



## CG GARAGE PODCAST #265 | KOKI NAGANO PRINCIPAL SCIENTIST, PINSCREEN

**Want to create a perfect digital human? USC alum and Pinscreen researcher Koki Nagano tells Chris how deepfakes and CGI could help us cross the uncanny valley.**

Joining Chris for this week's podcast is Koki Nagano, who has become an expert in virtual doppelgangers via his research on digital humans at USC and his work at Pinscreen, a company that uses AI to supercharge the creation of 3D personal avatars from a single photo. Koki gets under the skin (literally) of what it takes to create a detailed digital human and talks about how he's approached the uncanny valley from both sides.

### Contents

[Creating Digital Emily](#)

[Hollywood VFX vs. homemade deepfakes](#)

[Defining deepfakes](#)

[The benefits of digital humans](#)

### Useful links

[Koki's website >](#)

[Koki's TedX talk >](#)

[Pinscreen >](#)

[Pinscreen: paGAN real-time avatars >](#)

## CHAOSGROUP

- Chris Nichols I guess the first time we really met and talked was when we were working on the Wikihuman project, right?
- Koki Nagano Yeah, I think so.
- Chris Nichols Yeah.
- Koki Nagano I think we met for the first time. There was also Vlado, you and Paul, and maybe a few others at the ICT conference room.
- Chris Nichols Right, right, right.
- Koki Nagano Yeah.
- Chris Nichols Yeah, when we were talking about what we were going to do for the project and how we were going to do it.
- Koki Nagano Yeah.
- Chris Nichols It was a lot of fun, really interesting stuff that we did back then, and we have a lot to talk about, digital human work.
- Koki Nagano Yes.
- Chris Nichols But let's start with a little bit of origin story. How did you end up at ICT working on digital humans? What interests of yours led to that decision? You mentioned you were in architecture school, right?
- Koki Nagano Yeah, so my undergrad was environmental design.
- Chris Nichols Yeah, yeah, yeah.
- Koki Nagano So, it was no computer, just drawing on the paper and creating some tiny miniatures for the building. So I like all that art, design kind of things, but also at the same time, when I was very young, I had grown up watching *Godzilla*, some special effects movies. I remember watching the *Lord of the Rings*. I bought the DVD and it had the making of things. And I was growing up in the countryside of Japan, so there was not that much information. So I thought the only way to get to that industry was to somehow learn how to create art, major in those things. Somehow I ended up in some architectural measures. But when I was in undergrad I took some classes, and I made some media art, and I submitted to SIGGRAPH when I was in my third year in college.
- Chris Nichols What did you submit?

## CHAOSGROUP

- Koki Nagano It was a system which allows you to see two different images on the same screen. It uses polarizers to show, like, in a 3D movie, you can see a right image. So then you can have two different polarizers, so different people can see different content. So we extended that kind of thing to different applications.
- Chris Nichols Oh, interesting.
- Koki Nagano So that was a like small work, but that was my very first SIGGRAPH, 2010. I presented at SIGGRAPH poster. The SIGGRAPH 2010 was in L.A. and now I also wanted to study abroad for my graduate study. So my advisor contacted Paul, because it's close to ICT, so we got a chance to visit ICT. So, that's the first time I visited ICT. I think it was maybe only a few months since ICT moved to the current location.
- Chris Nichols The current location in Playa, right?
- Koki Nagano Yeah, yeah. Right.
- Chris Nichols Because it was in Marina del Rey before.
- Koki Nagano Yeah, yeah, yeah.
- Chris Nichols Okay. You know that building is also the building that they feature in Silicon Valley? That's the Hooli building. That's actually the ICT building.
- Koki Nagano Yeah, yeah. And then when I was getting my PhD, I saw a few people there shooting the movie. It was a blast. And also, that year, Avatar came out, and then Paul got an Academy Award. And I was really, really, really interested in how the technology to make those digital humans, how to create-
- Chris Nichols So you were interested in digital human work based on Paul's work?
- Koki Nagano Yes. Like *Avatar*, *The Lord of the Rings*, *The Matrix*.
- Chris Nichols Yeah. Cool. So that got you involved. So Paul basically got you at ICT right?
- Koki Nagano Right, yeah.
- Chris Nichols And so that became your graduate program, it was going to be at ICT?
- Koki Nagano Yeah.

## CHAOSGROUP

Chris Nichols      Okay. Because when I met you, you were getting your Ph.D. there. It was part of your research, right?

Koki Nagano      Yeah. I finished my bachelor's in Tokyo, and then I flew here in 2012 to start my Ph.D. at USC.

Chris Nichols      That's awesome. So what, I think, we should let people know is one of the things that you were doing your work on digital humans, I was working with Paul and Jay, and who else was there?

Koki Nagano      Xueming Yu.

Chris Nichols      Yeah.

Koki Nagano      Yeah. And Oleg was there.

Chris Nichols      Oleg was there. Yeah. And we were looking at creating this open data set that people could use for research, and we had the Digital Emily data, which was kind of raw, but we took that data, and we made it to something that was a little bit more usable, which ended up becoming the open-source data. But then you ended up using it for your own research, right?

Koki Nagano      Right, right.

Chris Nichols      So, can you explain to people what that research was that you were working on?

Koki Nagano      So there was a previous work by Paul Graham and different Ph.D. students on how to digitize the human skin at micro scales, like a hundredth of a millimeter.

Chris Nichols      Sub-pore level.

Koki Nagano      Yeah. Even beyond the Light Stage capability at that time. So, that worked. Can digitize the face at extremely super high resolution. I was looking into what happens to those structures when we, for example, smile. The skin stretches and compresses. And how the skin structures changes, and how it influences the skin reflectance, BRDF and all those things.

Chris Nichols      So just so people understand. You and I understand this, but some people... so basically, your pores have a certain structure, but if you look at your cheek, for example, it's broken up by all the pores and little micro-wrinkles put between those pores. But if you blow up your cheeks, those get stretched out and there's less break-up of the skin. It gets shinier.

## CHAOSGROUP

- Koki Nagano            Like a balloon.
- Chris Nichols            Yeah. With your research, because you guys had scanned that data, but you were like, "Well, what happens to that data when your face changes?" Right?
- Koki Nagano            Right, right. We were working on the project maybe for a year or so. Basically, extending the Light Stage capability to scan that at that resolution and with animation. So we submitted that work to SIGGRAPH 2015. That was accepted at SIGGRAPH's technical papers.
- Koki Nagano            So another thing was that I wanted to create a digital human, not just doing research, I wanted to see a lifelike digital human. Paul has many films that feature his papers, so I also wanted to do something similar. Then there was this opportunity he just gave me, and then we just have a new technique that can render the human skin at super high res. That was very nice.
- Chris Nichols            It was cool. People should check it out. I'm sure we'll put a link to the paper. But you guys can check out the videos that we were working on. It was really great to be able to support you and to help you get all the renders that you guys needed for that stuff because it ended up looking really great. I think it's an extremely interesting thing to think about. What you guys were doing back then at ICT was just the most high-resolution detail you could ever imagine on humans. Very, very, very, very, very high resolution. The interesting thing was if I have data that's that high res, is that going to make my digital human look better, right?
- Koki Nagano            Right, right, yeah. To basically create the digital human, I spent a lot of time rendering the Digital Emily thing, but I'm not an artist so I didn't have any previous experience in how to make these. But the idea was if we had an accurate database, then maybe someone like me can even set up a renderer, and can set the shader parameters, and all the settings right, and then maybe we can recreate the digital human without going through different or weird tricks to make it look right.
- Chris Nichols            That I 100% agree with. I think one of the things that's really cool about your stuff, if you're really trying to recreate a digital human, you get as much information as you know is a constant. If you reduce your variables, then you're in good shape, right? You guys calibrate all your colors, so you know that that color is correct. You have a scan of a person down to microns, so that's correct data. So all you really need to do is reverse-engineer what light passing through skin looks like and you're done.

## CHAOSGROUP

- Chris Nichols I think you definitely did it. That was the thing that was frustrating for me, as someone who is half-artist, half-scientist, I think that some of the artists would ignore the constant data that they would get at ICT and say, "Well, no. The color needs to be that, because it's been calibrated." And it got frustrating. But I got guidance from people like you and Graham like, "This is the color. And the way that the skin shader needs to be is like this." And therefore, don't think about it.
- Koki Nagano Yeah, just don't mess with the region that's already calibrated.
- Chris Nichols That's calibrated, et cetera. Yeah. And it works. Honestly, it works. There's a guy, an artist, that I work with named Ian Spriggs. I don't know if you know his work. He's not a technical guy at all, but one of the things I told him is, "As long as you paint your colors correctly, then everything else, just plug it into default, then it works." So he gets really good results because he uses defaults for everything.
- Koki Nagano Yeah, yeah, yeah. Default is the best.

### Creating Digital Emily

- Chris Nichols Defaults are the best. So you were, at that time, looking at extremely high res. The texture maps for Emily were like 16k, right?
- Koki Nagano 16k, yes.
- Chris Nichols Those are the displacement maps. That's crazy resolution.
- Koki Nagano Yeah, so we had to start the process by capturing the data. To capture details at that resolution we had to capture the skin samples at like five regions. Like forehead, temple, nose, cheek, and chin.
- Chris Nichols Right. Explain why.
- Koki Nagano Different regions of the skin, within the same person, is more different than the different people's skin patch of the same region. So my forehead and your forehead look more similar. Like for example, it moves up and down, anatomically. So it develops more vertical and horizontal-
- Chris Nichols Wrinkles, yeah.

## CHAOSGROUP

- Koki Nagano Yeah, wrinkles. But my cheek, it can move anywhere. So it's more isotropic. On the other hand, the nose doesn't move that much, so there's not that many wrinkles and-
- Chris Nichols Yeah, but you can't capture all of the face at that resolution.
- Koki Nagano Yeah, so if we can't, then we are basically done.
- Chris Nichols Right. But you can't, so you capture patches, right?
- Koki Nagano Yeah, so we capture the patches. Using these patches, we can propagate all these, basically the details of the entire face. And then we can create the 16k by 16k resolution details.
- Chris Nichols Yeah, which is pretty amazing. We've used that on the Digital Emily data, as well, and it looks incredible. And the Digital Mike one, as well. That's really kind of amazing.
- Chris Nichols Okay. So you created that microstructure data, and you used it to look at skin stretching and pulling. So how did you do that?
- Koki Nagano I think we tried all sorts of approaches we can possibly think of. Earlier in the project, we started looking at the physics simulation approach, because this is very much like the surface, so there's like a cloth simulation type of thing.
- Chris Nichols Oh, cloth simulation.
- Koki Nagano Yeah. Basically, it's just like wrinkle simulation. So there are many kinds of simulation technique; mass-spring cloth simulation, as rigid as possible. Basically, I tried all kinds of different things, but the skin is actually not just a single layer, not the surface. There's a volume, right? There's an underlying volume that needs to maintain that volume, and the surface needs to basically move in a way that maintains the volume. It has to also somehow consider that volume.
- Koki Nagano Then we started looking into a more volumetric approach, like a finite element method. One of the authors of the paper was Professor Jernej Barbic, at USC. He develops the FEM library. I also tried volumetric approaches, but the problem is that this resolution is too crazy, 16k by 16k. So you can't possibly get all the polygons and simulate regular physical simulation.
- Koki Nagano So, we went back to the capture, and then eventually, on the effect-wise, what's happening is when the surface gets stretched, basically it flattens out and when it compresses, it bunches up. Basically, if you just look at

## CHAOSGROUP

the surface from afar, it's just a statistical effect. So when it's stretched out, the surface normal looks vertical. So, if you put the point light on-

- Chris Nichols: So if you're using it as a normal map, basically you just reduced the normal maps, right?
- Koki Nagano: Yeah. Exactly. So we ended up with blurring and sharpening the displacement map to simulate the skin stretching and compressing.
- Chris Nichols: And compressing. Interesting.
- Koki Nagano: Yeah. And that can run on the GPU, as the blur and sharpen was extremely fast. So in 2015, 2014, we were able to render this in real-time.
- Chris Nichols: Wow.
- Koki Nagano: Yeah. 16k. Some of the videos in the paper were actually rendered by my own OpenGL GLSL shader in real-time. Many of those are rendered by V-Ray.
- Chris Nichols: Wow.
- Koki Nagano: Much better.
- Chris Nichols: But still, it's pretty amazing to be able to think about how you're able to do that because that's pretty incredible. The results were very cool. We'll definitely put a video link on there so people can check it out. And if you guys are interested in the paper, they should read that, since it's an open paper. But that was your Ph.D. paper, right?
- Koki Nagano: Yeah. That was one of my projects.
- Chris Nichols: That's cool. So when did you get your Ph.D.?
- Koki Nagano: 2017.
- Chris Nichols: 2017. Okay. That's great. That's really great. Since then, you were at ICT for a little while after that, but then you moved onto Pinscreen. So that is Hao's group right?
- Koki Nagano: Mm-hmm (affirmative). Yes.
- Chris Nichols: So you've known Hao, because Hao obviously was at ICT, as well. Tell us what Pinscreen is.
- Koki Nagano: Pinscreen is a startup started by Hao. Around maybe 2015, so like four or five years. We're based in Santa Monica. Mostly, my current work focuses



## CHAOSGROUP

on developing some cutting-edge technology for VFX, computer graphics and visual machine learning. Basically building a realistic digital avatar, something you can use for a virtual assistant, or it can be used for entertainment like TV or movie productions.

Chris Nichols Okay. So it's an avatar of yourself.

Koki Nagano Yeah.

Chris Nichols Okay. So you're moving that avatar through what?

Koki Nagano So it can be performance-capture based. I worked on some high fidelity, but real time, face tracking that can run on a desktop, where you can drive the avatars. We also have a lightweight version that can run on like iPhone 7 or so without depth sensors.

Chris Nichols Without depth sensors?

Koki Nagano Yeah.

Chris Nichols So all image based?

Koki Nagano All image based.

Chris Nichols Right. So you're capturing people's animations and faces, right?

Koki Nagano Yeah. Animations and faces. Also, to build an avatar you need whole body and hair. Basically, just looking at how to build the complete avatar.

Chris Nichols Does it make the avatar for you?

Koki Nagano Yeah. We have an app called Pinscreen. You can download it from the App Store. So how it works is, if you upload an image to the backend server, all the computation is done on the cloud. Then the computers digitize the 3D face and get the 3D hair, and also run some deep-learning stuff to get the textures and everything. Then you can get the assets back to the front end, and you can view the avatar and interact on the phone.

Chris Nichols So how is this different from the, what's the Apple one that they do?

Koki Nagano It's very related to Animoji. The difference is, Animoji you can create from a preset. So I can only customize very global, skin color, and-

Chris Nichols Right. Kind of like making your avatar on the Wii or something.

Koki Nagano Yeah. It's pretty. It's cute.

## CHAOSGROUP

- Chris Nichols: But yours is based on your actual image.
- Koki Nagano: Yes. Based on the actual image. Photoreal and personalized. That's much harder. So we can go to realistic, to stylized, but we cannot go the other way around.
- Chris Nichols: Right. You can't go stylized to realistic. So it is still a stylized version of yourself, right?
- Koki Nagano: Yeah. It looks like a game character, but eventually the goal is to make it something that looks indistinguishable from the real you.
- Chris Nichols: So you're cloning yourself digitally?
- Koki Nagano: Right, yeah.
- Chris Nichols: Okay. Compared to what we were talking about when we were going into the micro details of skin texture, this is a phone picture. This is a selfie that turns into an avatar. That's not the same detail that you're getting from before, right?
- Koki Nagano: Yeah.
- Chris Nichols: That's a challenge.
- Koki Nagano: Also, workflow-wise, it's very opposite. Basically, the traditional VFX, graphics pipeline is very bottom-up. So you have data, like a 3D model, albedo and all the like renderers, like lighting, everything needs to be set up to produce the rendering. And then, even with the best data, it's still very hard to create a very convincing digital character. It's a lot of work. Needs a lot of expertise.
- Koki Nagano: But on the other hand, some more recent approaches, like deep-learning and machine learning, were very top-down. It's very data-driven. How it works is often, given some images or videos, you throw them into some machine learning framework and then it learns how to create human facial expressions and textures and all those things without knowing BRDF and all the micro things.
- Chris Nichols: Right. It does it for you. Okay, so basically this is completely different than what you were doing before, because before you were doing everything by hand. You're gathering as many of the constants or variables that you needed from the scan data, but then you have to know the science of everything going back together, right?
- Koki Nagano: Right.

- Chris Nichols      The thing with machine learning is that it's going to guess all that information because that's the thing about machine learning, right? It's guessing.
- Koki Nagano        Right, right. Yeah.
- Chris Nichols        It's giving you the best-educated guess of what you need to do. And so that's very different.
- Koki Nagano        Yeah. I've been working on the exact same problem, but from different approaches.
- Chris Nichols        Do you think that the machine learning approach is going to replace the very laborious systems that we were doing before?

## Hollywood VFX vs. homemade deepfakes

- Koki Nagano        Yeah. So there was a very interesting comparison with *The Irishman*. So the de-aging was probably the best piece of that like VFX. But someone used deepfake to see how it might look like if the studio used deepfake to create De Niro. One is done by old experts, months of work, probably cost millions of dollars. But the other one looks convincing, but can be done by a guy-
- Chris Nichols        With his laptop.
- Koki Nagano        Yeah. Probably a desktop.
- Chris Nichols        So, I saw that, and I think we actually might have mentioned it on the podcast with Wael, as well. But I agree. I think, honestly speaking, I think that what was done by the guys at ILM, and huge respect for their work, but traditionally, the effects companies tend to over-engineer and overthink problems because they think that the more complex it is, the more likely they are to be rewarded for that work when that's not really the goal. And I think that machine learning feels like you're cheating.
- Koki Nagano        Right. Yeah.
- Chris Nichols        You know? But, at the same time, I think that was a very interesting idea. Honestly, I was thinking very closely about the work you were doing when I was seeing that because what the ILM approach did was attack the

## CHAOSGROUP

details. It went in there and basically only looked at the wrinkles and how to take the wrinkles out and smooth the skin appropriately, and reconstructed the face that way. But what the deep-learning thing did is that it actually spent more time rebuilding the face.

- Koki Nagano Right. It's basically de-creasing-
- Chris Nichols Right. The eyes weren't drooping.
- Koki Nagano Mm-hmm (affirmative).
- Chris Nichols You know? He looked like he was 55 the whole time, as opposed to 75. When he needed to look like he was 30 all the way to 75. So I think the deep-learning technique looked at the macro much more than the micro, you know?
- Koki Nagano Right, right. Yeah. Basically, it will reconstruct the entire face from scratch.
- Chris Nichols Right.
- Koki Nagano So, yeah. So if, I guess, the bone structure changes when a person gets older, or-
- Chris Nichols I don't know if the bone structure, I think your muscles-
- Koki Nagano Maybe muscle, yeah. But it'll change the overall facial structure-
- Chris Nichols The facial shape. Yeah.
- Koki Nagano Yeah. Which is harder for, maybe like the traditional approach, or like special makeup, you can probably remove the wrinkles and do something.
- Chris Nichols Right.
- Koki Nagano It's more difficult to change the actual face shape.
- Chris Nichols Right.
- Koki Nagano Yeah. But this approach can basically overwrite the entire thing.
- Chris Nichols So overall, did you think that the deepfake one was more convincing? I mean, I know there are only a few examples.
- Koki Nagano I think it's very comparable. Considering how much work was needed.
- Chris Nichols Right.

## CHAOSGROUP

Koki Nagano            Yeah.

Chris Nichols            Okay. So one of the things that was obvious to me is that the overall shape, like we just mentioned, was much more convincing, but there was something about the resolution that wasn't quite right. It was a little too smooth. It didn't blend in correctly. You didn't have all the details.

Koki Nagano            Right.

Chris Nichols            Is that currently a limitation of the machine-learning process, or what's going on there?

Koki Nagano            Yeah, so right now I think probably the current film, at least like 1080p or even higher, like 4k.

Chris Nichols            Yeah.

Koki Nagano            Yeah, the resolution's still an issue. So, there's a lot of frames in the film where the character comes really close to the camera. For that, I think, the technology will have some difficulties. But think about how this technology was last year. It was much lower resolution. Much worse quality, and now-

Chris Nichols            Right. So you think we will eventually get there.

Koki Nagano            I think so.

Chris Nichols            Okay. All right.

Koki Nagano            But I don't think it will replace the traditional approaches. It's just a different way. None of them are perfect, but we can-

Chris Nichols            We can use both.

Koki Nagano            Yeah, combine things.

Chris Nichols            Okay. Interesting. Right. So, you have been getting into some stuff with deepfakes recently, right?

Koki Nagano            Yeah.

Chris Nichols            So that's part of Pinscreen, as well?

Koki Nagano            Yeah.

Chris Nichols            You mentioned the avatar, right?

## CHAOSGROUP

Koki Nagano Mm-hmm (affirmative).

Chris Nichols That's not really a deepfake.

Koki Nagano Right.

Chris Nichols But there's stuff that you're using that's similar techniques to create deepfakes? Explain what you guys are doing in that area.

Koki Nagano So, there was one paper I did, maybe like one or two years ago. Basically, you can create arbitrary facial expressions of anyone from a single picture. So if you take a selfie, then you can use GAN, generative adversarial network. Imagine how, if I smile, if I open the mouth, it can change size, like "Peace." If I look left or right, it can also fill in the textures.

Chris Nichols So you take a single picture and then it figures out pretty much a 3D model of your face.

Koki Nagano Yeah, it's actually a hybrid approach. So we have an underlying 3D model but then, through this 3D model, basically the input to the neural net is also the 3D model. So we can use the traditional control, like a blend shape smile slider, and then we can make the person smile.

Chris Nichols So if you take a still of me not smiling you can use your software to make me smile?

Koki Nagano Yeah. So using this technology we can do a virtual FACS session from a single picture.

Chris Nichols Really?

Koki Nagano Yeah.

Chris Nichols So you can get all the FACS shapes?

Koki Nagano Yeah. From that angle, maybe you can do a little bit of the views.

Chris Nichols Okay. Can you change lighting?

Koki Nagano We can change the lighting, too.

Chris Nichols So you pretty much captured my face from a single image.

## CHAOSGROUP

Koki Nagano            Yeah. So if you have different textures for the smile, you can get the wrinkles from that picture. Then all you need to do is just blend them properly and then you can drive it, even in real-time.

Chris Nichols            Okay. Could you puppeteer it?

Koki Nagano            Yeah, we can puppeteer it, after that.

Chris Nichols            So you take a picture of me, just the way I am, then you could animate me from that single picture by doing whatever you're doing with your faces.

Koki Nagano            Yeah. Right. On the phone.

Chris Nichols            On a phone?

Koki Nagano            Yeah.

Chris Nichols            Okay. And what is this called?

Koki Nagano            It's called Pinscreen.

Chris Nichols            That's the Pinscreen one.

Koki Nagano            Yeah.

Chris Nichols            Yeah, that's crazy.

Koki Nagano            Yeah. So that was developed for building a 3D avatar, but basically this capability allows anyone to manipulate someone else's face in real time. So we got a lot of requests from all the press, like L.A. Times or like CBS and-

Chris Nichols            CNN.

Koki Nagano            Yeah. Interested in basically showcasing this new technology. They call it deepfake. It's a broader definition of deepfake. It's not the strict definition of-

## Defining deepfakes

- Chris Nichols      Is there a strict definition of deepfake? Because I think people are being very liberal with the term deepfake.
- Koki Nagano      Yeah, the very strict definition is a technology to create video of someone saying or doing something which they never said or did, using deep learning technology. But when some really technical person said deepfake, in the most technical meaning, it means using some kind of encoder and decoder to create fake images. There's also some kind of deep neural net.
- Chris Nichols      Deep neural net. Right.
- Koki Nagano      Yeah. But there are all kinds of different ways to create a fake video of someone. Some of them might involve different kinds of deep neural nets, or some of them might not even involve deep learning. It can be different kinds of machine learning technology or it can be traditional VFX. Sometimes people call it deepfake. Just definitions can sometimes get loose.
- Chris Nichols      Okay. All right. So this is interesting. But you're basically using deep learning technology. So what happened to technology that suddenly deep learning became a thing? Now it's in everything, right?
- Koki Nagano      Right.
- Chris Nichols      Is it all about the GPU and the RAM?
- Koki Nagano      Yeah, I think the GPU is certainly one of the things that made it happen. GPU, but also the data.
- Chris Nichols      The data? What are you using to learn your technique?
- Koki Nagano      For example, if you do just a pure deepfake you need maybe like tens of thousands of internet photos. Say you want to create a jpeg of Leonardo DiCaprio, maybe you can download some images, and then-
- Chris Nichols      Yeah. Just scrape the internet.
- Koki Nagano      Yeah. And then you can build a-
- Chris Nichols      A deepfake model.



## CHAOSGROUP

Koki Nagano Yeah, a deepfake model.

Chris Nichols Right. That's the data you guys used? You guys just scrape the internet, or?

Koki Nagano So, most of the time that's the only data that's-

Chris Nichols Available.

Koki Nagano Available.

Chris Nichols Okay.

Koki Nagano The data was one thing, so there is much more data on the internet. Like phones are getting better, I think, many things. GPU also gets better, so allow you to train bigger, more complex-

Chris Nichols Oh, right. So the quality of the pictures on the internet are getting better.

Koki Nagano Yeah.

Chris Nichols And the hardware is getting better at interpreting it.

Koki Nagano Yeah. So also, this algorithm is very data-driven, so it is a lot of images, or it can be audio, if you want to synthesize audio.

Chris Nichols Well, you can synthesize anything.

Koki Nagano Yeah.

Chris Nichols Yeah, you just have to learn. Whatever data you want it to learn, it'll learn that, right?

Koki Nagano Right, yeah.

Chris Nichols Okay. So you create a deep neural network that allows you to do that and now you say, "Okay, you've learned A, B, A, B, A, B. Now I give you a new A, and you tell me what B is."

Koki Nagano Mm-hmm (affirmative).

Chris Nichols Right? That's really the basic basic definition of that, right?

Koki Nagano Yeah, but deep learning itself has a broader definition. So it can be like scene recognition for self-driving cars. You have a video feed and the car needs to understand there's a pedestrian, there's a car coming, scene

## CHAOSGROUP

recognition, whatever you want to, like segments or human segmentation. Also, related to VFX, how to composite the images. Or like the text understanding. For example, given an image, you can describe the scene using text.

- Chris Nichols      Interesting.
- Koki Nagano      Yeah. Or facial recognition is also an application of deep learning.
- Chris Nichols      Right.
- Koki Nagano      So it's basically just a general machine learning technology, which uses a special kind of algorithm called deep neural networks.
- Chris Nichols      Right. Yeah, okay. So that's a big thing right now, and it's a hot topic. Obviously, the press wanted to talk to you guys about it based on the idea that you could puppeteer anyone. Not just yourself, but you could puppeteer other people. That became a concern, I'm sure, of the public. So what are your thoughts on that?
- Koki Nagano      So the technology, what it can do, is the same as what the VFX has been doing for a few decades.
- Chris Nichols      Right, but it can do it way easier.
- Koki Nagano      Yeah, so the difference is now the technology is accessible to anyone.
- Chris Nichols      Right.
- Koki Nagano      So, that's the difference. Maybe a similar thing happened before with Photoshop, maybe before it was difficult to create images. But now, anyone can do that. Now, this is coming to video.
- Chris Nichols      Right.
- Koki Nagano      So I think no one can stop the progress of the technology. So eventually it'll happen and I think we have to live with it. So yeah, but we have to do something at the same time, right? So we need to raise the awareness of this technology so people know that you can't just blindly trust the videos anymore. That's part of our work.
- Chris Nichols      So you make it fake, and then you show people it's fake.
- Koki Nagano      Yeah, we have to.
- Chris Nichols      Yeah, no. That's an excellent point and I've brought this up before. I think you're absolutely right. I think that now people don't blindly trust photos

## CHAOSGROUP

because everyone knows it's probably Photoshopped. In fact, Photoshopping photos is, in itself, fun and not always vindictive, right? The whole point is to make a joke out of it, right? So you guys are showing us these videos are possible and the more people know about it, the more they can recognize that it's real or not real, or at least question whether it's real or not real.

- Koki Nagano Right, right. Yeah, so last month we were invited to the World Economic Forum in Davos to demonstrate this deepfake in real time. So now this technology is being discussed on the same table as climate change and other things while the world leaders and industry leaders-
- Chris Nichols Do you have concerns about the technology?
- Koki Nagano So right now, people are careful about not putting the source code online, because then people can use this information and make other bad things.
- Chris Nichols But if you guys did it, someone else could discover it too, right?
- Koki Nagano Well, so we never release our source code. That's too dangerous. And also, at the same time, we developed technology to detect fakes, taking advantage of how the algorithm works.
- Chris Nichols Okay. But you're playing both sides? Because if you're creating technology to detect the deepfakes, and you are creating the deepfakes, can't you then make the deepfake technology better so it can't be detected?
- Koki Nagano Yes, exactly.
- Chris Nichols You're creating the virus and the antivirus at the same time.
- Koki Nagano Right. But I think some researchers discussed this. Some rules. Like when a researcher worked on this thing. So, I think one of them was saying if the researcher worked on this creation tech, they're also responsible for developing the technology to detect it.
- Chris Nichols Right, okay.
- Koki Nagano But I think eventually that detection won't be a single cure, because I think eventually it will be impossible to detect it. I think eventually it'll be very good to-
- Chris Nichols When do you think that's going to happen?
- Koki Nagano Well, some videos you can't even tell. So educating people is very important. Also, the detection tool is just a measure of how authentic this data is. Also, all the platforms, like Facebook and Twitter, they have to be

## CHAOSGROUP

careful not to spread fake news. So those platforms are also responsible for taking care of what's authentic or not. There's a study; fake news spreads faster than real news. Also, the regulations by law makers-

- Chris Nichols      So I'm going to get into that. So a lot of the deepfakes that you see there, like the ones done by Ctrl Shift Face, that's using open source software mostly, right?
- Koki Nagano      Right.
- Chris Nichols      So, that's out there. So you were just talking about source code. It's out there. It's possible, and he's probably a good artist, and he's very good at using that open source software, but it's still open source software. So you can't remove that from the internet.
- Koki Nagano      Right.
- Chris Nichols      Right? You also can't regulate that.
- Koki Nagano      Right. So there are good uses and bad uses, like Photoshop can be for good or bad. It's the same thing. But some report says that currently, 96% of deepfakes are porn videos, which is obviously harmful. I think there needs to be some sort of regulation for misuse of the technology.
- Chris Nichols      Okay. Right. I brought this up with Wael, as well. Right? It was interesting. So 96%, you said, of deepfake technology is being used for porn videos, right? That doesn't necessarily mean that the deepfake is to blame; it's the person, it's the act of what they're doing with the deepfake.
- Koki Nagano      Exactly.
- Chris Nichols      So if it's being used for malicious things, then that should be regulated, right?
- Koki Nagano      Right, right. Yeah.

## The benefits of digital humans

- Chris Nichols      So, because there are other uses of that technology that are beneficial, what are those?
- Koki Nagano      Well, entertainment is obviously, because you can create a realistic face. Entertainment, VFX, also like a virtual assistant.
- Chris Nichols      Virtual assistants. I know this is Mike Seymour's big thing, right?
- Koki Nagano      Yeah, yeah, yeah. Basically if you just put this together with some speech synthesis, you can have someone talking with realistic face. So yeah, the technology itself has just-
- Chris Nichols      Lots of uses.
- Koki Nagano      Yeah, potentials, but there's also, because it's so powerful, it can be used for something bad too.
- Chris Nichols      Right. Yes, there's a lot of bad things that happened on the internet. That doesn't mean we have to take away the internet.
- Koki Nagano      Right, right, yeah. Exactly.
- Chris Nichols      All right. So, I want to talk a little bit about when we were looking at the Wikihuman data, and when I was working at TD way back in the day, and we were creating these digital humans with super-high resolution scans, and every time we got better and better data from the light stage, from ICT. The data set would get better and better and better. But it was still a lot of work to get there. And the big challenge we were always dealing with was the uncanny valley problem.
- Chris Nichols      So the uncanny valley problem is like, digital humans, all right. Are we going to be able to get out of the uncanny valley? That was seen to be the number one thing. If you're looking at this deepfake idea, do you believe that you would be able to come out of the uncanny valley, or tackle that and create a very realistic digital human without dealing with that data set? And, actually, for VFX. Could you do VFX of a digital human and be out of the uncanny valley completely using that technology?
- Koki Nagano      Right, so this technology is more like, start with the other side of the uncanny valley, right? Because, it starts from the real human. So it's already crossed.

## CHAOSGROUP

- Chris Nichols Right.
- Koki Nagano Yeah, but the traditional way, it has to build up small things, then it has to pass the valley. Takes some time. But now this technology comes from the other side.
- Chris Nichols So you have to keep it so good that it stays on the right side of the valley.
- Koki Nagano Yeah, because the data given is the videos or images, so it learns how to basically mimic those identities of humans or expressions.
- Chris Nichols Do you think, for example, you could use this technology – I'm just going to come up with an idea, right: Let's say you have to create a digital human of some kind and you want it to look like a specific person. I'll just use me as an example. I'm going to be an actor in a film and you want to create a digital version of me for some reason. Can you make a digital version of myself that's pretty close. Doesn't have to be perfect. It can be okay. But then use the deep learning, and a deepfake technique, using my face to manipulate and make the other face look better?
- Koki Nagano Yeah. I think so, yeah.
- Chris Nichols And that means you're not actually rendering that face the traditional way. It's being derived from pixels. I'm not ray tracing, subsurface scattering and putting bump maps. It's all being done on the network, right?
- Koki Nagano Yeah, exactly.
- Chris Nichols So you think that could be possible?
- Koki Nagano Yeah, that can be possible. Yeah.
- Chris Nichols And it would look better?
- Koki Nagano I think so, yeah. So, one problem of this top-down approach is because it's so top-down, it's not easy to control, manipulate, like add a subtle smile or change the lighting.
- Chris Nichols But maybe you could use your technique that you used before, so once you've done it, then you can add a little more smile.
- Koki Nagano Right. Yeah.
- Chris Nichols Using your puppeteering technique that you talked about before, right?

## CHAOSGROUP

- Koki Nagano Right, right. Adding more controls are very useful for VFX and graphics applications. In a traditional approach, you can probably get first, like 70% in maybe like 20% to 30% of your time. But if you build up and get closer to the final goal, it's getting less effective. So if you put in more time, but improvements, the last 20%, 30% is very difficult. But if you combine this new learning technology, it can get you from that 70% to-
- Chris Nichols To 100%.
- Koki Nagano Yeah.
- Chris Nichols Interesting. Okay. That's great because that's a very interesting problem to deal with. So who is your customer? It's not just people buying just on the App Store, and you guys are making money from that, there's got to be other use cases that are going on here, right?
- Koki Nagano Right. So, I talked about virtual assistants. Virtual assistants can be used for many things. Like shopping online. I can talk to virtual assistants and they recommend what to buy. Or, this is also a popular application, doing virtual try-on.
- Chris Nichols Oh, right. You can create an avatar of yourself and then put the clothes on.
- Koki Nagano Yeah, exactly. So without traveling to the store, you can customize your clothing and, yeah, you can do that.
- Chris Nichols Yeah.
- Koki Nagano Yeah. Also, there's a big application in AR, VR. Avatar-based interaction for games.
- Chris Nichols Yeah. That what Magic Leap are trying to create their assistant, right?
- Koki Nagano Yeah. Basically, we need more digital people in-
- Chris Nichols Yeah. I think digital humans have a very big role to play in our future and it's not nefarious. I think that the challenge has always been, because we deal with virtual people, just even in a text box, all the time, right? But if you're able to see a face, and you're able to communicate to them in some other AI method, it may be easier.
- Koki Nagano Right, right, yeah.
- Chris Nichols Yeah.

## CHAOSGROUP

- Koki Nagano      Yeah, we have a facial tracker already on the phone. So yeah, it's a lot more intuitive and you can even customize the virtual assistant's look, hair or appearance.
- Chris Nichols      I think someone was thinking about making a virtual assistant sort of in a house for some elderly woman and made it look like their grandson, so she would talk to her grandson. She knew it wasn't really her grandson, but she was like, "Oh, yeah." It's like asking Siri or Google Assistant or whatever, but then it looks like someone familiar.
- Koki Nagano      Yeah, it can be more personalized.
- Chris Nichols      It can be more personalized. Yeah. That's not a bad thing, I don't think. That's not a bad thing. Well, this is really cool. I think it's very important for people to know this technology works. The fact that you guys are creating it and also making people aware of it is really good. If you are sitting there at Davos and showing people because I saw the videos of you guys on CNN. Did you go to Davos, or was it just Hao?
- Koki Nagano      I wanted to go, but my son is eight months, so...
- Chris Nichols      Yeah, your kid's too small.
- Koki Nagano      Yeah.
- Chris Nichols      But yeah, I saw the videos of you guys on CNN when they were demoing it to people, and showing people how it worked. And it was done in real-time, which was kind of interesting. Just to give people an idea of what it was, it was basically: You sit in front of a screen, you see your own face in front of the screen, just like you were looking at a webcam of some kind, and then you pick a different celebrity that you want to look like and then basically your face is replaced with that celebrity's face. So if you want to look at me, and I say, "I want to look like Will Smith." And then suddenly I look like Will Smith. That's kind of crazy.
- Koki Nagano      Yeah.
- Chris Nichols      Kind of crazy.
- Koki Nagano      Yeah, the subject's chosen by the conference.
- Chris Nichols      The conference?
- Koki Nagano      Yeah.



## CHAOSGROUP

- Chris Nichols Right. Yeah. I think it was really interesting to do that. I definitely believe that the type of technology that you guys are creating is an important part... I can't wait for it to be used in the visual effects world as well. I think that people keep talking about that and that's going to be very interesting. But also, sort of keeping people aware of that. The most important thing, I think, is to show people that this is even possible. Have you seen the Corridor Crew guys' YouTube videos? They made the deepfake of Tom Cruise coming into their office.
- Koki Nagano Maybe. Yeah, I have seen that.
- Chris Nichols Yeah. And it was pretty cool. It was pretty interesting. But what they did there is they like, "Hey, it's fake Tom Cruise coming in," but he looks like Tom Cruise. And it was pretty good, but they basically told you, "This is fake." So then people are aware of this technology. They also did another one. Do you remember the movie, it was a sequel to *Mummy*, *The Scorpion King*? Do you remember *The Scorpion King*?
- Koki Nagano Mm-hmm (affirmative).
- Chris Nichols And it was a really bad CG version of The Rock. And they used deepfakes to fix him.
- Koki Nagano Right, right, right.
- Chris Nichols Did you see that?
- Koki Nagano Yeah, yeah, yeah. Also, there was one from *Justice League*.
- Chris Nichols *Justice League*, yeah.
- Koki Nagano At the beginning, Henry Cavill.
- Chris Nichols That took the mustache off, right?
- Koki Nagano Yeah, right. So yeah, those are the fan usages, but I'm sure it'll be very useful for-
- Chris Nichols Yeah. I think it's interesting technology to think about. I'm not sure how it can be regulated. I don't think the technology needs to be regulated. I think the use of the technology needs to be regulated.
- Koki Nagano Yeah. Right. Yeah. And there's different kinds of deepfake technologies. Probably the most popular type of deepfake is the face-swapping.
- Chris Nichols Yes.

## CHAOSGROUP

Koki Nagano            So, I can put Will Smith's face on my face, but the rest of the body and everything is me. That's called face-swap deepfake. But also some technology can drive the entire video. So like you have a single picture with like an upper body in here, with a background. So then you can animate the entire video or image. So it's called puppetmaster deepfake. Another kind of deepfake, this was famous. So do you know the Obama-

Chris Nichols            Yeah, yeah. The Jordan Peele one.

Koki Nagano            That's also related. So the original paper came from-

Chris Nichols            That one's about mouth manipulation.

Koki Nagano            Yeah. So, that one is only the mouth region. So, the rest is basically the real Obama. Maybe it can look the most natural. So, that's called lip-sync deepfake. But it needs to work with the sound, so it has to work really well with the lip-sync. So there are all kinds of different deepfake techs.

Chris Nichols            Yeah. But I think, okay, let's use that as an example. In the early days of visual effects, lip-syncing wasn't always very good. But now, basically, you could say, "Don't worry about that, and we'll just do all the lip-syncing using that technology." And it's pretty convincing.

Koki Nagano            Yeah, and like rendering the teeth with occlusions.

Chris Nichols            Don't worry about it.

Koki Nagano            Yeah.

Chris Nichols            Yeah. Interesting.

Koki Nagano            Yeah, it can be done.

Chris Nichols            It can be done.

Koki Nagano            Right.

Chris Nichols            Okay.

Koki Nagano            Right. So when we worked on the detection, because of the way they work, some are easier than others. For example, the lip-sync one. The mouth is the only fake part of the video.

Chris Nichols            Mm-hmm (affirmative).

## CHAOSGROUP

- Koki Nagano            So without sound, it can be more difficult to detect a fake video. But with face-swapping, basically, we have only the face mask-
- Chris Nichols            Right. You can detect the edge of the mask.
- Koki Nagano            Right, yeah. So it's not very symmetrical, so you need to do some kind of composite to the background. Yeah, so some are easier to-
- Chris Nichols            To detect than others. Interesting. Interesting. Well, this has been really good and very educational. We're actually just over just about an hour, so that's perfect timing. But, Koki, this is fascinating work. So if anyone wants to know more about the different tools that Pinscreen has got, they just go to Pinscreen.com, I'm assuming?
- Koki Nagano            Mm-hmm (affirmative).
- Chris Nichols            Right. And there's some really cool stuff there. There's a lot of stuff on the news about Pinscreen, as well. I'm sure you guys can follow that. We're in a world right now where a lot of people are concerned about where this technology is going and what it's going on, and I think we have a responsibility like we're doing right now, to educate everyone so that they know that this is possible. And the more they know, the less scary it will be, because they'll be able to use their education to make that assumption, so I think it's very important to have these conversations.
- Koki Nagano            Mm-hmm (affirmative).
- Chris Nichols            Well, thanks so much, Koki, for doing this.
- Koki Nagano            Yeah. Thank you so much for having me, Chris.
- Chris Nichols            Perfect.