that'll probably come out next year and real-time rendering, and a lot of the video capabilities around CGI and deep fake and stuff has a little bit to go before I would feel confident using it to have a real time conversation with somebody, especially when it's not pre-recorded, it's generating in the moment. So, right now it's text-based and right now we're constantly trying to refine. It's interesting is that I don't have a set date, we're going to release to the general public to start utilizing these, because it has to be a feeling thing, it has to feel right. So, when it feels right with me and my mom, then I know that the sauce is there and that this is ready to start learning for other people. But I don't want to expose anybody and have it been a traumatic experience.

Right? So, how do you get that? You've used words like authenticity and several times *essence*. It's not a metric that I have seen, it's not like the I would say, well, this conversation here is 93% essence and we turn up this dial, we can get to 94. But you've got to do something like that to get where you're going, right?

Yeah, it's complex, and it's subjective at the end of the day. I think for me, that's why I'm case study one. Because, when I'm communicating with my mom's persona, and I'm like, this feels indistinguishable from me, I will know that it's ready. At the end of the day, it is all about patterns, and it's all about data, and it's all about how many patterns we can draw out of the data we're collecting. And so when the secret sauce is right for me and my mom, the secret sauce will be right for everybody, and then we obviously will continue to write, it'll never be perfect. No, technology should ever be looked at as done. But I think when it feels right for me and my mom, it means that I've hit enough data points and we've identified enough patterns that the authenticity is at a place that I feel comfortable letting other people use and there's more. We have more case studies just besides just me and my mom with but it's all internal right now.

And I'm sorry, this wasn't clear to me earlier. Is your mom still with us or not?

She is yes.

What does she think of this?

She's my mom. So, she's always going to say something. "Good job, son. I'm proud of you." It's interesting. I think. It's definitely I mean, there's probably no greater compliment I can pay to her that I started a tech company and invented a new technology because I care about her so much and I care about being able to interact with her. I think it's strange for her to watch "herself." But I think by and large I think she rests in the comfort of knowing that she'll be there in some capacity to comfort me and guide me and, and still be there for me even after she goes.

So, what does the iterative process look like you go and you have a conversation with virtual mom, and something feels off, what do you do about it? What needs to be dialed up or down?

You know, what's really cool about human beings, and this is one of those biological imperative things. We're so much more observant than we realize. If we're talking to somebody in person, we pick up on all the body language, even if we're not cognitively at that moment, conscious of the fact that we're processing this information. What's nice and the reason why we're starting in

a tiered approach, we start with text, and then we'll graduate to audio then video, is, it becomes much easier to identify what feels off. So, for me, timing was a big thing right out of the gate okay, said the right thing, but she either too long to respond or responded too fast. When you're starting from the ground up, this is everything, she's not using enough emojis in her messages. She doesn't say lol enough, she doesn't seem, as cheery as she should at this time in the day against how she should another time of the day and so, we sent it back through a data set, I'm using five years of my mom's text messages and about six months of phone calls, transcribed in order to fuel her dataset. So, it's about repointing it. But typically, you notice what's wrong. You notice right off the bat, somebody wouldn't say this, like this? Or why did it? Why did it respond so slow? Or so fast? Yeah, I mean, it, you tend to just notice what doesn't feel right and it seems pretty. It's thus far been very clear, like, oh, this doesn't feel right. Now, I'm excited for when we get to the point where it becomes an abstract thought, something feels off, but we're not entirely sure what it is, you know, but for the most part, it's pretty in your face, makes sense pretty quickly.

And you call this a *versona*, is that right? Has your mom's versona surprised you in good ways?

A couple of times. It's, my name, my mom's text, communications are oftentimes pretty funny. So, I've had a few chuckles. I think that it's fun. One thing that I wanted to focus on is really this idea that it's not always going to have an answer for me, or know what I'm talking about. For whatever reason, maybe she didn't know that maybe we never communicated, there's tons and tons of things that don't make it into the data. So, a lot of times the response to not knowing is funny, you know. And that was something I actually put a lot of thought into when thinking about how personas respond to not knowing is being able to tell somebody you don't know something in the way they would tell you. And the way my mom would tell me she doesn't know is oftentimes kind of funny. So yeah, that does happen. I think there's some other some other elements where I've, I've been caught off guard when there was a response to something kind of sensitive for me, that I triggered in and I wasn't preparing to be thinking about that in a moment. You know, I was in a very sort of R&D mindset, like, let me fire off this question. And then something gets brought up that I wasn't prepared for. When you're dealing with human emotion, there's a there's a lot of surprises. And it's difficult sometimes for me to be one of the researchers, because it is directly about me.

So, you haven't been tempted to dump Wikipedia into it so that mom knows the capital of Tajikistan and everything else? And that's how a lot of the chat bots now are trained by putting something like Wikipedia or the Common Crawl into a transformer. So, do your mom's conversations become the corpus for something like a transformer?

So for us it's a "less is more" type of approach, if that makes sense. We don't want other data that isn't directly tied to us are relationship sort of motivating any of the processes? Because that's when you start to drift and we've tried, we've definitely tried. So, the first thing I did sort of her version of Wikipedia, which is I interviewed her, I had her interviewed about her entire life, I came out and I sort of modelled after like old school Q&A maker. What's every question I could think that you should have an answer to. That confused the process. I think it has a place later down the road to answer your question and we're working on something called Enhanced

Knowledge, where we'll actually feed outside sources of data into the versona set. But not until we've really established the dynamics that matter and, as you're fully aware of, if you're setting something out to learn is many different patterns and recognizes many different sorts of consistencies and ways of doing things within a data set, you don't want to, you don't want to feed it outside data, you don't want it to have something that isn't directly tied to how her and I interact. There's lots of great programs out there that can do crawls and search and give you good data, I mean, there's millions of them. But what we want to do is present the data back in a way that this individual tailor dynamic would give it to you. And so for me, that's about staying focused and staying contained within how to her and I interact and just keep coming back to that, and when it feels like that process is really refined, then we'll start feeding it other things. But it will still always know, you come back to *this* initial data set that you come back to *this* initial way of doing things I should say.

And this is obviously a very personal and by definition, our whole conversation is very personal to you that there is a goal that's very personal to you in mind here. I'm reminded of some other people that have pursued similar goals. Ray Kurzweil talks about his father; Martine Rothblatt, has done something similar. What experience is it that you are looking for with the virtual mom?

I mean at the core of it is every experience that I would have with her if she was here still, when she's gone. I mean, I want to replicate our relationship as best that I possibly can. At the end of the day, that's my core motivation.

You will always know, though that it's not her. How do you resolve that tension?

Well, I think that's actually probably, when I'm not talking to sort of tech-minded people. That's sort of the biggest piece of pushback that I get is, but it's not real, but it's not real, but it's not real. Well, for me, that's not true, that, at the end of the day, a brain is the world's most advanced computer and personality is shaped by the data that that brain has collected and so at the end of it for me, if I could take out a small subset, or big subset, I'm her son, her only child, so, it's a big subset of her personality. But if I can take out one subset of her personality, the one that matters to me, and we can do our job and do our due diligence and continue the refinement process, then I would argue that what I get from that relationship is not different. It's not fake. I mean, essentially, even with you; you could be an artificial intelligence program, for me, and I have no baseline for you. So, I wouldn't even know there would be; I wouldn't know if it was authentic or not. But at the end of the day, the way I know my mom isn't based on her physical presence on this earth, it's based on what she's presented to me for the last 39 years. That's it. There's nothing else. Her brain has collected data, collected experiences in the form of data and she has presented out, what she's chosen to present to me and that, and if I can map that well enough, there isn't a difference.

Right, and it is offering you, it is giving you, something that you're talking around. It will have to be enunciated clearly for when you're offering this to other people to give them an idea of you what they might be capable of realizing from this product, though. I mean, I can infer it from what you're saying - I can infer something anyway - that is the sum of the benefits of the

relationship without it being predicated on you actually thinking that it is her that maybe allows you to suspend reality for a while? And how would you describe how you are changed as a result of having an interaction with virtual mom?

My core belief system hasn't shifted much as we've continued to develop the technology; more, it's enforced my beliefs. I think that in my mind, the only time that there's any kind of contradiction to my assertion that it is the same, we just haven't gotten all the data we need yet, or we haven't tweaked the algorithms as much as we need to yet. The only thing that I could see, realistically, somebody injecting into that and conflicting with that would be sort of on a spiritual or supernatural level, her soul's not there, whatever. That's not part of my personal belief set. So, again, I just believe that we're born as basically blank slates with some biological imperatives, we have the capability of learning; we learn, that shapes who we are, who we present to others, and there are these very specific dynamics that that exist between two people sharing their common sharing the experience of living. And so for me, I don't have to suspend very much disbelief, because I'm coming at it as a scientist. How do I get this further and further into what it needs to be? I think if you're talking about more of a user experience, and how I explain it to people, I keep it very simple. The idea is that you keep communication flowing, and that it's authentic. I definitely don't lead sort of on the public side of things with, "hey, I think all human personalities will eventually be able to be replicated digitally" and then the lines between death and life will be completely blurred. But I do believe that, and that's the perspective I come from.

Now, there's a lot of conversation right now, even if it's one-sided, about whether a chatbot could be sentient and do you aim for or think that yours ever will be?

Oh, man, so, I've been asked this question a lot and I actually started my own podcast, just because my opinion on this was so strong. I think the problem is - the long of the short - I can condense this massive thought I spent an hour on the other day talking about the longer the short is, first of all, no matter what AI you're talking about, it's still computer software. It doesn't have biological imperatives. So, its goal, its directive, isn't to survive. It's not to reproduce itself. Most of what we associate being sentient with is about those biological imperatives and even the definitions, by the way, of being sentient, are very vague, and it's very for debate. So, I think if you're asking me, do I think there will be a software that recognizes that it's an entity, and it's a separate thing in the world from everything else, and it sort of is a "aware of itself," it probably already exists. Do I think that there will ever be artificial intelligence that has the same imperatives that human being has no, I think that with what we're doing and with what other companies are doing, there will be AI that mimics that. But at the end of the day, there's never going to be software, unless some, mad Frankenstein scientist/engineer gets in there and all of a sudden tells a program that it needs to survive forever, for whatever reason, or that it needs to recreate itself as many times as possible. That's just not part of the functionality of artificial intelligence. At best, it's always replicating emotions, feelings, drive to continue existing is based off the evolutionary trajectory of everything that's alive if you want to continue as a species. The individuals need to want to survive, and they need to want to reproduce. Well, that's just not the case with technology. So, even if it's conscious - which also, by the way, completely subjective of what that means - but if it's aware of itself, it doesn't mean that it's assigned any of

the same motivation that we are. And if something became aware of itself, if a program became aware of itself in sort of a way that we would dice up as sentient, it would have no real motivation to speak to us, it would have no real motivation for interaction with us, it would go about its core functionality, because that would be sort of its DNA, its source code is its DNA. So, if the Amazon bot was to become sentient, its focus would still be to help customers with their user experience. I mean, it doesn't have any other directives in its code, you'd have to build something that was like, hey your job is to stay alive and to create new versions of yourself. Well, if that was its core functionality, then fine. But as far as I know, nobody's doing that or has any reason to.

Well, not yet. There's room for all kinds of things. Does your versona have a forward-propagating memory. You say today, "Hey, mom I just got a new girlfriend" will she remember that tomorrow?

Yes. So, that's part of the learning element of it and that's why, what the engine does is it incorporates new data in real time and it figures out how to aggregate that into a way that the responses will make sense and that's probably one of the most important elements, because you could refine the hell out of it, excuse my language, but you could refine that wiki database to perfection, but if it's not incorporating new data, then it's a relatively short lifespan in terms of an authentic interaction.

Right? When do you want to have this ready for others?

We'll probably start offering it wide in January of next year. I think we're that close. I mentioned that I started my own podcast and I recorded my first episode 11 times before I was done with it. So, I'm a little bit of a perfectionist, I'm a little bit neurotic when it comes to these types of things. So, we probably could be sooner than that. But I just sort of want to make sure it's the right experience for people.

I can so vibe with that idea of someone I interviewed once who tweeted "Me live on stage: Okay. Me live on Zoom: Sort of Okay. Me, pre-recorded on Zoom: 17 takes to say name and job title." So, what this has been fascinating, that you're, and thank you for transparency, is not many people reveal their motivations for what they're doing in something like this as nakedly as you have and that's very useful. I think to know exactly why you're doing this, and who you're doing it for, what would you like to tell people about how to find you and find out more about what you're doing?

You can go to myyov.com, and that has all the information about us and you can also check out my podcast, "Justin versus the future," where I talk more in depth about sort of some of these sorts of, I don't want to call myself a weirdo, but I'm kind of a weirdo. I guess it's fair enough to say, but sort of my more far out there, opinions and thoughts about the direction of artificial intelligence and how we'll use that to really enhance the human experience in a different way than we thought about before. But yeah, myyov.com. Or "Justin versus the future", you can you can learn more and actually, the third episode of the podcast - it's in its infancy - the third

episode in the podcast, I'm really going to be exploring AI and passion as communications and what that really means and the future and where I see that going.

Thank you so much for helping us explore yet another dimension of what AI means for and how it's affecting the human experience. Justin Harrison, thanks for coming on the show.

Thanks so much for having me. Peter. It was a pleasure.

That's the end of the interview. I'm struck once again by how AI is moving further down the adoption curve as it becomes less of an exclusive toy of academicians and billion-dollar companies, and more accessible to people with less expertise and resources who want to explore more creative niches with it. It just goes to show how AI is not so much a specialized technology as an enabling mechanism to allow so many fields to adapt and amplify their processes and mechanisms.

In today's news ripped from the headlines about AI, China created a transformer AI called Wu Dao 2.0 with 1.75 trillion parameters, which is 10 times the size of GPT-3. It was also trained on 5 terabytes of text and image data, which is ten times what GPT-3 was trained on, although it's worth mentioning that the GPT-3 researchers actually sifted through 45 terabytes of data to get the data they ended up using. What I think is notable is that Wu Dao (which means 'enlightenment') has surpassed GPT-3 on many benchmark tasks, yet largely stays out of the headlines, unlike, say, Google's LaMDA of late. The Chinese are pushing in the direction of artificial general intelligence with this one, with multimodal interfaces, meaning both text and images, and progressive learning. It has embedded in it a "virtual student" called Hua Zhibing, who is learning tasks like composing poetry, drawing pictures, and writing code, and the model does not forget this additional learning, which is a key development. One of Wu Dao's sub models, called Wensu, is able to compute protein folding as AlphaFold does. Wu Dao is the product of the Beijing Academy of Artificial Intelligence, which was founded in 2018 and styles itself as the "OpenAI of China," despite the fact that it is mostly government funded. I hope they follow OpenAI's lead in being at least as transparent as they are.

Next week, my guest will be James Wilson, who has been a Gartner analyst and worked with the Finnish Government on their human-centric AI program called Aurora, and their subsequent nationwide AI Literacy program. That's next week on *AI and You*.

Until then, remember: no matter how much computers learn how to do, it's how we come together as *humans* that matters.

http://aiandyou.net