Noah Kravitz:

Hello and welcome to the NVIDIA AI Podcast. I'm your host, Noah Kravitz. This past August, three of NVIDIA's research leaders gave a special address at SIGGRAPH, the annual International Computer Graphics and Interactive Techniques Conference that's been running since 1974. One of those people is here with us today.

Sanja Fidler is a VP of AI Research at NVIDIA, where she leads the NVIDIA's Spatial Intelligence Lab in Toronto, Ontario, Canada. Sanja is here to tell us about the lab, to talk about the research she's most excited about right now, including what was presented at SIGGRAPH. And to share a little bit about her own journey through the worlds of research and artificial intelligence.

So without further ado, let's get to it. Sanja Fiddler, welcome and thanks for joining the AI podcast.

Sanja Fidler:

Hi Noah and hi audience. I'm very excited to be on this AI podcast.

Noah Kravitz:

We are very excited to have you. Thanks for taking the time. There's a lot going on obviously, and congratulations on the special address and everything else at SIGGRAPH.

So we wanted to start with a little bit about your own journey. You followed your passion for computer vision and artificial intelligence across Europe and into North America. Can you tell us a little bit about what first got you interested in the field and how your journey took you to Toronto?

Sanja Fidler:

So, maybe I'll start with my youth and there were actually three important break points that led me to where I am.

And the first one actually starts with my dad. So my dad would sit on a chair next to my sister and me and tell us bed time stories. And surprisingly, he was very good at it. He was a scientist. And so he would tell us stories about scientists, for example.

He would tell us about Nikola Tesla, who was born in Croatia. And my mom was also born in Croatia. I was born in Slovenia.

Noah Kravitz:

And was this in Slovenia? Where did you grow up?

[02 min 04 sec] **Sanja Fidler:**

Yeah. I grew up in Slovenia, so my dad. Okay. Born in Slovenia. My mom was born in Croatia. My dad was born in Slovenia and I was born in Slovenia. Okay. So he would tell us stories about, you know, how at a young age Nikola jumped from the roof of their house holding an open umbrella, thinking he would fly, you know? And every night there would be a new episode about his inventions that led to the creation of radio, and alternating current obviously we didn't understand, but he made it sound fun. And the competition with Thomas Edison, right?

It would be almost like a Netflix series, and for a child it was very exciting.

[02 min 41 sec] **Noah Kravitz:**

That's amazing. Yeah.

[02 min 42 sec] **Sanja Fidler:**

I could not wait to hear more till the next day. So kind of like, my childhood heroes were not the movie stars or music stars, they were scientists.

[02 min 52 sec] **Noah Kravitz:**

Awesome.

[02 min 53 sec] **Sanja Fidler:**

And that really kind of shaped me.

So perhaps not surprising, you know, one day I appear in front of my parents and proclaim I want to be an inventor. And there was even a photo of me and my sister, actually my sister dressed as a robot. It's quite a bit of money for her to put some cardboard boxes around her. And maybe not surprisingly, she became an economist. So this, you know, moment pretty much settled my profession. I was gonna be right, a new inventor, and this was at a very young age.

The second moment was really thanks to my mom and this was in primary school. I was, you know, very young and at some point I got pretty ill. It was something like COVID almost, I think it was called Whipping Cough or something. I was home with a fever and coughing two to three months, and I basically missed a lot of school. I missed, you know, fractions, a whole big chapter on math, right? I had no idea about it.

So I come back to school and of course I didn't understand anything they were talking about, and I developed some sort of resistance to going to school. And before a math test, I threw a tantrum, like crying hysterically on the floor. I hate math. I don't wanna go back to school. And my mom is actually a teacher. And of course, you know, having a school hating child, that was not an option.

Noah Kravitz:

Yeah, it happens.

Sanja Fidler:

So even though she was an English teacher, she would sit down with me and work with me on math, and she made it really interesting. She had this really nice way of teaching me through giving me puzzles, math puzzles. And I began to really love it. You know, as kids understand things, they grow to love it. And I think to this day, what drives me at the core is solving problems. And I think that still kind of stuck with me.

Noah Kravitz:

Yeah.

Sanja Fidler:

It was pretty much settled what I would study, I was determined at age, I don't know, 12, 13, I was gonna study math.

And the third moment was really kind of thanks to my grandma. I was already doing my PhD and I decided to work on computer vision. I saw this one talk. It started actually with math, even my PhD, and I saw this talk on someone recognizing cats and dogs, you know, it was very early AI at that point.

And it just kind of spoke to me. You know, I was always kind of dreaming of robots and computer vision felt like the first step. I was there doing my PhD and my, my grandma, she was a very smart woman. She was actually one of the first female plastic surgeons in Yugoslavia.

Noah Kravitz:

Oh, wow.

Sanja Fidler:

Sanja Fidler: Yeah, yeah, yeah. She was always telling me these stories, how she graduated med school and the day they were graduating and they were out having fun and, sirens came out. World War II started, and she had to basically just go to the operating room and that was her next four years.

Noah Kravitz:

Wow.

[06 min 01 sec]   **Sanja Fidler:**

Basically fear became alien to her. And it wasn't for me, really. So I studied my PhD in Slovenia really for the fear of leaving to the wide open world alone, as a woman. I was somehow not encouraged. My mom would scar the hell out of me of doing that. So towards the end of my PhD I was working on this AI, something similar to deep networks, just my own take on it. I was presenting at a conference and a famous professor at UC Berkeley stopped by the poster, really liked it, and invited me to visit his group at Berkeley.

And I was beyond excited, but I still carried this kind of weight of fear and hesitation. I talked to my grandma and she said, 'You know, Sanja, don't listen to your mom. Just go.' Actually, she passed away a few months later. That was January 13th, 2009. And the next thing I remember, I'm sitting on a plane, I look at my plane ticket to California. And it was January 13th, 2010. It was exactly one year later.

[07 min 12 sec]   **Noah Kravitz:**

Exactly a year.

[07 min 14 sec]   **Sanja Fidler:**

I am not kidding. Like this was exactly. It was.

[07 min 17 sec]   **Noah Kravitz:**

It was meant to be.

[07 min 18 sec]   **Sanja Fidler:**

Meant to be. I was both scared and excited, but you know, chapter two of my life was really about to start.

[07 min 25 sec]   **Noah Kravitz:**

Was that your first time traveling abroad?

[07 min 28 sec]   **Sanja Fidler:**

No, I would go before, you know, just visiting New York with my family. This was the first time I went alone and was leaving, you know? Packed my bags and here it was, you know? It was scary, but …

[07 min 41 sec]   **Noah Kravitz:**

And you landed in Berkeley of all places.

[07 min 44 sec]   **Sanja Fidler:**

I stayed there a few months, maybe seven, eight months, and came back, graduated, and then I did my postdoc and I was at U of T. So that's kind of what brought me to Toronto.

[07 min 55 sec]   **Noah Kravitz:**

Right. Amazing. I feel like the graphics and interactive industry owes a big thank you to many members of your family for all the — inspiring you and then grandma kind of giving you that nudge and everything. That's amazing.

Why Toronto? What was the link that brought you to Toronto?

[08 min 13 sec]   **Sanja Fidler:**

Actually the U of T, University of Toronto was doing really great stuff in deep learning. Like I said before, that was kind of my PhD. I was really inspired by doing this hierarchical representation to recognize objects. I was reading all these neuroscience papers that basically said this is how the brain works, right?

And I was, and I was isolated to kind of have my own take of how that would look like. And then I was reading this deep learning paper, well papers, and that was really appealing to me. I was kind of going between Berkeley and U of T and I decided to go to U of T to learn from Geoff Hinton and the other people like that. And that's why I landed here.

[08 min 57 sec]   **Noah Kravitz:**

Amazing. Great choice. And so you've been in Toronto since?

[09 min 01 sec]   **Sanja Fidler:**

Yeah, I mean, I was there for a postdoc and then I got a research assistant professorship in Chicago. So I did a small stop there for a year and a half. Okay. Then a position, a faculty position opened at U of T and then I came back in 2014.

[09 min 19 sec]   **Noah Kravitz:**

Amazing. And so now you head up the Nvidia Spatial Intelligence Lab in Toronto.

[09 min 24 sec]

Sanja Fidler:

That's right. I joined NVIDIA, that was 2018.

[09 min 27 sec]

Noah Kravitz:

It's about seven years ago.

[09 min 29 sec]

Sanja Fidler:

Seven years ago. I actually met Jensen at a computer vision conference — that was 2017 — and we had a really great chat about simulation. I was already working on simulation for robotics at that time, and I was telling him about it, and I think he was also thinking about it. So it was a great conversation. Then later he gave me a call or, we went on a call, and he said come work with me. And I had other options, but the fact that he said come work with me. and not for me, just told me everything about this place I was joining and that was it.

[10 min 04 sec]

Noah Kravitz:

What a great story. That's fantastic. So tell us about the lab. What, what is… For those listening who might not fully get the term, what does spatial intelligence mean and what's the charter of your team? What are you doing at the lab and how — you may have just said, sorry, I was imagining Jensen and that whole conversation — but when was the lab founded?

[10 min 24 sec]

Sanja Fidler:

2018. May. That basically started with me, and then we slowly grew and also increased scope. So we recently renamed ourselves to spatial intelligence. I would say it's a new encompassing word.

So spatial intelligence essentially denotes intelligence in 3D, right? Intelligence in the 3D world. So the same as we have LLMs representing intelligence in language, you have this, all these families of vision language models for intelligence in 2D images. Now we need to build the same capabilities, but in 3D. And the question is of course, what that is and why.

Maybe I'll motivate with robots because really, that's one of the prime motivations. At the end of the day, robots need to operate in the physical world, in our world. And this world is three dimensional and conforms to the laws of physics and there's humans inside that we need to interact with. We typically hear the term, such ai that operates in a real physical world as physical ai, so I'll maybe use that term quite a lot, right?

[11 min 31 sec]

Noah Kravitz:

Yep.

[11 min 32 sec]

Sanja Fidler:

The physical AI is really kind of the upcoming big industry. Very likely larger than generative, agentic ai. Jensen typically says everything that moves, all devices that move will be autonomous, right? So that's kind of the vision.

So a robot to operate in the real world obviously needs to understand the world. What am I seeing? What is everything I'm seeing doing? How is it going to react to my actions? So understanding it needs to act... you know, if I want to drive you from A to B, make you dinner, I need to actually control that robot to make an action.

But then there are two other capabilities needed that are perhaps a bit less obvious. So basically it's 3D, virtual creation, and modeling and simulation. And the reason is that robots. need to have a virtual playground that almost perfectly — or, we would like it to mimic the real world as faithfully as possible — where basically they can train their skills and also test their skills before we are going to deploy them in the real world, right?

Like this is basically the critical thing we need to solve for deployment of robots. Basically spatial intelligence kind of comprises these four core capabilities, which are modeling, so creation of a virtual world, but then also modeling it, how it evolves in time based on our action.

Understanding and actioning through the world. And obviously applications are more than robots, you know, architecture, construction, gaming, everyone that has 3D data, 3D world data. We first started with this virtual world creation, so content creation, and then we explore, because in order to develop the spatial intelligence you also need physics, which evolves in time and understanding.

[13 min 22 sec]

Noah Kravitz:

A year or so ago, maybe less... There's so much that has happened with Generative AI in particular in the past few years that it kind of blurs together sometimes when I talk about it. But I remember when video models started coming out, the first, you know, Sora from Open AI and some of the other ones, and discussion around, well, these video models are actually also physics simulations. We're discovering — we thought we were making a video model, but now we're realizing that there are properties of physics happening inside of the videos that are output and, and all of these things. What makes a good physics model? And when you're talking about modeling things that are going to happen in the future, I've also heard that described as, 'Well, what an AI model does is really predicting what's going to happen in the future." And if it's a video. It's output, it's sort of frame-by-frame.

How do you think about the four things you just described relating to one another? And I don't know, maybe you can talk a little bit about physical AI in particular and, and how the evolution of how these models came to be so accurate that we can now use them in simulations.

Sanja Fidler:

Yeah, so in NVIDIA Cosmos and the models you're describing, right, Sora, Veo 3, and so on, learn their capabilities from videos. And, especially Nvidia Cosmos is targeting Physical AI, which really means that it's doubling down on modeling physics, not necessarily the creative aspects, but physics capturing how our world works. So it's forming these world simulation capabilities by learning purely with videos, and we specifically target collecting videos that are real-world recordings. There is no human editing involved and, and if there's any graphics data, it's actually all physically simulated.

How we're using physics is mainly for benchmarks, actually. So, you want to create, because you have full control, right? I can have two bouncing balls, three bouncing balls with this material, and a more complex wall. And there you can really go like, you know, every single test. How good are you with that? How good are you with that? And that's our test. And you kind of, hill climb that performance,

It's an evolution of models, right? The first world model came out … I think it was Jürgen Schmidt Huber, 2019. It was almost parallel to us. Ours came like a few months later. Where the idea was really kind of like AI replaces the game engine, you know, AI creates the world, you have the user interaction, and the next frame is not human written code, it's the AI.

Noah Kravitz:

Right, it's generated.

Sanja Fidler:

Obviously that was early on. It was… I forgot exactly what they were using. We were using GaN, ours were called Game GaN, so, Pacman, you know, so you could actually play Pacman on a keyboard like those videos for AI.

Noah Kravitz:

We had an episode of the podcast with somebody who created, um, GaN Theft Auto. So like Grand Theft Auto, but being generated…

Sanja Fidler:

That was actually us!

Noah Kravitz:

Oh, that was yours! Okay.

[16 min 26 sec]   **Sanja Fidler:**

Yeah, that was our stuff. So that's cool.

[16 min 27 sec]   **Noah Kravitz:**

Yeah, yeah, great! I, forgive me, I don't remember offhand who the guest was, but yep.
That was so cool.

[16 min 32 sec]   **Sanja Fidler:**

So, you know, people just got crazy and it was amazing to see what, you know, where it, it
went.

[016 min 37 sec]   **Noah Kravitz:**

Yeah.

[16 min 38 sec]   **Sanja Fidler:**

We actually also applied it to driving, that was 2021. It was called Drive Again. Some
similar technology, but just a lot of autonomous driving videos. And it almost kind
became a driving simulator, you know?

Cosmos really took to new heights, but at the time it was kind of like imagining how this
could be useful for physical applications. So that was all kind of GaN based with all kinds
of known limitations. And you know, in the meantime, diffusion models came out and it
was clear that, you know, like that's also the next big leap in video modeling.

And actually in 2023, we kind of partnered up with some of the students that did the
latent diffusion. That really was kind of a big breakthrough in images because you didn't
model pixels anymore, but this kind of latent code made it significantly more efficient. So
we kind of applied that and extended that to video, and that led to video LDM (Latent
Diffusion Models), which really became, you know, you could see the future by looking at
those results. Obviously it was not so yet, or Cosmos, but we were onto something.

And then the industry actually kind of switched to this latent diffusion architecture. And
then it's about scaling and obviously the architecture changed a little bit behind the
scenes and data and so on.

And that basically is creating the modern age models.

[18 min 03 sec]   **Noah Kravitz:**

So I understand that your lab has grown recently. Can you talk a little bit about the new
areas that the lab's now encompassing and how that kind of furthers the overall goals?
The overall charter of the lab?

Sanja Fidler:

So when I joined, we joined Rev's organization and Rev was building Omniverse. Omniverse is this state-of-the-art simulation platform where robots can be robots, as Jensen says.

Noah Kravitz:

Right.

Sanja Fidler:

nd talking to Rev at the time, he mentioned there was a huge team working on it. Obviously they were able to render really fast, they had real time ray tracing and so on. So really kind of the key missing piece was content. Bear in mind this was like 2018, right? It was like baby times for that. And that's how we started. I said, okay, how can we actually make this platform workable? Especially for Physical AI where it's really about modeling the world, which is messy, diverse — it's really challenging.

So we started with content and we developed a bunch of techniques for that. And through kind of the period of our lab, we became more and more ambitious and we realized that the pipeline for Physical AI or this 3D spatial intelligence also needed to change because you need to have better physics algorithms. Physics algorithms interact with each other, you know? I've got this bottle here, hopefully the audience can hear it [shakes bottle]. It's plastic, and whether there's water inside, I can put it on fire and, you know, there is no cheating. Like in a game where I can kind of stage it. This needs to all be simulated.

Noah Kravitz:

It's real.

Sanja Fidler:

Yeah, it's real! It needs to feel real, right? I can put my finger on it. Bad things happen, right? Like the robot, if it's training, then it needs to kind of experience it in this way.

So it was clear that we needed the next evolution of physics, and I can join the team. And also, you know, perception is obviously important. And Laura joined the team and she was, she's very interested in 3D perception, but going towards open world. Meaning, anything in this room, I should be able to recognize it and understand my affordances with it. And then that can lead to a better action. So we expanded the team basically by building blocks that we actually need.

Building the full stack for spatial intelligence.

[20 min 29 sec]   **Noah Kravitz:**

And you mentioned Omniverse. Your lab has been very involved with the creation of Omniverse. What are some of the innovations, some of the research breakthroughs — you mentioned physics models improving, what are some of the other innovations that really made Omniverse possible and helped to grow into what it is today?

[20 min 47 sec]   **Sanja Fidler:**

I think first of all, Omniverse is created by many teams at Nvidia. Much, much, much larger than any single team, really kind of the vision of Jensen Rev. It has a mountain of technology for real-time tracing powered by DLSS size that makes, AI in the loop, AI powered physics, solvers like I was saying... so that's just scratching the surface and I really can't take credit for any of that.

I can maybe tell you a little bit about where, what we were thinking when we started with our 3D content creation work.

[21 min 22 sec]   **Noah Kravitz:**

Okay.

[21 min 23 sec]   **Sanja Fidler:**

And I would really say that we doubled down on two directions which both turned out to be very important in the end, and it's really kind of this perseverance through time that created something of value. So the first one was, okay, clearly there's a graphics pipeline. We know everything in how that works. So why don't we lift images and videos to 3D to be fully compatible with existing graphics pipelines, and we really double down on differentiable rendering as this foundational technology. Meaning, graphics goes from 3D and renders to images if this is differentiable, meaning kind of like amenable to AI.

So this path led to one of the first image-to-3D models that we'll call gamblers, one of the first generative models of 3D assets, gas, 3D, and as the latest achievement. We also made foundational improvements for 3D pls. I don't know whether I need to explain it in all the detail, but essentially it's a really — like a new neural graphics primitive that you can easily optimize from videos. And we added tracing capabilities to it. And at SIGGRAPH we actually announced integration of, we call it 3D GR00T. 3D GR00T omniverse. So basically you now can download Omniverse or Isaac, which basically helps to train robots.

You can scan, with your phone or whatnot, this environment. And boom, you have it in Isaac and you can start training robots in here. Like there is no, you know, you don't take weeks for it.

[22 min 57 sec]   **Noah Kravitz:**

Noah Kravitz: It's amazing. It all makes sense in terms of, you know, looking at the way you're describing the way things have built up and the building blocks and adding features and, oh, cool.

That makes sense. And then I sort of listen, you describe like, oh, take your phone, wave it around the room, and now the robot can trade in the room. And it's, it's still, it's just so exciting. It's so mind blowing. It's,

[23 min 16 sec]

Sanja Fidler:

Yeah. Very cool. Yeah. And exciting, but, but that's basically what you want, right? Like scale — I want to just go and take what's up when we hear and, and sim, right? And, and boom, the robot is training.

So the second one, the second path is I, we kind of saw the fundamental, some fundamental limitations of this graphics pipeline because you need to also model agents and physics. It all kind of also felt daunting.

So we also made this bold approach of AI that it's basically the world model that does the whole content creation, world simulation based on user interaction and it's all one. And that was the chain of models that I described earlier. So two different things that all now kind of came together in really, I think, useful capabilities.

[24 min 01 sec]

Noah Kravitz:

So how has the advent of AI and 3D content creation and, sort of specifically in workflows changed the way that people get the work done? The way that researchers or designers can create objects and create scenes and kind of manipulate things? What's the impact of AI been so far on these workflows?

[24 min 20 sec]

Sanja Fidler:

I think this technology really democratizes access to these tools and basically it gives everyone the chance to become a creator. I have no idea how to use 3D software. I might try it a few times, but now I could be reasonable. I could reasonably if I wanted to, actually use robotics. I can reasonably get this object in a simulated world.

The cool thing is that it also gives additional superpowers to creators that have the talent, so artists, designers that can actually use this technology to now do many more creative things. I have seen so much amazing stuff coming out that I wouldn't even think of. I think it's really kind of empowering to the entire population, in different ways, which is great to see.

[25 min 10 sec]

Noah Kravitz:

We had Danny Wu from Canva on recently. He's the head of AI products there, if I got that right. And he was describing a similar thing, but kind of more on a level I can relate to, because I write, I talk, I, I mostly work with words. I can't draw or paint to save my life,

and so that ability, having that superpower, if I want to see how something might look, an idea, it lets me do that now. And so I can only imagine the 3D physical world with 3D design, talking about simulations. The stuff you've seen must be pretty cool.

Sanja Fidler:

Yeah, I think so.

Noah Kravitz:

So talking about this 3D world in Physical AI, you spoke to it a little bit earlier, but how are all of these advances with the technology, and computer vision included, enabling robotics, autonomous vehicles? You talked about it a little bit, but maybe you can kind of put a point on how physical AI has really started to take off.

Sanja Fidler:

If there's anything that people take away from this talk, it's that physical AI can scale through real-world trial and error. It's simply not possible to put my car out there or a robot out there and it is going to mess up my kitchen here by bumping everything and so on. This is super expensive, unsafe, and it's just going to take us forever to get there.

So yeah, simulation is really the answer here, right? And if we do it right, if we are actually able to use computer vision and other techniques to basically go somehow create these virtual walls that feel real, then it's possible to train this kind of parallel virtual universe, and safely, essentially, basically accelerating time before we can deploy robots and also bringing the overall cost down, right? Because now we are doing it in the cloud as opposed to having...

Noah Kravitz:

To remodel your kitchen after every test. Yeah.

So what are some of the methods that are key to making simulations physically accurate or more physically accurate as we go?

Sanja Fidler:

I think the jury is still out. Okay. How exactly to achieve physically accurate, like something I can completely trust, simulation at the scale, diversity and realism of the real world. It's hard in the traditional way with different physics solvers, you know, that's hard, right? For the models, it's also kind of hard. There's still hallucinations and sometimes objects disappear, go in one another, and obviously that's going to keep improving.

So the likely success is going to come in some sort of a combination of both. And obviously we're going to keep pushing on each direction, each that right, be as good as

possible, and maybe in between until we reach a point. There is a combination of that, right? Using these world models with a small, traditional approach that really make sure that physics and simulation is correct.

And the other very important message is that in computer vision and robotics, the really big breakthrough is VLM. So, visual language model that is able to reason, okay? This is basically how humans navigate the long tail of very diverse and rare scenarios. So the physical world. So we're kind of bringing that knowledge from language into the physical world. We are encountering completely new situations that we have never seen before in training. And now the VLM could come to our rescue, basically like really saw this long tail. And that is really kind of the discontinuity from before. That is a tool that we have now that before was missing.

So that's probably like the, the most bold statement I can make right now.

[29 min 01 sec]

Noah Kravitz:

Fair enough. Do the traditional methods get baked into these models, or how do you go about combining them? The two approaches?

[29 min 08 sec]

Sanja Fidler:

What you could do, for example, is you can use kind of the traditional, way, which is also not full of AI everywhere, to kind of have a coarse simulation in 3D with solvers that we know how to model certain effects. So you can make that simulation, you can render it out, and that becomes a guidance to a world model. I think that is input. Right? Okay. Correct. It's kind of telling me, oh, should be roughly here and here, and then it becomes much more feasible to create pretty pixels out of that, both in time and space.

So that's kind of like what we're thinking right now.

[29 min 45 sec]

Noah Kravitz:

I'm speaking with Sanja Fiddler. Sanja is vice president of AI research at Nvidia, and we're talking about the work that her spatial intelligence lab in Toronto has been doing, along with the evolution of AI and models and solvers in the mix, and all the things that go into making these models more accurate so we can rely on them and trust them as Sanja, as you were saying.

And we've also talked a little bit about SIGGRAPH, and I mentioned at the top that you gave the keynote a special research address alongside some other NVIDIAns. At this year's SIGGRAPH, what are some of the notable things from NVIDIA's presence at the show that maybe we can impart to listeners here?

What are some of the things that they should take away from what NVIDIA did at SIGGRAPH?

Sanja Fidler:

This year at SIGGRAPH we really tried to send a message on Physical AI in the keynote. And the reason is because this is a really important area with a big impact. And the community has a lot to give. A lot to give.

A lot of expertise is already there. The gaussians flats, nerves, I mean, that all comes out of that, right? We discussed how simulation is key. Like that literally means simulation is key. So everyone in the audience should feel empowered to help us in this quest of robotics, you know?

And the cool thing is that it feels all so early stage. Like I said before, it is open-ended. We don't kind of know yet. You know, we are hypothesizing. So, we really hope the audience kind of connects with here's a new challenge for you.

Noah Kravitz:

Yeah, yeah.

Sanja Fidler:

How can you, you know, what do I do next? And graphics is very mature, but here's a new challenge for you that maybe needs to out, you know, think outside the box.

So I think the key to success, we suspect will be the combination of Cosmos — and so this is NVIDIA's world foundation platform, and this is both video generation, so simulation of the world, as well as reasoning. Reasoning about the laws of physics is reasoning about all the agents in the scene, the scenarios and so on, and, and physics simulation. So all these three pieces interacting together, that's our bet of creating really physically, but also semantically, accurate simulations of the real world in the future.

I think that's really kind of like what I hope the SIGGRAPH audience takes away from the keynote.

Noah Kravitz:

That, that spirit of — I mean, it even goes back to what you said about when you met Jensen and he invited you to come work with him. That spirit of collaboration — open source being what it is, conferences obviously, but we've had guests across all different industries come on the show and talk about how important it is to share research and trade notes with other people who are on the other side of the world, working in other industries, et cetera.

As AI continues to touch and evolve and change virtually every industry you can think of, how important is that? Or, what are you getting from this experience of working across so many industries, and does it feel like AI is kind of bringing industries together or does it feel like it different industries are kind of hunkering down and siloing in their own approach to how they use these emerging technologies?

[33 min 00 sec] **Sanja Fidler:**

Yeah, it's definitely bringing them in, bringing them together, right? Because yeah, the workflows essentially are very similar. At some point they're very similar, and the difference is the data and the expertise, domain expertise, that's different.

And actually there is even sharing, you know how I do autonomous driving versus humanized robots versus factory simulation architecture. There's some commonalities between things that could be shared and, having this data-driven approach to simulation could really bring industries together and benefit from one another and build tech that can essentially make all of us, all of these industries better and open source.

You mentioned open source. I am a believer in open source and Nvidia is also a big believer in open source. Like I said, in a lot of areas we are also still early on and that's the only way to keep progress going, and really build up these capabilities.

[33 min 56 sec] **Noah Kravitz:**

Absolutely. Even though in a lot of ways, the timeframe that we've been talking about, the last seven, eight years in particular, isn't all that long, in AI terms, and particularly in this recent generative AI revolution, it's a long time. So you've been doing this almost since the beginning of early object recognition and AI kind of now through to everything we're talking about today.

What's next on the horizon? Is there a breakthrough that you're either waiting for or you know, maybe more secretly kind of thinking, I think this is going to happen soon? Whether it's something that's particular to your own work that you're doing, or more broadly, what's the next big breakthrough in AI that you are looking forward to or maybe just hoping to see?

[34 min 39 sec] **Sanja Fidler:**

Well, I think it's gonna be robots, right? So I started the story with my sister in a car,

[34 min 45 sec] **Noah Kravitz:**

Your sister, of course,

[34 min 37 sec] **Sanja Fidler:**

And me dreaming about a robot taking the dog out in the morning. My parents made me do it, and I really like to sleep in the morning. That was kind of the early dream.

My grandma lived 30 years alone. My grandfather died quite early. So the first talks I gave as a faculty were all started with a grandma and a Wall-E cute robot in a kitchen talking to each other, and the robot helping. It's just a kind of common thread of let's build this technology because it can be really powerful and useful. And I now believe after many years in this field that we're likely going to see that in our lifetime.

[35 min 32 sec]	**Noah Kravitz:**

Yeah.

[35 min 32 sec]	**Sanja Fidler:**

Robots in some form — autonomous cars are already, out there to some extent. And more is coming. So that's the breakthrough I am looking forward to.

[35 min 46 sec]	**Noah Kravitz:**

Have you ever had a robot in your home with you for a period of time — have you ever lived with a robot?

[35 min 52 sec]	**Sanja Fidler:**

I have a robot that wipes the floor. I would love to have something that does more than that.

[35 min 59 sec]	**Noah Kravitz:**

Noah Kravitz: So, Sanja, as we look to wrap up here — and this has been fantastic, again, thank you for taking the time —what advice would you give to researchers out there who are interested in the work the Spatial Intelligence Lab is doing? Who might be interested in collaborating, working with you, and in some way joining the lab, or collaborating from afar? And then in particular, what are some of the skills and research areas that you think are becoming increasingly important now and will continue to be, at least for the next few years?

[36 min 29 sec]	**Sanja Fidler:**

Sanja Fidler: Actually the bar that we have is both low and high, so I'll explain what I mean. I think what we are looking for is people with immense passion. I feel like I still haven't lost the passion of the first day. You — I wake up and I am excited. So I think the passion is what drives us forward. The energy.

[36 min 52 sec]	**Noah Kravitz:**

This is only a podcast, but the passion comes through in your voice. So, yeah, I'd say you're doing all right.

[36 min 57 sec]	**Sanja Fidler:**

I mean, that's what drives it, because as a researcher, life is not easy. Most of the time

things don't work, you know? That's basic do. It could be six months or a year where you don't get results. So you're really that kind of like, you know, passion and energy that makes you keep going. I think wanting and having the ability to go technically deep is very important. Not jump from one thing to another as things get hard, but like, let's learn really the fundamentals to the level we need so we can innovate.

And I guess maybe to my first point, the high level of perseverance, right? We want to keep going and there is no wall thick enough. Right? The rest, I think we can teach people. Like if you have these basic things, a lot of the other stuff comes along. So in terms of like, you know, it's mostly also interest, right?

So we are very interested in this 3D world modeling and understanding 3D worlds. So people that are interested are share a passion for the same topic, please contact us. We would be very happy to work with you.

[38 min 07 sec] **Noah Kravitz:**

Fantastic. For listeners who want to learn more about the lab, about the work we've been talking about, where are the best places to go online?

Is there a homepage for the lab? Is it on the Nvidia site? Any social media handles to follow? Where would you send them?

[38 min 23 sec] **Sanja Fidler:**

It's all on the NVIDIA website. Probably if you, if you search for Spatial Intelligence Lab, Nvidia or Toronto, Nvidia, which is our old name, it should. Pop out. Yeah.

[38 min 35 sec] **Noah Kravitz:**

Fabulous.

Sanja, again, this has been great. I think that the story, just imagining, your dad telling you those stories as a kid and with your sister, and then the advice your grandma gave you: Overcome your fear. Get out there. Just absolutely fantastic.

Congratulations to you and all of the team, everyone you work with for all the work you've been doing, and SIGGRAPH of course. And we really look forward to following your progress in the future. Best of luck.

[39 min 01 sec] **Sanja Fidler:**

Thanks, Noah. It was really fun talking to you.