[Scientific Innovation Series 9]

The Metaverse and Digital Realities Transcript

Date: 08/27/2021 (Released)
Editors: Ji Soo KIM, Jooseop LEE, Youwon PARK

Introduction

Yongtaek HONG: Welcome to the Chey Institute's Scientific Innovation Series. Today, in the 9th iteration of the series, we focus on the Metaverse and Digital Realities. I am Yongtaek Hong, a Professor of Electrical and Computer Engineering at Seoul National University. I am particularly excited to moderate today's webinar with the leading experts and scholars on the metaverse, a buzzword that has especially gained momentum during the online-everything shift of the pandemic. Today, we have Dr. Michael Kass and Dr. Douglas Lanman joining us from the United States. And we have Professor Byoungho Lee and Professor Woontack Woo joining us from Korea. Now, I will introduce you to our opening Plenary Speaker. Dr. Michael Kass is a senior distinguished engineer at NVIDIA and the overall software architect of NVIDIA Omniverse, NVIDIA's platform and pipeline for collaborative 3D content creation based on USD. He is also the recipient of distinguished awards, including the 2005 Scientific and Technical Academy Award and the 2009 SIGGRAPH Computer Graphics Achievement Award.

Plenary Session

Michael KASS: So, my name is Michael Kass. I'm a distinguished engineer from NVIDIA. And today we'll be talking about NVIDIA's view of the metaverse and how we need an open metaverse. And we believe that the core of that metaverse should be USD, Pixar's Universal Theme Description. Now, I don't think I have to really do much to introduce the metaverse to this group, but the original name goes back to Neal Stephenson's novel Snow Crash in 1992, and the original idea probably goes back further. And the metaverse is about people coming together inside of a virtual space. It can be for a variety of different purposes. It could be for entertainment. It could be for collaboration. It could be for community. It could be to get practical things done. The idea of the metaverse is such a big idea that it's not possible for one company or one

organization to control it all. So, we're confident that there will be multiple metaverses in the future. Some of them will be very specialized. But in addition to whatever specialized metaverses end up being created in the future, we think it's critical that there be at least one open metaverse that is a metaverse based on open standards to which anybody can connect and over which new single interest individual organization has control. So what would that look like and how would we get there? We think that the path to the metaverse is really to the open metaverse. It's really much like the path to creating a true 3D web.

What would that involve and what does that mean? Well, we need to adopt an open, powerful, expressive 3D scene description. We need to make it possible to connect to that description from a wide variety of sources. And we need to make sure that every piece of that is based on open source technology. People have spent a lot of time - spent decades - trying to bring 3D into our current Web. The efforts go back as far as 1994 with VRML, virtual reality modeling language. In 2001, we had X3D. In 2015, glFT. Then, in 2019, WebXR. And in 2021, WebGPU. Now, all these efforts have been important, influential and pushed forward the idea of 3D in the Web. But we believe that they're never going to really get us where we want to on a time scale that we want in order to form a true 3D web. And why is that? Because these are incremental improvements in a system where 2D is really at the core and 3D is being added. We believe that's backwards. So, in order to create a true open metaverse that anybody can connect to, we think there are four differences in approach that we have to take versus the kind of efforts that I mentioned that were going on in the past.

First of all, we need to start with the right 3D representation. So in the past, we started with the Web as it was and tried to add to it. And at the center of the Web is HTML. But really, 3D is much, much harder than 2D. So what you really want to do is start with 3D and embed 2D inside because the complexity is much, much better on the 3D side. Now, given that complexity and given the size and subtlety of those models, we think it's critical that the 3D representation has to be incremental at its core. What do I mean by that? Well, when a Web page changes, you can, generally speaking, resend the whole page if necessary. The amount of data is just not that great. And so you see a lot of situations where Web pages just refresh entirely no matter what it is that has been modified. This can work in 2D. It has no chance in 3D. Because in 3D, the amount of shared data that describes the scene could be hundreds of megabytes. It could be

gigabytes. And there's just no way that you can retransmit all of that in an effective way every time something changes. So, whatever we want at the core of a true 3D web needs to be incremental. It needs to be that I send you a representation and then I can send you any number of changes so that we can keep up to date without having to start from scratch. And fundamentally, HTML and most other Web technologies at the center of the 2D Web are not fully incremental in this way, where you can just modify anything you want little by little.

A third direction we think is really important to create an effective 3D web is to pay attention to high end hardware. The tendency in the Web these days has been to try to make sure that everything looks the same everywhere. And that requires basically going down to the least common denominator of hardware. And there are good reasons to want to do that, the best of which is to try to make sure that whatever the content is universally accessible. And we think that's a laudable goal. But today, it's possible, relatively inexpensively, to do remote rendering. You don't have to do all the rendering on the client. A single server on a low cloud platform these days costs less than a dollar per hour, and that server is potentially capable of providing a service to many different users. So the cost of doing it in a remote location, either on the edge or in a data center, it's really not that prohibitive. And so, we think it's important to target where we want to be and allow that rendering to happen in the data center if it needs to now and then over time, it could migrate to the client.

The next thing we think is really important is to make sure that whatever that representation is, that it supports offering at all times. Now, most of us are familiar with game engines, which tend to have a baking step that they want to make sure that everything runs as fast as possible, they take the representation of the world that was used for creating it, and they boil it down. They bake it into something which is optimized for performance. Obviously, there are reasons to do optimizations, and this is this is very critical. But we think that the kinds of experiences that we'll have in the open metaverse are going to be very varied and not just limited in the way that many of today's games are. And in order to keep open the possibility to make anything still potentially achievable, we think it's important to maintain the kind of representation that you want for offering and be able to swap in and out that representation as needed for any part that could potentially be desired to change.

OK, so if we're going to achieve this, if we're going to build an open metaverse based on the kind of ideas that I mentioned, we need a common representation of the virtual world and one that is powerful and one that is extensible. And we think that representation already exists. We think that representation is USD, the universal theme description that was developed by Pixar and very kindly of them, open source to the entire community. So, Pixar developed USD to meet the needs of their own computer graphics productions. And it evolved over the course of many of their films and has a very long list of features, which I won't go through individually, but which are sufficient for the complexity of the world that they created. And I think most of us would agree that if an open metaverse can come close to that complexity, it would be delightful. So, from that point of view, it's really very, very appropriate. Now, I want people to understand basically how USD works and its foundation. It describes the world as a set of properties and then values for those properties. And then collections of properties together represent higher level objects. In addition, it puts those collections of properties and values onto a set of layers, and by dividing those properties into layers, for Pixar, it was important enabler of collaboration for large teams. But for an open metaverse, it's exactly what you want so that different parts of the world can be modified simultaneously by different actors or by different services or by different portions of the metaverse.

So, what is it about USD that makes us believe it's the right solution? It begins with a set of schemas so the schemers can, for example, describe geometry, cameras, lights, even simple rigging for characters, and it allows you to describe relationships among objects. Very importantly, it includes inheritance, so you can create classes inside of USD. So, for example, you could build a class to describe trees. You could build a class to describe a certain kind of tree, like pine trees that you might find in a forest. You could then have another class to describe rigged objects, and then you might have a particular tree which would inherit certain things from the fact that it's just a tree at all, other properties from being a pine tree, and then have specializations on top of that that are due to where it is specifically in the world or what its purpose is and its unique properties. These are the kinds of things that are very important to make it flexible for creating content. You want to be able to make changes in one place that affect a lot of different things together based on what the relationships are. USD is also able to scale very well to work to quite large data sets. So, it includes features for lazy loading, includes different variants of the same object, the same asset which can be used for

level of detail and the variety of other purposes. It is very nicely extensible, so you can add new schemas, you can add new input output formats and a variety of other ways that you can extend the library itself. One of the very key properties that it has is that it's based on layering.

And this is really important to maintain the ability to offer things with the full richness that you want in offering representation. These layers are in some ways analogous to layers in Photoshop. So, basically, each layer has a collection of properties and values for those properties that are referred to as 'opinions.' And the layers are in a stack in terms of their strength and the strongest layer wins out for the value of the property. These layers are all composed together dynamically to give you the final result. So, different people can control things on different layers and you can collaborate at different objects together from a variety of different sources. So, USD has most of the things we think are necessary for the basis of an open metaverse. There are a few things that are missing and we've been working hard to try to fill those missing holes. One thing that has been missing from USD, as it was originally distributed by Pixar, is that it was designed for their purposes which are primarily not interactive. Their main purpose was to facilitate film development. And so, we've done a lot of working with USD to make it support very fast updates. And we've managed to get the speed of those updates to the point where we have been able to base our autonomous driving simulator on a USD core.

The other thing is that I mentioned that there was a requirement for the foundation of an open metaverse to be incremental. And so we built a library on top of the USD library where you can do anything you want through the USD API. You can create objects, you can create relationships, you can create lights and measures and move them around and do whatever it is that you want. And it keeps a record essentially of what has changed. And so at any point, you can ask it to synchronize to a central database server or another user and construct the difference between what used to be and what is now. And that's the thing that's required at its core for whatever representation we want to be the shared representation of an open metaverse.

The next thing that USD doesn’t build out of the box is materials. And in part that's because Pixar didn't want to force any particular material choice on anybody. So, we've done work with a material library developed in media called MDL (Material Definition Language), which we've open sourced. And we spent a lot of time making sure that

there's a very good integration between the MDL library, allowing it to provide very high-quality materials, as you see here in this picture, and for those to be fully interchangeable, fully open/available to all.

USD itself is pretty much entirely declarative. So, it states that, you know, there's this kind of object that has this property that lives over here, but there's very little in the USD that defines behavior. So we've been developing a procedural modeling tool called Omnigraph, which is layered on top of USD and makes it possible to specify procedural generation of geometry or procedural behavior. And lastly, USD does not include any way natively to specify physical simulation. So, we have open sourced PhysX, the NVIDIA physical simulator, which is used in many game engines. We've opened sourced it and we've done an integration with USD. And we've worked with Apple and Pixar to specify in USD the required physical parameters so that we can create a realistic virtual world in a way which is universally understandable.

So with these tools, with USD at the center, how do we create an open metaverse? What does it look like? Well, the center is going to be some kind of database. On the edge, there are different ways to connect into it. You have a variety of different client applications. On the other end, you have a variety of different ways to experience the content. We want to be able to experience it in AR, in VR, from a desktop, in whatever the best way to connect to this virtual world is for a particular person. And to enable the kinds of experiences that we want, we want to have a variety of different technologies to connect in. So, we try to dig a little bit deeper. I can show you, now, our implementation of these ideas, which is inside of what we call NVIDIA's Omniverse.

So, we built a central server, a database server we call the nucleus server. And the nucleus server is the one standard of truth, it describes the entire metaverse - virtual world in USD. It operates by a publish and subscribe mechanism, so every different piece of USD, every layer, is independently published and subscribed to inside of the nucleus server. And we've built bidirectional connectors to a variety of different client applications that you can see on the left from Maya, Max, Blender, Unreel, Revit, et cetera, which allow those applications to synchronize their internal representation with the nucleus server by sending only what changes back and forth. We've created a framework where we can buy services to the virtual world nucleus server, micro services that could do format conversions if necessary, search for different kinds of

applications, do offline rendering, levels of detail, a variety of different services that you could imagine. And then, we have a process which can look at the virtual world, subscribe to it and render the results either to a workstation, to VR, to AR, inter Web browser or whatever is appropriate for a given user.

So, how does this work if you have an existing client application? Well, what we have typically done is we've taken like this the application Scene Graph and we've introduced plug ins. So in the plug in, we bring in the USD library and we try to mirror between the USD library and the internal representation of the application. Typically, these applications have APIs that allow us to be notified when something happens to the Scene Graph. For example, somebody hit some button in the user interface and creates something. When that happens, we create the same corresponding object in the USD library. And we also go the other direction. If there's a change to the USD library, we then mirror it into the Scene Graph of the client application. And on top of the USD library, we have something we call the Omniverse Connect Library and that allows the application to connect into one of these nucleus database servers and create subscriptions and publish various things. And inside of that library is this technology that I mentioned that keeps track of what's changed and only sends the differences. So, what happens is when everything is connected live, somebody in application 1 can hit a button, can move something, and make any change that they want, changes in their scene graph that changes to the USD library inside of the plug in in that application that then gets published as a diff to the nucleus server. Those changes can get broadcast to all of the applications that subscribe. Then, for example, in client application 2, the Omniverse Connect Library would send an update to the USD library, make the appropriate changes and those changes would be mirrored into the scene graph of the second application.

So here, I'm going to show you the omniverse at work where everything is being synchronized with USD. We have three applications on the left. We have Unreal Engine 4. We have Adobe Substance Painter and we have Autodesk Maya. And then on the right, we have Omniverse Create doing real time path-tracing using RTX Renderer. So, here you see the users simultaneously making updates. The Maya user at the bottom is doing modelling, which plays to Maya's strength. The Substance user in the middle is painting textures. You see those live updated on the right. And then at the top left, you have Unreal being used for layout. And notice that they're not seeing necessarily all the

same things. So, the Maya users only working on one particular asset, the table. And all the other pieces are being put together. So you can subscribe to just a piece or to the entirety of any scene.

And we believe that a true 3D web shouldn't really be reduced to the least common denominator. And what's convinced us is that we've been very successful at doing real-time rendering. And I'm going to show this to you in a moment where we can render from this nucleus database and stream it to a mobile device. You can stream it to a webpage. We can stream it XR, whether nviit's VR or AR. And by doing that, it allows us to really have a lot of freedom and see quite high quality as we venture into virtual experiences which can be shared and be part of the real world or completely independent of it.

So, here's an example, where we created this virtual world on the moon and it's been driven at the moment by AI Pose Estimation. So, we have a camera looking at the user, the camera is creating a skeleton. That skeleton is being described in the USD, it's being published to the nucleus server. It's then being used to drive this model of the astronaut. And now we see somebody with a tablet, the same virtual scene. Uses a camera and AR. So we're using our AR kit to export a camera there. And you may have noticed, it quickly went by, there is somebody at the desktop. He's actually moving a flag around. All these users interacting with the scene at the same time. And here, we take it a little bit further. So, those people are actually physically in the same room with each other. This works just as well as people far away. Here, we have user in St. Louis and one user in Santa Clara. And now they have XR motion suits on for XR motion capture. And even though they're in completely different places, they're able to be part of the same scene. They see each other and they can speak to each other, they can interact. Obviously, they could be fully engaged in VR suits if they wanted to, but we have a wide variety of different options.

So, here's another scene. We're going to see the power of the universal representation. We've taken the scene described in USD, we’re path-tracing it in real time. We've created parameters in the USD that describe the physical properties. We've now brought in a VR user. He's represented by this avatar, and he can interact freely with this three dimensional virtual world as if it were real. Knocking it around and you'll see, in a moment, he will take the rocking horse, come up to, play around with it as if it's a

real physical entity. And we've completely standardized the representation of the physical properties and provided an open source simulator that can be used as part of anybody's simulation of a metaverse.

As I mentioned, USD represents things in layers and those layers can be independent and composited on top of each other as 3D properties. And this is just one example of how you can use this composition to build up a character. You can start with an initial sculpt, for example, down in ZBrush. You can add props on another layer that were created. For example, in 3D Studio Max. You can add lighting to the scene using, for example, Omniverse Create, which is our Omniverse native application. You can add materials to it on a separate layer, for example, using Adobe Substance. And you can put it in the scene and the scene can be built up of different pieces that came from all kinds of different places as well. Each of these layers is addressed by a URL. So, none of these things need to be local. They can come from anywhere, just like on the 2D Web. And just as you can create 2D Web pages that bring in pieces from all over the Internet, you can do the same here with a nucleus server and with our view of the of the open metaverse.

So, our idea here is to put USD at the center. To help and extend USD and as much of a standardized way as we can. To be based on open standards so nobody has to be afraid that their data will be captive or that they'll be at somebody else's mercy. And in doing so, we think that we can drive a path where everybody comes together around an adequate collection of open standards that allow a very rich variety of experiences that go far beyond gaming, far beyond some of the more restricted applications that many people have talked about and do things which are which are useful for industry, that are practical for getting real work done, and that also provide the ability to create community and real deep, important human interactions. So, we invite people to look at USD, to come and tell us what they think it needs to cover their needs as well and try to help standardize that and make that an open addition to the USD ecosystem. So, thank you very much.

Yongtaek HONG: Thank you, Dr. Kass, for pointing out the importance of open standards and providing a comprehensive overview of how USD can be extended to meet the need of an open metaverse. Now, I will introduce our second plenary speaker, Dr. Douglas Lanman. He is the director of Display Systems Research at Facebook

Reality Labs, where he leads investigations into advanced display and imaging technologies for augmented and virtual reality. His prior research has focused on head mounted displays, glasses, free 3D displays, light field cameras and active illumination for a 3D reconstruction and interaction. Today, his lecture title is 'How to Pass the Visual Turing Test.' Dr. Lanman, please.

Douglas LANMAN: Thank you for the introduction. So, again, my name is Douglas Lanman and I'm coming to you from Facebook Reality Labs. For the last 10 years of my career, I've had a singular focus, which is in a sense to create the ultimate display, at least the ultimate display within my lifetime. And what I mean by that is something that is as immersive, as realistic as the human visual system is able to perceive. And so, for the last seven years of this mission, I've been joined by a team I now lead at FRL called Display Systems Research. And our goal in creating this ultimate display is so that we can experience anything visual, anything captured from the real world or created synthetically with the full fidelity possible. And so for those of you in the audience, I encourage you to think about the question I ask every member from an intern to a very experienced professor who joins my team. What is the ultimate display? So, we're talking about the metaverse at this event. Ultimately, the output matters. How do we view this, right? Is it a tablet? Is it a cell phone? All of these things exist and have value. But what drives me in my career is to find the ultimate display, the one that is most compatible with the human visual system. And so Ivan Sutherland actually launched many of us on the path we are on with AR and VR. His answer in 1965, a very short memo I recommend you all read, I'm just going to read an excerpt here. He found that the ultimate display would of course be a room within which the computer could control the existence of matter itself. So he aimed high from the very beginning. Control matter and its full generality. A chair in such a room would be good enough to sit in handcuffs in such a room would be confining. A bullet displayed in such a room would be fatal and with appropriate programming such a display could literally be the Wonderland in which Alice walked. This is not the metaverse we speak of at the moment. The metaverse we speak of as a digital representation where embodiment is primarily audio visual with some light haptics. The degree of haptics Ivan imagines is beyond the reach potentially of our lifetime. And so Ivan took a step back, as we all are, even decades later, and tried to imagine what would be the ultimate display. And so in 1965, he essentially built what we're still working on, the first head mounted display. So, it's interesting that his thought experiment naturally led to a room. To control matter itself, you probably need a

lot of infrastructure. But to control photons, even in 1965, Ivan did not truly need a room. You needed a room for the compute and the tracking system, but he only needed a head mounted display. So, here it is, the world's first full 6 DoF (Degree of Freedom) tracked head mounted display. And so yeah, it's a room, but it's not a room, it's just a headset. So I think that's telling that to create something that could deliver the full 6 DoF fidelity of the human visual system. Ivan actually did begin with a headset. And it makes some sense that if you are trying to create a stereoscopic display, one display per eye placed close to the eye is sort of the logical first architecture you could imagine that would work. But we took a detour for a very long time in our field. The primary academic community I published and it's SIGGRAPH, and when I began my career more than 10 years ago, the state of the art virtual reality systems were Ivan's rooms. Cave virtual environments. Shutter glasses, combined with high end projectors would make the most immersive display you could possibly experience 20 or 30 years ago. But it's still a room. How many rooms are we going to have in our world of the future? I think this vision of a cave is not compatible with where computing evolved. It's personal, it's portable, it's not infrastructure heavy. Not all of us have the square footage to even create such a room, but we still want what Ivan wanted. We want to display that we can experience anything on, at least anything visual. And so that leads to really what's driving my career and my team. We step back and we say, what can we do to pass the Visual Turing Test. All the way from capture through rendering, transmission, compression, reconstruction and ultimately display - end to end, what is necessary to make and recreate reality accurately through the human visual system. That's the goal.

How do we get there? Well, first, let's take a step back. Where are we at? How close are we to creating the ultimate display? Well, those of you in the audience have a major role in creating the displays of the world. Since 1965, by my estimate, this is what we got, everything from cinema through televisions, monitors, laptops, tablets, phones and now wearables. This is the ultimate destiny of all computer graphics. It's going to be displayed on one of these two dimensional surfaces. I think this is really what drives me in my career. I grew up playing video games. I'm very passionate about interactive storytelling. A flat, two dimensional canvas is not what humans are ultimately built for. We view a stereoscopic world that is much wider, brighter, higher contrast and more immersive than any of these devices. But the world does want these things, and it's for a good reason. When I was just beginning high school, every computer for the most part was a desktop, at best. And portability became the driving factor of display technology

over the last 30 years. And also being personal. We no longer share our mainframes with everyone. We have a personal computer that is portable and that's what one out. But in the process we gave up, for those of us who care about computer graphics, nearly everything. We gave up immersion. Of course, we'd love to have a portable cinema screen, but how could you build such a thing? A cave is not intrinsically portable. And so we've had to trade off portability and personalization for immersion. This is really, I think in some sense, the best expression. I couldn't say it 10 years ago, but this is why I started working on displays. Computer graphics wasn't the problem. It was the output device. So, let's go to the left. Let's look at emersion. The devices we have are generally cinema. All of us have experienced cinema. Even those of us working in technology, very small percentages have actually experienced a good VR Cave and until very recently, almost no one had experienced VR. And hopefully now, you can get why at least in my lifetime, I think the answer of what is the ultimate display has to be a wearable display. And here's why. There's a singularity all the way to the left here. All of a sudden, when we go past a cave to a VR headset, we have something very bizarre happen where we loop around and we have a portable and personal device. Again, it's not a shared mainframe. It's not a heavy installation device. It's something I can take anywhere and experience anything with. And it's personal, which has a display engineer means every photon I create can be directed to your retina. No wasted light in theory. And the other magic trick is even compared to a cave, I think this has the real chance of realism by having a controlled lighting environment, good 6 DoF tracking. We can hope to reach things we've never achieved in displays. Faithful dynamic range to reality, wide color gametes, retinal resolution across the full human field of view. This is why, at least to the left, certainly VR and AR is the most immersive display we'll experience in our lifetime. But then going back, on the right, something special also happens that if you think about AR, AR is both personal and portable. So, we can do all the things we want to display engineer. Every photon once again can fill our vision. Now, we can go to a coffee shop and have an immersive desktop system and not be forced to work on small tablets or small laptops. Clearly, this sort of defines the endpoints of displays, at least as I see it. And then the last part, the Turing test. Because we have a stereoscopic display on both ends of the spectrum, we can actually achieve true visual realism. So, there you have it. I don't think this will change. The story is sort of set, at least for my career.

AV, VR are the ultimate displays. So, the question now is just what do we work on? What are the actual key challenges? And I think in some sense, starting to begin another 10 years of my career. I don't think we should do things like this, to be honest. Early in my career, I look at one individual thing at a time. How do we get high dynamic range? How do we get a small form factor? Over this talk, I'll share some recent work from my team that tries to push the dimensions of the visual experience in AR and VR, but jointly. It's not enough to find one magic trick. You have to find them all together to make a better display system that pulls all the visual axes forward. So, a better way to start on research is from first principles. Where we act.

If you buy my story so far, VR is the ultimate display at the moment. So, what's wrong with it or what limits it? Well, it's obvious. Take it out of the box, weigh it. It's around 500 grams. Obviously, we'd like that to be sub-100 grams to be more like eyeglasses. And it's an order of magnitude thicker than eyeglasses. And so, just from the get go, we know that form factor and comfort are key driving factors to make not just the ultimate visual experience, but an ergonomic one that you actually want to replace other display devices with. So these are sort of the five driving factors of comfort I see for display researchers like me. We have to make these things as thin and as light and as comfortable as eyeglasses - form factor. Accommodation is a problem that no other display had to solve because it's already correctly focused. A two dimensional plane presented an actual physical distance, has no vergence accommodation conflict. It's only once you have a three dimensional display where accommodation becomes a core research challenge for display technology. Also, everywhere else in the spectrum, we have a direct view display device. It's only at the ends where we start using viewing optics. And so, that brings a new challenge that we did not need to solve, except for in projection displays, which are optical distortions. Things like pupil swim, where the distortion depends on the position of the eye. And then new factors. We want to wear these devices for hours a day. We have to solve prescription in a good form factor. And finally, we want to bridge both augmented, mixed and virtual reality. We need good passthrough. We need to see the environment and the environment needs to see us. So these, again, are natural human devices that support natural interaction. If we can do all these things, now, we get to play the display engineer game. We can look into one of these headsets and we can see what's wrong with them, what needs to be worked on.

So VR, we're chasing that ultimate VR Cave. Super immersive, wide fields of view, but rather low resolutions because it's just a pixel game at the moment. You can never have enough pixels, it seems. And then on the right, in AR, we're constrained by form factor. If you believe that only glasses are acceptable, then within modern waveguide technology we have limited fields of view and quite non-uniform imagery compared to a direct view display. So, these become the second pillar of challenges you need to solve, in a sense, to create the ultimate display for the metaverse. Resolution. First and foremost, every other display in our life is 20-20 visual acuity, but not so in AR and VR. Our field of view. Again, my passion is to create the ultimate display, which means the full human field of view, or at least the full field of view supported by eyewear. Full human dynamic range, full color gamut of reality. If we're talking about AR, as much as people want to believe AR is the future, it has a major problem. It can't subtract light from reality. And so, solving occlusion is going to be a key roadblock to achieving parity with the contrast of direct view displays. And then, of course, passthrough I mentioned earlier is not just about comfort, it is about realism. The most realistic content you're going to get is reality itself. Capturing the environment you're in, passing it through that the full display stack and recreating reality is the necessary first step for emmersion.

And if you can do this, why did we do it? At least for me, the answer is, of course, everything is what I haven't said. Everything we could imagine would be made possible by such displays. But what drives me, even before this pandemic, is the idea of eliminating distance. Right? If we can actually feel present together. When I sit down on my couch, a relative from the other side of the earth could be there with me and I really believe it, really believe they're physically present. That they occlude the light from the couch, the shadows are cast correctly. This would transform, you know. It would eliminate distance. We could work from anywhere. We could be with our loved ones no matter our separation. This is a grandiose vision. It's not the full vision of the metaverse, but it's what drives me personally because it does not exist. That degree of telecommunication has never existed in my life. That's why I get on airplanes. All right, so hopefully at this point, at least you can understand where I'm coming from as a researcher. We have an ultimate display. We know the problems.

Now, it's just a question of you have a finite number of years in your career. What do you work on next? And the long answer is all of this together jointly and so all of these topics I've lectured on over the last seven years and even before we're working on all of

these topics. But this is a rather short presentation today. So, I'll give you some highlights and then connect the threads at the end. And I'll emphasize work I haven't presented as recently. It's almost more interesting to hear the news.

So, let's start with my personal feeling personally, the thing I feel we have to work on. Think about it for a moment. Which one do you think I'm going to choose? Which one would you choose? I think it has to be form factor. This is not everyone's choice. But again, if we're really creating the ultimate display, I want it to be something you might choose instead of your laptop, instead of your cell phone. It has to be so light and so comfortable that you'd wear it for hours a day. This is a challenge I think that's worth worth more attention than others. And so, it's what I actually started on. The first AR/VR product or research project I ever worked on was this one in 2013 which tried to address form factor. And so, we found that by shrinking and tiling the viewing optics for a VR headset, we could get to those eyeglasses I think are so necessary. The way we did it is not particularly important. It was just to show it's possible that VR is not this bulky shoebox. It has the potential to actually be not a headset nor goggles, but actual real deal glasses. So here's what it looked like at the time. If you look into it, decent image, at least for a rather new industrial researcher. Decent image, but nowhere close to 20-20 visual acuity. And so this was my first taste of the cruel trade space of optical science. That, yes, we we shrunk the form factor, but we gave up 20-20 visual acuity and perhaps with a high enough resolution screen, just the right microlens array, we might might get it. Or diffraction might actually put a wall on what we can achieve. And so, this helped me know for the next eight years after I did this project, it helped set a path for me and my team to avoid clever ideas that that find a dead end. I don't know if this is truly a dead end, but at least for now, it has been. It got us form factor, but gave up too much of the visual experience in getting there.

So, here is attempt number 2, eight years later at trying to create a pair of glasses. Here's where we're at a few years ago, five years ago with the Oculus Rift. This is what my team put together. There's a lot of caveats here. We gutted the electronics. This is just the bare minimum. It is the viewing optic and a panel, driver electronics would need to be external. But we believe this is possible. And so it achieves the same eyeglasses-like form factor, but it does it without giving up so much. And so how we did this was we used a very old optical technology known as a holographic optical element. And this wasn't something that was really on my radar eight years ago. And we combine that

with polarization optical folding. And in a moment you'll see that that actually allowed us to take a modern VR architecture without many changes and actually make glasses. And so, it means we didn't compromise so much. We retain the field of view, we retain the resolution. It's not really a different architecture. It's just a different variant. So, here's what the lens looks like. It's just a holographic film that's been carefully exposed. You can think of it as having, in a sense, the hologram of a lens. And here's how it works. So, on the left, we have what most VR systems have today, a refractive singlet lens with a big, long back focal length separated by displays. This might be somewhere around five centimeters or a little less thick in a modern VR headset. To get that down to sub-centimeter, first, we folded it. This is a concept that many people are exploring, known as polarization folding. It's more than 30 years old at this point, but that gets you a third of the thickness sort of instantly. You can fold by a factor of three. And then at that point, the thickness of a pancake type polarization folded lens is entirely driven by the thickness of the panel and the thickness of the lens. So, by flattening the lens into a hologram, we take something that's many millimeters thick down to hundreds of microns or less. And then we're just down to the display in the air being the thickest elements. Here's what are we are viewing optics looks like. So, again, just to make this clear, it's a display panel. Right on top of that is a holographic lens, an air gap, and then reflective polarizers. So, the lens is actually counterintuitively on the display, but optically close to your eye. And this achieves not quite the form factor of what I did eight years ago, but it achieves very high resolution, very high contrast. So, it's about three times thicker than what I did before but still on the order of eyeglasses. Here's what the prototype looks like. So, you can see, it's a proper VR headsets. It's a research prototype. This one is laser illuminated in green, but we have built full color prototypes.

So, now's the part where I try to step back and say, well, what did I learn in almost a decade of being a researcher in the field? Well, I learned you do have to cut with the grain. We have to find these ideas where everything moves together. And so, once we built this, no one actually commented on how thin and light it was. Everyone who saw this demo commented on how saturated the colors were. Because this is a laser illuminated headset, it slams the primaries all the way to the edge of the horseshoe. And so we have this very brilliant display. You've never probably none of you have experienced the laser illuminated display or very few of you. It's beautiful. And so, what that meant is we actually, without really going in with that plan, we figured out a high color gamut display at the same time as form factor. So, now you've got to ask how

many things did we also check off. And so very briefly, I'll explain, I think it's absolutely critical in addition to form factor, we need retinal resolution. If we're to read text, which is the primary content most of us consume, it has to be at retinal resolution. 20-20 visual acuity. Now, those of you who work in the display industry, these numbers are not the right numbers. The right numbers are pixels per degree. And so, where the industry wants to get to is 60 pixels per degree. Where we're at today is this red zone, 10 ish pixels per degree. We're seeing emerging products and concepts that get us in the green middle zone. But many optical concepts actually struggle to create an optical transfer function that's sharp enough to resolve a 60 pixel per degree underlying display.

And so this is just one example. We've worked on many high resolution headsets on my team, but this is a variant of what I just showed you, where we replace the outer reflective polarizer with a polarization volume hologram. So, we added the second holographic optical element that happens to be polarization sensitive. And if you place a chrome mask beneath it, that's the equivalent of a 5K display, something that doesn't routinely exist now but could easily exist in the near future. It can actually resolve this 5K display, not just straight ahead, but even on the far periphery. And so this is an experimental result. But if you look at the bar chart here, even quite far off the optical axis, probably far enough that your eyes normally wouldn't gaze that angle, it's actually resolving a one arcminute display. And so again, this starts to feel like a next generation display system. We get form factor, we get resolution, we get color gamut.

So where do we go from here? And so I think where we go from here is to think about the limits of eyeglasses themselves. And that would be field of view. So, let's take a moment to think about this, but change gears for this presentation. So we've been talking a lot about VR. AR, of course, has an equal stake in the future all the way on the right hand side of my original figure. And there, I think the key challenge is field of view. Right? If you look at modern wave guides, do they have enough field of view to give you a good visual experience? Remember, one of my driving things is telepresence. Do we have enough field of view to even look at a natural one to one scale human in an AR device? Are we going to have to shrink our virtual humans down and start losing some of the benefits of a three dimensional spatial computing device? So this is a very simple question. It is one of the first ones I ask along my path of joining Facebook Reality Labs is just, is it enough? Right? Of course, I know what the road maps are in the industry. I

know what fields of view are possible, but I don't really care. I want to know if any of those fields are going to be enough. And so let me tell you a different take. Again, sort of eight years in my mission of studying augmented virtual reality, rather than worrying so much about publications, I started to worry about the experience. And so, here's a little thought experiment I did. This is probably four years old at this point. Are sunglasses even enough? The field of sunglasses, people always say, oh, glasses are enough. The field of view of sunglasses is not actually as large as you think. So one of my colleagues, Brian Wheelwright, took a pair of sunglasses outdoors and he measured the field of view. Now, the first thing, I think, is you need to ask an optical scientists. Many individuals report field of view in degrees. It's the wrong unit because it doesn't give you a good sense. If you double the number of degrees, does it mean you really doubled the immersion? The correct unit, those of you who have a physics background, is of course the one my colleague Brian made, which is steradian, solid angle. So if you measure the solid angle binocularly of a pair of standard sunglasses, you get about 3.3 steradians. Human vision is actually 5.3 steradians binocularly, depending on how you count it. Which means that sunglasses, the ultimate glasses, if you will, are only 63% of the human field of view. And so if you're really trying to create the ultimate immersive device, it already means that the form factor of sunglasses will limit you. But for AR, I think we'd accept these need to be fashionable devices. So let's just go with sunglasses and accept that the one true metric for AR is what percent of the solid angle is your device. 100% is sunglasses. We'd all agree it's not the ultimate VR immersion but is the ultimate AR one. And I did this, eight years ago right out of my postdoc to understand field of view. You get these cardboard tubes and you cut them and you walk around the lab the whole day. You go and get a coffee, you talk to coworkers, you try to edit a document, and most of the time you try to convince yourself that the very low AR field of view are enough. It's going to be OK. I can edit a document. I never felt satisfied by that because it wasn't actually a time machine of the experience, it didn't give you your peripheral field of view. So, this is something we built four years ago that I haven't quite seen built before, even though it's obvious. We took a pair of shutter glasses, like you'd have in a VR cave. We added motion capture balls to them. And if you think about just some off the shelf components, first it gives you 100% the field of view, nearly 100% the field of view of sunglasses. If you have good projectors and you stand a meter or two away, you can easily get quite close to retinal resolution and you can wear this all day. It's 48 grams. No problem. I can work an eight hour day in it. So we built this into, I think, the first real time machine for AR glasses. So, here's what we built again. Starts

off just like a VR Cave. We have shutter glasses tracked by motion capture. And up above, it's like one wall of a VR caves, we have a high resolution projector. But the key difference (again, I haven't quite seen it, which is a surprise to me when we built it. I thought many people had done this before) replace the wall of a VR cave with a low cost, translucent screen like you'd have in theatrical projection. The reason to do this is now you have a real AR experience. Rather than using VR glasses to emulate AR, where you have latency, contrast loss, resolution loss. Reality is real, the augmentations are augmented and you have the full field of view of eyeglasses. So, we call this an AR Cave and I recommend those of you working in the field, it's easy to put one of these together. It takes a few weeks. And you actually get to experience the ultimate AR display. I'm probably the only one in the audience who's actually experience full AR like I'll have 20 years from now. 100% of sunglasses' field of view. I've read a Google document that at retinal resolution, I check the weather, read the news. It's an interesting experience that you can actually get rid of your monitor and have this virtual monitor.

And so now you can ask the question, this is why I built this with my team. It was to answer a simple question, what field of view is needed for telepresence? And so, here you can see remember, the one true metric is percent of sunglasses' field of view, not degrees. Most AR devices on the market today are about 5-10% field of view of sunglasses. If we had 20% field of view of sunglasses, that would be a monumental leap in our display technology. And so look at the green box. That's beyond state of the art or beyond known path of state of the art AR. And I have two co-workers standing with me and I'm just looking at what field of view would be needed. And you'll see, just trying to have a natural conversation with two individuals, which someday will be remote. You can't quite keep them framed, certainly not at the current state of the art and even at a much larger AR field of view. And so, I think this goes to what visual experience is, what user experiences we have in AR if we can't have multiple participants framed at natural scale? This makes a design challenge, not a research challenge. And so, just to nail home the point, this is the demo I gave everyone in the lab. So here, I actually emulate the future AR telepresence call, where I have two participants that are here, just some basic computer graphics. You're viewing them at 20%. So, a little beyond state of the art AR field of view. And you can see you can see one individual when you're looking at them, but you're actually missing the second individual in the conversation. So, you're losing those emotional cues, the body

language, how they react to the other individual. So, this this helps to really inspire me. This simple little demo that took a few weeks to build helps you know that field of view is a core challenge of AR. We all knew it, but now we can start actually running studies and understanding what is enough. So, that's that's the AR cave.

Let me take a few more moments and tell you one more challenge in AR. And I think this is even more fundamental and even harder to solve than field of view. Sure, 20%, 50%, 100% of sunglasses, incredibly hard for those of you working on waveguide and related technologies. But occlusion is the one that makes the right hand and left hand side of my diagram collapse. If AR glasses could remove light from reality, there is no difference. VR and AR are one in the same, other than the form factor. It's either glasses or goggles or something else. But if you can subtract light and add light, you've created the ultimate XR device. So, how hard is occlusion? How hard is it to subtract light? Well, first of all, just like field of view, most people convince themselves it's going to be OK. We'll just crank up the brightness. So, this is a cloudy day in Seattle and we emulated a weather app, just trying to overlay say you're taking a walk and the overlay ends up where it is.

It's quite hard to see. And of course, it's this struggle. You know, you can boost the brightness, but at some point, you're going to either flood the eye with so much light, they dilate or you just can't quite win there. What you want to do, of course, is remove light. And there's a lot of work. Many of you in the audience have worked on this topic. Roughly, I'd classify this into three buckets. In the middle, is the obvious solution, which is an optical relay. The problem here is we want to put a spatial light modulator out in the world, but we can't really afford to push around a two meter screen everywhere we go. So, we have to relay some small spatial light modulator, like an LCD optically to the world, which means you end up with these periscope constructions. Even if you use the clever holographic optics we showed earlier, you probably will struggle to create an optical relay that can put a spatial light modulator conjugate to a distant plane in a small eyeglasses-like form factor. So I think the relayed ones are great as time machines. You can build a pair of binoculars, you can walk around and see what occlusion looks like in AR. I recommend you do that. That's very interesting. You could go to the light fields, like I worked on eight to ten years ago. Problem with that is once again you're making an optical relay by tiling. And so, that introduces for AR a lot of limitations on field of view, on resolution. Diffraction is going to cause artifacts for the see through

environment. It won't look like eyeglasses. And so I think this path academically is fascinating. It doesn't feel like a practical path we're actually on. And so, what we're left with is the most obvious idea that almost anyone comes up with in five minutes of discussion, which is the idea of local dimming. Simply take whatever your AR display is and slap a spatial light modulator directly to the front and take what you can get because it might be better than nothing. So let's talk about this one for a moment. I think it's a pretty common idea in the field. But this event for a larger audience. So the idea would be we can create our augmentation using any device, off-axis combiner, waveguide, whatever it is, right after that device on the world side put a liquid crystal or some other spatial light modulator, a transparency in front of the world and that will give you local occlusion. So, your pupil has a non-zero size. So, you're going to get this blurry occlusion mask. Probably won't help you pass the ultimate Visual Turing Test because you can't get a sharp edged boundaries. But you could certainly make a plaintive text, a virtual monitor, much higher contrast. And so, again, these are these are one of those things I've learned you talk about in the hallways at a conference. Everyone knows the idea, but no one's seen it. Just sort of like the AR Cave I mentioned earlier. Everyone can imagine what they are is no one's actually seen it. So, I think this is, again, 10 years into this, a key thing I identify we need to do in our community, which is demo more. Build these time machines that are actually good enough that you learn more than a hallway conversation can intuit. So, here you go. We have a sort of mockup for those of you just to remind you what we're doing. We have a combiner, whether that's a waveguide or some other device creating a virtual image say 2 meters. Then right in front of it, we have a sharp occluder that becomes effectively a blurry occluder on the world side. And so, we've built a bunch of prototypes of this. But again, I haven't shared this one publicly before. So again, this is a different take on things. It's not really publishable research, but it's the first time I ever experienced this. We built an off-axis combiner system so we could have a very wide AR field of view, so much higher than the 20% we saw earlier. So very immersive AR. We added accommodation technologies so we have a varifocal system in this. And then we added an occlusion mask. We didn't worry about making this look pretty. It was all about getting to experience it. And so, hopefully after this pandemic, people can visit our lab and we can do some demos. But this is, to my knowledge, the first local dimming demo ever built with the complete 6 FoD headset. So you can see we're just drawing the simple initials of the lab. But this is what you would get without local dimming. This is what you do with. And so all of a sudden, these hallways speculations of like, "I don't

know, I think it'll be OK" or like, "No, you won't possibly be good enough. You should just globally dim everything or nothing." You end those debates and you simply see it and decide whether designers can use this concept for value or not. So, there you have it. I'm going to wrap things up. I would love to go for another hour and tell you about all of our aspects. But you can find those all over the Internet in various places. I'm very proud that FRL has given me a home for the last seven years to explore all these things.

As a research scientist, I have to be honest that I'm increasingly I'm trying to find a path forward where all these things can be moved together. But the truth is, we haven't found it as an industry. We have things that are are quite established in VR, like a refractive singlet architecture or a pancake lens architecture. And in AR, we have things like waveguides. But pushing beyond that has proven very difficult. And so that's really the core challenge anyone's working in this field is we are making the ultimate display device. It only matters if we move everything or nearly everything together. But what we have done as a community since Ivan Sutherland started us in 1965, we found hundreds of building blocks, from holographic optical elements through light field displays. We have so many concepts, we don't lack for those. We're finding new ones every day. Again, I think I'm trying to encourage our community that what we need next are new architectures. Things like the holographic optical concept where multiple axes can move forward together. And the other thing, I think we should be more open to demos of old ideas. Much of what I spoke about is an old idea. I didn't speak about it today, but I've spent many years working on varifocal displays, retinal resolution displays, high resolution insets, AR testbeds. All of these things are obvious in some sense, but they need to be done and they haven't all been done. I haven't done all the work. And so I really think being more open, even academically to time machines, perceptual testbeds, this is how we can end debates and start moving in a coherent direction towards the ultimate display, which is, of course, AR and VR. Thank you.

Yongtaek HONG: Dr. Lanman. Thank you for sharing your thoughts on the importance of the ultimate display for AR/VR system and for reviewing challenges of wearable AR/VR display system and Facebook's efforts to solve these technical challenges. Thank you. Now, our third speaker is Professor Byoungho Lee. I think he can also touch on technical challenges and solutions in the optical side. Professor Byoungho Lee is a Professor of Electrical and Computer Engineering at Seoul National University.

Professor Lee is a fellow of IEEE, SPIE, OSA, and SID. Professor Lee also serves as a senior member of the National Academy of Engineering of Korea and a member of the Korean Academy of Science and Technology. Today, Professor Lee will be discussing 'For Real Impact of Metaverse.' Please, Professor Lee.

Keynote Session

Byungho LEE: Thank you, Professor Hong. I'm Byungho Lee at Seoul National University, and it is my great pleasure and honor to give a talk today. The title of my presentation is "For Real Impact of Metaverse." This is the outline of my presentation. After a short introduction of the Metaverse, let me talk about why it is drawing attention and what improvements are needed for real social impact. And then I'll talk about optical challenges, current optical technologies and critical issues for AR/VR headsets, and advanced technologies for AR/VR. And then I will conclude my talk.

Most of you are familiar with the terminology such as augmented reality or virtual reality. The computer generated virtual environment provides us a good opportunity for virtual reality. And here's the real world. In the case of augmented reality, most environments are real-world-based and only partial environments are artificial from the computer. There's a different level of mixture between the real world and virtual world. It can be called the mixed reality. And actually in the mixed reality, interaction is very important. And there's another terminology called the extended reality or XR. In XR, also interaction is more stressed. X can be an A, a V or an M. Recently, the Metaverse is a hot issue, and as Michael Kass mentioned, the terminology of the Metaverse first appeared in the novel by Neal Stephensen, although the concept was suggested quite earlier. The Metaverse is coined from the word "meta" and "universe." "Meta" means "beyond." So if it can be considered as a kind of collective virtual shared space, and the virtual environment and the physical reality are linked and shared by the Internet.

There are unique characteristics of the Metaverse, which are also the cause for drawing much attention very recently. You have almost no limitation in action and no limitation in the number of participants. And users can create their own avatars and environments and make them interact with each other. More than that, the created the world can be saved in the physically sustainable virtual space. Very famous examples of such Metaverse platforms are Roblox, Zepetto, and Decentraland. Also, the arbitrary

environment can be created and can be applied to various purposes, for example, an entrance or a graduation ceremony for universities, enterprises, global marketing, entertainment, and also some celebrity concerts. There are many such applications. Monetizing is also very interesting and fascinating in the Metaverse. Global enterprises collaborate with Metaverse platforms and release virtual products. For example, Gucci is cooperating with Metaverse platforms and we can purchase Gucci suits or accessories, and then my avatar can put on the Gucci suits with small money. With the development of blockchain technologies and non-fungible tokens, which are usually called the NFTs, users can earn digital currency or purchase virtual products or even virtual land. On the platform of Decentraland, we can purchase the virtual land. Actually, they limit the whole area of the virtual land, so the price of the land is increasing as in the real-world situation.

Of course, COVID-19 affects quite much on the Metaverse. Social distancing is ongoing worldwide due to the wide spread of the COVID-19 and depression due to the pandemic and the lack of social interactions accelerated the adoption of the Metaverse. Also the importance of individuals' hobbies is increasing, which also affects the Metaverse.

Advances in graphic technologies are also important, as explained by Michael Kass. Entertainment enterprises produce high-quality virtual content that can be loaded on Metaverse platforms and realistic movements of objects and human characters make the Metaverse immersive. The Metaverse is a very hot topic, especially in Korea. These are examples of searching results with the term Metaverse on Naver or Daum. There are lots of news on Metaverse every day. But there is also a limitation now because most users are teenagers except for some events utilizing the Metaverse such as graduation ceremonies, educations, or entertainment concerts. So actually there are some people who actually suspect whether the Metaverse can have a real social impact or just a bubble. Actually, many years ago, there was a kind of platform called Second Life, which is still alive. It's like the early version of the Metaverse, but it did not make a great success. So, there should be a kind of driving force for the new trend of the Metaverse for it not to follow previous examples.

Then what improvements are needed for the real social impact of the Metaverse? First contents are mainly focused on games or short-term events now, so those contents are still being consumed by only limited people of certain age groups. So sustainable

contents for daily uses or killer applications fulfilling public needs should be developed. These are some examples of works by Korean companies, which make photorealistic rendering. They are wonderful but actually, it gives much burden on computing and data transfer. So there could be some agreement depending on applications. In some cases, we really need the photorealistic image rendering. But for other applications high quality rendering might cause some difficulty of widespread applications or popularity.

There's another issue of high-speed rendering and data transfer. If I use this kind of head-mounted displays for VR entertainment, if I move my head, actually the sensor should sense the movement of my head and then the image should be rendered in real time within the response time of the nervous system. Typically, current VR headsets have the average latency of 30-50 milliseconds, but it should be 20 ms or less than 20 ms, or even 10 ms. So, latency results in visual fatigue or nausea. So it is a very critical issue. And regarding the computational load, actually the computation can be done on server, but then there should be high-speed communication technology connecting the device and the server computers. For example, 5G and 6G high-speed communication should be used.

There is an issue of ethics or securities such as privacy violations, virtual property robbery, hacking and data sniffing, sexual or racial harassment, and commercial fraud. Of course, many of them are typical issues on the Internet. But there could be new legal issues, such as the copyrights over creation in the Metaverse, because users can make their own avatars and create facilities and buildings on the Metaverse. There could be some copyright issues. Also, AI-assisted creation is possible in some Metaverse platforms, so copyrights to be a legal issue.

Another critical issue is the AR/VR devices, and current AR/VR devices still have issues of bulky form factor, as Douglass Lanman explained, narrow field of view and eye box, and also their focus cue is limited. These kinds of limitations hinder the immersive Metaverse experience. So actually, many characteristics or properties are full of trade-offs and there are lots of challenges, as Doug explained.

Now, let me talk about the optical challenges in VR/AR devices. For the virtual reality headset, this is a typical structure and here's the display panel and here's the lens. The separation between the lens and the display panel is smaller than focal length of the

lens. Then the image is floating at a far distance with magnification like this. Also we have two eyes. So there are two panels and the images overlap like this. We usually provide stereo images for the left eye and the right eye, and this is the overlap region. Due to this optical characteristics, actually there should be some space i. Although the space between the lens and the display panel is an empty space, we cannot arbitrarily reduce the space. So, basically the virtual reality headset becomes bulky. These are well-known VR headsets, and they have a quite good field of view, and these are the penal resolutions, but usually they are bulky.

For AR devices, actually, it is more difficult to make the system compared with VR devices. That is because for AR devices, the outside light should come in as well as the light from the display panel. So this is a typical, simple example of the structure. And here is the display panel and the light is going in this direction. And there is the concave mirror. The reflected light goes into the eye after being reflected again at the half-mirror. The outside light is coming through the half-mirror and goes into the eye. But this structure has a small field of view and there are diverse versions of these kinds of AR devices. This is the case off-axis imaging and the light guide. In this example, the image is projected from this direction and it is guided by total reflection and there are partial mirrors embedded in the light guide and the light is partially reflected and goes into the eye. In this case, the wave guide is thinner than the light guide, several millimeters typically, and the image is projected to from here and it is deflected like this, and with total internal reflection, it propagates through the wave guide and goes into the eye. Here are very interesting devices called the in-coupler and the out-coupler, and the light is deflected or defracted in this direction, and also the light is defracted in this direction. Those kinds of devices can be implemented with defractive optical elements or surface relief gratings (SRGs) or it can be also implemented with holographic optical elements as Douglass Lanman explained. But anyway, there are lots of technologies, but they have some advantages, but also they have some disadvantages.

These are very well-known examples of AR headsets. Microsoft Hololens has a relatively smaller field of view compared with VR headsets. Generally speaking, the field of view of AR devices is smaller than VR devices because for VR devices, we can block the outside light, so we can easily make a wideview lens in front of our eyes and then the field of view can be enlarged. But for AR devices, such a case will distort the light

from outside view. So the design is more difficult and this kind of specification is not as good as VR devices.

There are many key performance factors. Actually Doug explained these factors, and more than that, he explained many important factors. So I will briefly mention these kinds of factors. The field of view is important. The form factor or the size of the device or the weight of the device are also important. There are some efforts to use a metal surface to flat lens for AR devices. The eye box is also an issue. In the general case, there is an eye box denoted as the blue box here and the images can be observed correctly within this eye box. And here's a box, and here is the eye pupil. If the eye pupil is moving out of the eye box, the images cannot be observed. So it is important to enlarge the eye box, but to enlarge the eye box, the field of view or resolution can be sacrificed. So there are many efforts, for example, to move the eye box, with the sensing of the eye pupil location. This is an example of such cases, although the eye pupil is moving, the eye box is following the location of the eye pupil and the correct images can be observed.

The image quality is also important. To minimize the aberration is also important. The aberration can also cause fatigue in observation. So the focus cue is important. In stereo imaging, the left eye image and the right eye image are separately provided, and then the floating image is intended to float here. Then, we observe the images. So this is the vergence distance and converging angle between the two eyes. But for the two eyes to observe clear images, actually they look at the screen or floating images from the head-mounted displays. So there's a difference between the vergence distance and the accommodation distance, which causes fatigue. So this is a very important factor to be resolved.

So generally speaking, for the VR/AR headsets, these are typically state-of-the-art specifications of the devices. The weights are 100-500g or more. The thickness is 3-5 cm or even larger. The depth of the 3D image is fixed at a far distance of 2 meters typically. So the ideal desired VR/AR glasses could have these kinds of properties. First, the weight should be very small, like the usual, typical glasses. The thickness should be small. For example, the resolution should be higher and the contrast should be much higher than the current status. It is desirable that the 3D depths can vary from 30 cm to infinity, for example. The price is also a very important factor. Also the battery

life is very important. Typically, the battery life now is 2 hours or less. It is better to have a longer life and high brightness is also a very critical issue. Typical AR devices can be used only indoor, but for outdoor applications, we should have higher brightness.

Now, let me talk about a few examples of advanced VR technologies, mostly the work from my lab and in this case, we are using focus-tunable lens and then we change the focal length. Here the display image for one image frame. Actually, what we want to do is that we want to implement 80 depths of image planes. So for a single image plane actually we illuminate back-light for different depths, over the 80 layers of depth in synchronization with the focal length change of the lens. So we can do that in real time if we are using micrometer devices which can operate very fast. So these are the experimental results. The videos are taken at 1 m and 20 cm. So in that case, actually, the conflict between vergence and the accommodation can be minimized. Here is a thin VR structure. Doug also explained very wonderful devices. We also modified the devices and are using the light path folding and lenslet arrays to further decrease the gap. We made a device with 3.3 mm gap or 8 mm total thickness. We introduce our new compact VR optics design, Lenslet VR. Lenslet VR structure is a glasses-like form factor. At the same time, it has a great optical performance for VR, such as field of view, eye box, and eye relief. Let me introduce the principle and demonstration results. Currently, our devices have a form factor of a headset. Ironically, most volume of the conventional single lens system is occupied by the air. This thick air gap between the display panel and the lens is only necessary to secure the optical pathways. In order to reduce the gap, we utilize the lenslet array. Since the lenslet array can have a much shorter focal length than the single lens, we can reduce the thickness of the gap to 10 mm. Additionally, we also utilized the optical path folding systems by inserting a few polarization controlling films inside, we can make the light bounce back and forth.

Let me show you the experimental results.

The eye box size of our prototype is 8.8 mm by 8.8 mm. As the camera position moves, we can observe a fine continuous image within the eye box.

Now we are within the eye box, but now we are moving out of the eye box. Anyway, we were successful in making very thin VR devices. There are lots of research going on in my group and other groups, especially at Stanford University, Facebook Reality Labs,

MIT, and so on. So in this case, we are trying to correct the aberrations of lens or errors in the refractive power of the devices. So there are lots of works being done on the image processing for these kinds of applications. And also there are lots of research on holographic displays. The eyewear devices can also have holographic images rather than just stereo images. If we use holographic images, the vergence-accommodation mismatch can be minimized. But usually, the calculation time or images rendering time of holographic images is quite huge. So there are lots of research to adopt AI technology or Deep Learning technologies in hologram generation.

So now let me conclude my thought. There are several key factors for the Metaverse, which are the diverse and sustainable contents. Those kinds of contents should be generated more and more for the popular applications of Metaverse platforms. High-speed communication is needed not only for the low latency algorithm, but also for large computational load on server. There are ethics and security issues. Also, I think the final issue will be the optical breakthrough. So minimizing the visual fatigue, compact system size, and wide field of view and eye box for AR/VR devices will be needed for the wider popularization of the Metaverse. There are many IT companies who are very eagerly working on AR/VR devices and also the Metaverse. So I think the future of the Metaverse is bright and many more researchers are needed and many research funding is also needed as well. Thank you very much for your attention.

Yongtaek HONG: Thank you for touching on the societal implications of the Metaverse and for reviewing technical challenges and your achievement in optical engineering for a realistic metaverse. Our final keynote speaker today is Prof. Woontack Woo. Prof. Woontack Woo is a Professor and the head of the Graduate School of Culture Technology (GSCT) at Korea Advanced Institute of Science and Technology (KAIST). He also serves as the director of both CT Research Institute and KI-ITC Augmented Reality Research Center.

Woontack WOO: Morning, good afternoon, and good evening, first of all, I would like to thank the SK CHEY Institute for Advanced Studies for inviting me here. It is my pleasure to have a chance to share my experience and thoughts on the future direction of Metaverse. I will present the possibility of reality and virtuality convergence platforms in terms of the digital twin, metaverse, and extended reality. Today, I will talk about forests rather than trees, key concepts rather than the detailed individual technologies. Then I

will talk about the possible future directions. Some are already curious whether the policy is a bubble or not. Some ask me whether winter comes again. How can we make a metaverse sustainable? How can we make it socially beneficial?

It is pretty much pretty much an important issue. But the core message of my talk is the metaverse goes beyond the just phone play space. It is a real virtual convergence platform. At the same time, it is a social media and economic platform that can anyone can easily access and utilize in their daily life. So, let me explain what I have done during the last 20 years. Actually coined the terminology 'Ubiquitous VR' in 2001, but what is the 'Ubiquitous VR'? The basic concept of UVR is a virtual reality experience in smart physical space. It can be realized in the context of augmented reality. How is it possible? In order to achieve 'Ubiquitous VR', we need three key components: first, the 3D link between the real and virtual, and the augmentation of the context of interest.

We can augment real physical space with several different types of content. It includes information and knowledge. The last one is bi-directional interaction and collaboration for human-to-human communication. So UVR is a way to experience and utilize the metaverse in our daily life. Let me explain what I have done in the last 20 years and then what is the plan for the next 10 years. In the first 10 years, I did research on the realization of augmented reality in the ubiquitous environment. For this purpose. I collected a signal from various sensors and then analyze or interpret the signal, then augmented the physical space with the integrated information. In the last decade, I conducted research on how AR can be used to expand humans' capability in terms of humans' physical ability, intellectual ability, and social ability. Next 10 years, I will do research on how to expand the capabilities of our society beyond the people. It is closely related to the future direction of Metaverse. Nowadays, Jensen is pretty much a famous person in Korea because he claims "Metaverse is coming". I think the more accurate expression might be "Metaverse is coming back in Korea", at least in Korea, the metaverse is coming back. It is true because the Korean government already invested in Metaverse in the year 2000.

Since 2010, it has disappeared from the people's attention and then forgotten. He is mainly talking about the virtual world. However, the more important thing is how to utilize the Metaverse in our daily life. Let's see the interesting plots, this is the Google trend on the terminology on metaverse, as you can see in this plot when the Metaverse

community started the discussion in 2004, there are some activities. And then this year, many people showed some interest in this area. But the attention is decreasing nowadays all over the world. However, if we look at the phenomena in Korea, it is pretty much interesting. As you can see, in mid-2000, there were several spikes. Even the Korean government spends a lot of money in this direction. And then recently the attention is increasing every day. That is the interesting point. I will skip to the other plot. So recently there are so many opinions about the metaverse in Korea, but the VR community claims the second life is the origin of the metaverse. Recently, Infinite Office, supporting office work without going to the company is regarded as the representative example of the metaverse. The gaming community Claims the Roblox or Fortnite are examples of the Metaverse. Some define the Metaverse as a virtual world experience through the Avatars. Some called the Metaverse a virtual space where the virtual assets can be traded. Some consider the Metaverse is Web 3.0. Then what's the Metaverse? There are various opinions and it might be the story of an elephant touched by a blind person.

So how do you understand the elephant?. The topic I want to introduce today is how to understand the elephant. OK. In order to understand the relationship between extended reality and the Metaverse, we need to understand the reality and virtuality continuum. This one is a well-known continuum proposed by Paul Milgram at Toronto university. We are living in physical reality. And then we can generate a virtualized reality; nowadays we call it Digital Twin. If a real-world and the corresponding virtualized reality exist, then we can merge two together. If we emphasize reality, that can be augmented reality. If we emphasize virtualized reality, that is augmented virtuality. Also, we can remove what we don't want to see in physical space; that is a diminished reality. The same thing can happen in virtualized reality. Diminished virtuality might be possible. Also, we can mediate the physical object to the other information or space that is called mediated reality. The mediated reality is somehow linked to the metaverse. You know, the vertical direction in this figure, the augmented reality and augmented virtuality form the mixed reality. The mixed reality can be extended to the vertical direction. Based on the humans' imagination and creativity, this space can be experienced and explored through avatars. I define core as extended reality or ubiquitous VR. This is a well-known fact that the terminology metaverse was introduced by the novel Snow Crash.

That is not exact; this novel did not define the terminology, just introduced the concept. The first community defined the Metaverse as the future internet, a kind of a shared and open standard-based virtual space. Since then, the ASF has defined the Metaverse roadmap. In that document, the Metaverse is the convergence of virtually-enhanced physical reality and physically persistent virtual space. The Metaverse has four key components. The left side is the physical space, the right side is society. The upper side is an augmentation of information. The bottom side is a simulation. By combining these four components, we can make interesting UX scenarios: lifelogging, virtual world, the mirror world. Augmented reality is some component of the metaverse. OK. So the first version of Metaverse was introduced in the novel, then in the middle, as I explained in the previous slide in the mid-2000, the metaverse roadmap was introduced. Nowadays it comes back and is it possible to be used in our daily life? Still, there is a big question. In order to make this new concept in our daily life, there are several issues we have to solve. How can we use this kind of a new concept in our daily life? There might be several possible scenarios. One might be using extended reality or augmented reality in physical space. In order to achieve that scenario, we may need glasses as introduced in previous presenters.

By the way, is it possible for the wearable Metaverse to reopen the digital gold rush nowadays? In order to pave the way to the digital gold rush era, the Metaverse ecosystem must be well established in advance so that the reality & virtuality convergence space becomes both a social platform and thus an economic platform. To realize such a platform, it is necessary to build and utilize integrated infrastructure such as IoT, 5G, Data, metaverse, Non-Fungible Tokens, or extended reality. Then, in order to make a sustainable metaverse, the integrated platform must be open, reliable, and interoperable. At the same time, it is necessary for us to prepare for various social issues, economic issues, and legal issues as explained by prof. Lee. Then how can we use this Metaverse ecosystem in our daily lives? Augmented the city might be a good candidate. The augmented city is a smart city that improves our daily lives by organically integrating ICT technologies based on the Digital Twin, and then manage and solve city problems through the citizens' participation. So in order to achieve such an augmented city, there are several core elements. The first is the organic linkage between reality and the virtualized reality. We can build virtualized reality using the Digital Twin, 3D augmentation reflecting context as well as a location. In order to use

the information in real life, the information needs to be provided based on the location. Location is a good example of context.

We need more intelligence according to the user's request. Also, you have to handle users' intentions. The other important thing is connecting people through this Metaverse. So bidirectional interaction and collaboration are pretty much important. The last one is a participatory problem-solving platform. What can we do using this kind of platform, especially the augmented city platform? First of all, the augmented city platform can be used for city planning, monitoring, and operations. Moreover, it can provide the citizens with a new way of experiencing, communicating, and collaborating with the people in the Metaverse. This will allow us to extend the capabilities of the citizen, cities, and then the society. How is it possible? Originally, this idea was proposed to the Busan EDC project. We can put the IoT sensors here and there in the city through the high-speed network, neighboring Edge and Cloud can collect signals. AI can interpret such a signal to generate information and knowledge. Such information and knowledge can be visualized in Digital Twin using such information and knowledge, we can simulate in the Metaverse, then the result can be feed to the city through the AR and VR platform. If the citizen has wearable glasses, the citizen can fully utilize the augmented city, as you can see nowadays, high-speed networks, especially in Korea, SK and KT provide a fast network between the device and nearing Edge and Cloud.

So, the XR glasses can provide important information for the user to make a smart decision. Such information, sometimes processed in the glasses, sometimes processed in neighboring Edge and Cloud, everything can be done on the fly through the high-speed network. Through this process, citizen ability can be expanded in various ways. There are several possible applications. First of all, learning is a pretty much important component. Working, playing, shopping, even managing the city is a good example of metaverse applications. So what if we apply this idea to the new learning systems? We can build learning space on the Metaverse. In this space, there are no borders, no grades, no classes, no teachers in this new space. We just can connect and communicate and collaborate to learn new knowledge there. Then, we can apply the knowledge to solve the problems we encountered in the real world. There are several technical challenges. There are issues we have to solve. I will skip that deep explanation on the details. It can be applied to many other areas: museums, concerts, parks, and gyms. Those things also can be a good application. Even if we cannot go to

the museum, still we can enjoy the museum in the metaverse. The concert is the same. This will drastically change the whole life for our whole lives.

It extends humans' ability in various sences, especially we can expand humans' body ability. Basically Metaverse is beyond the space and time, so we can move our body using the avartars. Also, we can support brain activities. The knowledge can be provided just in time to make a proper decision, even Metaverse can help to connect the people so we can extend the social relationship ability of the human. It transforms the way we live and interact with the world. Even it revolutionize the industry in terms of productivity, efficiency, and safety. There are so many more different applications. Still, there are several challenges. In order to make the Metaverse ecosystem, there are several key components: video avartars, content authoring, Digital Twin, AI, metaverse, extended reality platform, wearable UI, networks, wearable devices, even industrial ecosystem. Still, there are several issues we have to solve. 3 Presenteres already explained such technical issues. So, the remaining question is: how can we make open, sustainable reality-virtuality convergence platform, but that requires to provide the standardization of data Especially data expression, is pretty much important for the internetworking. Internetworking is a heterogeneous platform. It is another important issue for the seamless XR experience. Another important issue is the economic activity; how can we connect online to offline bidirectional economic activity? Is it possible to exchange the digital asset across the different platforms? Also, the data right is another important issue.

Through these complicated issues, we can build virtualized cities and then we can make several different layers of Metaverse. By applying these kind of things, we can experience several different things in the metaverse. So what are going to be happening in the augmented city? If we can realize that kind of things, we are living in physical space, we can make Digital Twin, just copy physical world to the digital world, then we can do several different simulation there. Then what's going to happen in the augmented city? First question is, this kind of technology improves quality of life in the city? Are the citizens of the city feel happy? So we have to think about how to make a balance between smart city and happy citizen. If we don't consider such kind issues, humans just use technologies and the machines learning, getting smarter and smarter and smarter every day. So in order to make happy citizens, we may need to change the metrics. Evaluation metrics, in terms of the technology, in the viewpoint of the

technology, we always try to make it efficient, to improve the efficiency. We try to make a useful solutions, usable solutions.

However, such a kind of solution may not be good for the human, human beings. Sometimes, inefficient solutions need to be selected for the human. Sometimes, inconvenient solution might be selected for the human. Sometimes, unintelligent suggestion might be needed for the human. If we include human beings in the loop, the. the measurement should be different. So we are developing interesting technologies every day, technologies improve every day. Then why we do the research, why we do the development? Why do we make better technologies? The ultimate goal is to make people happy. In that viewpoint, we can use Metaverse, which we have to think about how we can use Metaverse for the human beings. So there are several challenges. We have to make Metaverse sustainable. Many people worry about the bubble. As I showed in the previous slide, more than 15 years ago in Korea, Korean goverment spent a lot of money on that direction. OK, this year, Korean goverment is planning several research activities for next coming years. And we have to learn from the previous history. We have to try to make a sustainable metaverse. In order to make a sustainable Metaverse, we have to chip in thinking about the Metaverse. Many people thought Metaverse is the fun place. Place for the games. This is good for the game. However, we have to think about how can we make it useful for the human beings. In order to use Metaverse for the human beings, we have to think about Metaverse is the social platform, especially using the reality-virtuality convergence. This is also the economy platform. We can do several interesting activity there. Also, you can use Metaverse for the problem-solving testbed, as I explained, the main mediation can be occured on the Digital Twin. Digital Twin is just a copy of physical world. Then, by collecting the signal from the physical world, you can do several interesting simulation there. So if we encounter any social issues, then we can simulate such issues in the Metaverse, many citizen can participate in the Metaverse to find out proper solutions. In order to achieve such goals, we have to think about Metaverse is Social Overhead Capital. So government need to invest the money for that purpose. Government need to support innovation, need to open ecosystem, and prepare the economic problem there. Still, there are remaining homeworks in this area? How can you make Metaverse for social good? How can we make it sustainable? How can we make people smile in that space? New challenge will have mistake and a failure, and said that, even so those who never give up and take on the challenge to change the world for the better. So I hope

we can take on the challenges together to build a new Metaverse with a reality-virtuality convergence platform. This is it for the presentation. Thank you.

Yongtaek HONG: Thank you, Professor Woo for your comprehensive lecture of current status and the future direction of the metaverse technology. Now, we will move into a discussion session with our four brilliant panelists. The audience that pre-registered for this Webinar has left many interesting questions. Based on these questions, I have selected a few questions to be discussed.

Discussion Session

Yongtaek HONG: I think Professor Byoungho Lee mentioned a little bit about this. Our real world has become increasingly digitized during the Covid-19 pandemic. How is the pandemic changed to the scope/future of the metaverse business and what are other important driving forces behind its sudden popularity and growth? Who can start give a comment and maybe very short comments should be fine.

Byoungho Lee: Because you mentioned my name.

Yongtaek HONG: Ok, yeah, I saw in your slide that mentioned about covid-19.

Byoungho Lee: As you said, the current covid-19 situation has functioned as a kind of catalyst for rapid expansion of the base of metaverse. The tendency to seek social interactions in the metaverse space has expanded as the number of real world meetings decreases. But another aspect is that there exist many difficulties in the real world, especially under the current economic situation. While people can do many things in the virtual world in metaverse. Besides the covid-19 situation, I think there are other factors driving the popularity of metaverse. The strategy of metaverse platforms such as Roblox or Zepeto was good. There is no entrance barrier, so it is easy and free to join and their business model is also good. The users can create own avatars and decorate them with, for example, Gucci costumes or accessories. And another example is that users can make their own games on Roblox and and others can use the games as well. And there's another thing. There is a tendency or trend among young people to freely express their thoughts and individuality becomes very important. I think that kind of trend is also a driving force for the metaverse ecosystem. And also technically, a

computer graphic's power has improved significantly in recent years, which is also a very important factor.

Yongtaek HONG: Ok, thank you. Dr. Kass?

Michael KASS: Yeah, I think that to has certainly created a situation where it's more important than ever for people who are far away to be able to virtually be in the same place, to work on the same virtual designs, virtual assets, virtual items. And it becomes even more urgent for people to get their tools to work seamlessly with each other across that distance. And so, what we've tried to build with Omniverse is a way to integrate everything across those barriers by turning all the assets, all the representations into a common and open form. And so, we have had just an explosion of interest from all manner of different sectors of the economy and different organizations across culture and and different purposes who want to work with each other in a seamless way. And so, we're trying to build that out as fast as we can.

Yongtaek HONG: Ok, Dr. Lanman, do you have some comments on that?

Douglas LANMAN: Yeah, I certainly see the pandemic as what it is, which is a step change across many aspects of our society. But in this regard, in terms of sort of bringing about the metaverse earlier, right? There's there's our professional lives, right? Where we present our work and we share the context of what we're doing with peers that really understand the context. But all of us know the truth of it is like when you have a holiday, you see your relatives and they ask you, "Hey, what do you do?" And for many years I've been working on AR and VR. And I would tell the same story that one day this this thing I put on my face will somehow let you live anywhere and be in a workplace distributed across the entire globe. My relatives would always sort of roll their eyes like, "You love science fiction. Yeah, maybe someday." But now with the pandemic, those of us who are lucky enough to be able to work from home, largely like tech workers and information workers, we have lived it now. Right? And that that probably would have taken a decade or more to come about naturally. Major companies are now allowing work from home where they would not have contemplated it before. And so, we ran the experiment and largely it was successful. Right? I can just from my small vantage point of leading a research team, we somehow made do. Right? Even though our hardware driven team, we actually worked effectively remotely. And now, it's

no longer the science fiction story, it's you actually get to the real part of it. We did the demo and now it's like, what are we missing right now? Right? I'm missing a trip to Korea where I could meet everyone in person and I could see the amazing city and go to restaurants with you. And that part we can't replicate in a virtual reality headset. Right? We can't recreate a full restaurant experience. There are things we can't do. But there are there things we can now do and we all know it. And so I think in many ways it it's convinced me like what I was saying is not just a good story. And it does seem like the truth of perhaps this is a way we can reduce our carbon footprint. We don't need to be on planes as much when it matters. We need to, but we can work productively remotely, which means we can open up economic opportunities to individuals that don't live in capital cities. I mean, much of what I tried to sell my relatives is coming true. And I think it did pull it in by a decade or more. You know, we'd still be talking about the metaverse for all the reasons Professor Lee mentioned. There are many factors of society moving that way. But the pandemic definitely accelerated the part I'm connected to. And it's perhaps one of the few silver linings to it is is this tool is now being more actively developed than ever.

Yongtaek HONG: Professor Woo?

Woontack WOO: I also agree on others opinion. I also have a similar opinion on that issue. And then we accelerated the widespread use of the metaverse in Korea. The more important the driving force might be the digital transformation technology matured. The hardware and software technology progressed. And also new generation might be a driving force. They are born in digital. They are pretty much familiar with the digital technologies. They are good at using the smartphones. So currently if we only focus on the game area, especially Roblox or Fortnite in Korea, the kids are the main driving force. 50s and 60s, they don't care about Fortnite and Roblox. Nowadays many people try to use Zepeto in Korea, but they just gave up pretty soon. The generation gap is also big driving force making the metaverse ubiquitous in Korea.

Yongtaek HONG: Thank you. If I can expand a little bit, then it seems like the metaverse is currently developing into a global business based on your response. And what commercialization strategy should be combined with this platform business and how do you envision the business model for the metaverse? So I think this one, maybe Facebook has some business models for that. You can start.

Douglas LANMAN: But I don't presume to represent all of Facebook. I'm a display engineer. I can say personally for myself, Professor Lee mentioned several things I find interesting that are very tiny. It's like saying, what is the business model for the Internet? I mean, it's a new platform. The metaverse is a 3D Internet in some sense. And so we'll see an infinite number of businesses and the evolutions of the 2D versions we already have. But things that I think are interesting Professor Lee mentioned, that content creation has evolved. Small anecdote is one of my family traditions is I build a video game for my children every birthday and the fact that I can do that as one individual and make an entire 3-5 hour video game over the course of many evenings, that was not possible 20 years ago. And so the content creation tools are phenomenal. And some of the things we saw Dr. Doctor present that if we can really create the metaverse as we were sort of discussing, there is new business opportunity. Again, this is a tiny little thing. But I personally find it amazing that an artist could create an element that goes out across the metaverse and make a real living. And, you know, just as someone who does make video games for a hobby, I now understand why you might pay 40 US dollars for a single model of a single chair. It's incredibly hard work that deserves recognition. And so this this idea of sculpting of artwork having intrinsic value, I think we'll see this idea continue to evolve. And this is not even half a fraction of a percent of what the metaverse is. But it's something I'm excited about because I grew up playing with video games and to see this democratization where everyone is now a creator. The tools are so powerful even I, who is not a great artist, can create a whole universe and share it. This is exciting, especially when it becomes this shared experience across the entire globe. So again, that's not the Facebook business answer. That's just my personal - what I want to see come about, because I think it's fascinating.

Yongtaek HONG: Dr. Kass, do you have any comments on the business models or other side predictions?

Michael KASS: There are a lot of different business models potentially at work here. And, you know, what did we see with the Internet explosion early on? It was that complete interoperability and connectivity and free flow of data creates enormous value. And that's all you can be unlocked in a wide variety of different ways. So we think that's really the first thing is to standardize on the protocols, standardize on the interchange and make sure that that extends to as many domains as possible. And once that's the

case, then if you have a service to provide and you connect into the open metaverse, you now connect it in to everybody. You don't have to create a different version of that service for every possible user. And that just creates enormous economic efficiencies. The other thing and one of our areas of focus is what I think both Professor Lee and Professor Woo have discussed, which is the idea of digital twins. We think that that's going to be transformative, that you have a piece of infrastructure and you want to understand where it's simulated over its lifetime. In the past, you built a building. You may have used CAD (Computer-Aided Design and Drafting) tools for that. After it was built, you pretty much threw away the CAD model. But today things are dynamic. You want to understand how that building is operating. For example, if it's a factory, it's going to be reconfigured many times and you want to optimize the use of the factory. So all of those are done best in simulation and in a simulation that can bring together information from a wide variety of different sources. And that's one of the big sort of motivations for building Omniverse and for creating these open standards for everybody to connect into.

Yongtaek HONG: Professor Woo, Dr. Kass mentioned the digital twins and you talked about the digital twins and also the future of the metaverse. Do you have any comments on business models?

Woontack WOO: Actually I have no idea about business model. But in the viewpoint of a mediation between the physical world and virtual world, digital twin is a good platform to connecting both worlds. If we consider the metaverse as a platform, the platform itself is a business model, like Facebook or YouTube. Many people will get together and then there might be several similar activity there. Metaverse itself might be a business model. That is my guess.

Yongtaek HONG: And Professor Lee, do you have any comment on that?

Byoungho Lee: Because the metaverse is an ecosystem, I think the hurdle of accessability should be low. So with the most current metaverse platforms, a strategy to expand the base for users by providing a free platform seems to be a good strategy. And in addition, probably it is necessary to operate the server workstation by taking a small percentage of fee from the users through content purchase in the metaverse and by developing and releasing VR/AR devices with various specifications like the

smartphones. Various models of smartphones - the hardware companies might get the success in commercializing the devices. And the monetizing strategy, like the case of the Decentraland is also very interesting. They limit the total area in the virtual land, which is divided in sections and sold in virtual space. And the prices vary, maybe taxes in virtual estate might be a sustainable business model for the future world. It's really interesting to the users to keep their wealth or digital asset in the metaverse space.

Yongtaek HONG: Thank you. For global business, actually, one area of debate is whether a true metaverse can have a single operator or requires a heavily decentralized platform built upon community based standards and protocols. How do you foresee the interoperability of the metaverse? And if in the case, how plausible will it be to establish industry wide standards around topics such as data security, data, persistence and forward compatible code evolution? Dr. Kass, do you have any comments on standards in the future for the metaverse area?

Michael KASS: Yeah, absolutely. So, you know, we we know that there are a lot of players that have an interest in creating walled gardens and trying to meet everybody's needs and keep them inside. But, you know, as we've seen from the Web, there's tremendous value in getting to a situation where everything can talk to everything else. And we think that value is so large that it's inevitable that it will happen. And we're doing our best to support it. So we've taken a standard that exists, open sourced Pixar's USD. We've open sourced additional pieces that we think are critical to making that all work. We've tried to make sure that the nice protocols for updating those models in real time, we're making it available to individuals for free to try out and do whatever they want with it. And, you know, we're convinced that people who understand that this is not a walled garden, who can see that all the basic data formats and protocols are open, will be willing to invest in it because they're not worried that somebody is going to come along and charge a very large tax on it. And as a result, we think that's that's the way for it. And we invite other large players to come and join us in looking at what needs to be added to the ecosystem so that it can have even even broader applicability. But absolutely, our view is that the metaverse itself is far too big for anybody to own, for any one player to own. There will be specialized metaverses that have their own proprietary standards, but our greatest interest is in one based around open standards. And we're doing everything we can to to push that it

Yongtaek HONG: Thank you. Any comments from other experts about the standardization for the technology?

Byoungho Lee: I have a comment. It would be good to think of the current smartphone case. In the case of Apple, it has built its own operating system that is a compatible only with its own mobile phones. But Google's Android has created a platform that is compatible with smartphones of various companies with an open strategy. So the metaverse platform needs to follow a similar model of Android. Image rendering programs such as physically based rendering important in developing a realistic metaverse platforms. In the case of image engines, there are several representative engines that are widely used worldwide, such as Unity and Unreal engine. And they adopt an open strategy similar to Android. And the difficulty of creating a virtual environment is alleviated and the quality is improving. And we saw very impressive demonstrations of USD in Dr. Kass' talk. So, it is expected that more users will use those kind of engines. And since most AR/VR devices are already compatible with this kind of engine, the setting standards for security or compatibility shouldn't be difficult, I think. And platforms with unique compatibility standards, such as Apple, may may also emerge. But for that kind of case, platform development companies need to establish precise standards themselves.

Yongtaek HONG: Thank you. If I can change gears a little bit toward the general public, I also use movie clip on my presentation whenever I explain the paradigm shift of wearable device from accessory tie to textile and also body attachable, stretchable wearable devices, etc. It seems like it takes about 30 years when you actually see the technology in our real life based on the Hollywood movie. So the general public's most common conceptions of the metaverse stem from science fiction, such as the virtual worlds portrayed in Ready Player one and The Matrix. So how will our future, maybe ultimate form of metaverse, compare to these fantastical visions captured by science fiction authors? I thought I saw Professor Woo mentioned about red or green pill. Right? I don't know at this moment which one should I take. But can you have a comment on that? The extreme, ultimate form of metaverse in the future?

Yongtaek HONG: In my viewpoint, nowadays, there is no common definition of metaverse. Each one has a different definition of metaverse. Originally metaverse comes from the virtual world. But the more important thing is using such information or

knowledge or experience in the physical space. So as long as we use here such information or knowledge in physical space, that may not be the same as the movie. My main interest is how we make such new space in the physical space, not the virtual space. So I hope we can use in everyday life. In daily life, there are so many things can be distorted using the metaverse. We can simulate and then, we can use results on the fly in the physical space and make a new opportunity in physical space. So, physical world might be more fun and interesting place when compared to the virtual. Even if we make a virtual realistic, still that is virtual. So I hope we can make such fun thing in the real physical space. What do you think, NVIDIA?

Michael KASS: Yes, so I think that obviously Hollywood is drawn by good storytelling. So whatever makes a compelling story, whatever is going to sell tickets. That's that's what Hollywood wants to present. I think if you look back at the early days of the Internet, all of a sudden all these computers were being connected around the world. And what did Hollywood give us? They gave us a prediction of computers taking over, launching missiles and all kinds of doomsday scenarios. And what they didn't tell us is the very ordinary things that started to happen, like the fact that you can now buy airplane tickets by yourself very easily. And so, there's been a disruption in travel agents, for example. And, that's just one of thousands of examples of ways in which the connectivity of the Internet has changed how we do business. And I think that when three dimensional worlds are shared widely, all of a sudden we're going to find thousands of new opportunities that nobody can predict right now, where people start to work more effectively together and things get built. I think it's going to revolutionize architecture. You're going to be able to go from conceptual design through real simulation to figure out what you're actually going to get. You see now, where furniture companies are starting to help you visualize the furniture in your own home, but eventually that's going to be a full metaverse activity. So, you're going to walk into this virtual world that's going to include your home and a decorator is going to help you lay things out. And it's going to be a full 3D experience. And it's going to be much more convenient, much easier than it is today. And it's going to change the whole way that we do interior decorating and the way that we build architecture and the way we plan cities and the way we navigate and simulate potential ways to to get all kinds of things done. So, I'm not so concerned about the science fiction novels and the Hollywood movies that portray gloom and doom because, you know, that's their job. I think it's it's going to be immensely practical. And we're seeing our manufacturing partners tell us how much

it's going to help them to be able to do everything in simulation before they spend the money to make a mistake. And those efficiencies that's going to benefit everybody. All of a sudden we're going to be able to reduce the cost of all kinds of goods and services. So, I'm very excited about metaverse based future in a digital world.

Yongtaek HONG: Thank you. Doug, do you have any comment?

Douglas LANMAN: Yes. I mean, it used to be at least when I joined Oculus, it's still a startup, everyone got a copy of Ready Player One. And we like to read these things. And many of us grew up reading science fiction. So going to your question of which things may they get right or wrong? I mean, step away from the dystopia and the societal things. I'm a hardware driven person. So looking at their hardware, I think Ready Player One is the most mundane, but probably the most accurate in the sense that Ready Player Two, the Matrix, they both assume beyond the ultimate display I talked about, which is a brain machine interface and an invasive one at that in The Matrix. And this sort of idea that you could replace Ivan Sutherland's instrumented room with a direct neural interface is correct. We are just brains in a box. That's obviously, I think, nowhere on the horizon in our lifetimes. And so, if you go to the boring old Ready Player Ene, you have headsets, which we have today. You have limited haptics, maybe a suit if you're really into like extreme haptics. I think they they sort of got that arc. You know, you mentioned it being 20 years. My friend Mary Lou Jepsen says the same. Any new idea in the display industry - it's usually 20 years from the first paper to seeing a consumer product. And so, I think that's where Ready Player One perhaps did get it right. That we already have most of the things written in there and the metaverse we've been talking about is pretty well-imagined in there. I mean, it's geared heavily towards gaming and education, but I think that's actually a reasonable view of where things could easily get within 10 years. Ready Player Two, not so much.

Yongtaek HONG: Okay. Professor Lee, do you have some comments?

Byoungho Lee: Yeah, so actually the current popular metasurface platforms like Zepeto or Roblox or or Decentraland are quite different from the metasurface as shown in the movie of Ready Player One. In the movie of Ready Player One, they used to head-mount displays and earphones, but for the current popular metasurface platforms, we do not use such devices yet. So, I think finally we have to go to that direction of

using the headset devices because the more vivid interaction is needed. So firstly, it is essential to have a ultra high resolution display devices to provide very high quality and wide viewing 3D-depth images, which could be as natural as the real world. And when it is accompanied by advanced sensor technology and tracking technology, they can interact naturally in the virtual world. For example, in the movie of Ready Player One, the main character of the movie shakes hands with other characters in the metaverse. He can feel as real through sensors and stimuli which allows the metaverse to be like an immersive real world. So, currently, displays, sensor and tracking technologies fall short of the ideal technology is shown in the science fiction. But a lot of research is being done, so, I think it will be possible to produce such vivid metaverse in the not-so-far far future.

Yongtaek HONG: Ok, this is going to be the last quick question. The general public sees a lot of the emerging technology in the area. You guys are really experts in this field. How long do you think it takes before general public really feels that 'Oh, the metaverse is here' in our real life? In terms of, you know, you mentioned about a lot of challenges in the hardware, AR/VR headset, . probably computational abilities and also the communications performance. Right? So, please just take turns on your vision, how long does it take when you know that I feel (I'm not the expert in this area) that the metaverse is here around us. OK, maybe start with Doug?

Douglas LANMAN: It's an interesting question because, you know, there's that famous saying "However long you predict, it's always twice as long." In some sense, it's been here my whole life in many ways, right? It's in kind it's here. And it's just degrees we're talking about, right? I grew up with the Nintendo Entertainment System and that that brought a shared but not simultaneously distributed virtual environment to me and that evolved into a local area network gaming like Doom and Quake and that that kept evolving into the massive World of Warcraft and Fortnite and Minecraft and what have you. And so, I guess I'd argue we don't have to wait in that sense. And AR/VR is just the most immersive way and I do think it does something special, which is it achieves presence. I don't I don't think you can do that with other devices. But it's already here. The metaverse has been here for quite a while and it's in that fragmented form of piecewise walled gardens. But it's here. And I think we'll just see it continue every year to just sort of evolve the same way home entertainment, like the Nintendo entertainment system. Right? It's a direct line to today's consoles. They didn't have a step change at

any point. And I don't think the metaverse really will have a step change. I think AR and VR is the same. Right? You can get a headset today, you can put it on and you can see three dimensional graphics better than you've ever seen. And in some sense, you know, the things we talk about, especially Dr. Lee and I, the technical challenges, while important, we're steadily moving forward. And so, I don't think there'll be one day where we magically cross the line and say, "Oh, this is the metaverse." We may, in fact, look back and say, "Oh, this obscure video game evolved into the metaverse" or how it actually comes about, epecially the most prominent, widely used one, will be a unique and interesting story that probably hasn't started yet. But I guess my answer is it's already been here for quite a while.

Yongtaek HONG: Professor Woo?

Woontack WOO: Considering the hardware platform experience, experiencing the metaverse, glass might be a good one to experience the metaverse. Considering the VR glasses, Facebook Oculus has to make a big success nowadays. Samsung made by. I'm awaiting the Apple Glasses, which might appear sometime next year. The more important thing is AR glasses. When AR glasses make a big success that will really be the time for the metaverse.

Yongtaek HONG: Ok, thank you. Professor Lee?

Byoungho Lee: I think it depends on the activities of NVIDIA and Facebook. So, I'm an optics person. So let me talk about the AR/VR glass devices. So, the demand for AR/VR devices has been steadily increasing from the past. And I think the need for a AR/VR devices will increase significantly as contents become more diverse within the metaverse platform. However, the reason why the current penetration rate of AR/VR headset is not high is that there are many, many hurdles yet, such as a high price, big size and weight of the devices that Doug and I explained. But there are lots of efforts going on in this field. AR/VR devices are rapidly developing. Facebook acquired Oculus in 2014 and launched the VR platform service, Facebook Horizon, in 2019. And within the company, Facebook Reality Labs, which researches VR/AR devices, is investing a lot of money and manpower in research. I think Microsoft introduced its own augmented reality head mounted device, Hololens, and upgraded version Hololens 2 and collaborated with companies in various fields. Equity is also investing a lot of money in

the development of AR/VR display devices. So, improvements in the performance rate and price of AR/VR through investments by such big high tech companies can speed up the popularization of the metaverse to the general public, although I cannot tell when it would be.

Yongtaek HONG: Thank you. Dr. Kass?

Michael KASS: Yeah, so, I totally agree with what Dr. Lanman said. That, you know, the metaverse has been around in some sense and some very limited sense for a very long time. But I don't think that's the point at which it really impacts our society in a big way. And I'll take this measure, which is, you know, today, at least in the United States, to be a fully functioning member of society, you basically need to be able to access the web. For example, let's say all you want to do is schedule a COVID shot in the U.S.. Well, the way you interact with the world, the way you make that happen is with the 2D Web. And I'm not so concerned with exactly what devices we use to interact with our virtual worlds. For some people, it's going to be traditional screens because they just don't want to put anything on their heads. For some people, they're going to dive into VR headsets. For other people, it's going to be AR. But the transformation to me that's key is when everybody starts to build 3D representations of whatever it is that's important to them and that those 3D representations are all tied together with a common infrastructure. And that, I think, is going to start to happen probably within five years. It's going to be pretty widespread by 10 years. I think it's going to be everywhere that no matter what it is, you're to be in the 3D business. Right now, whatever it is that you do, you probably need to have a website and it's a 2D website. There's a future where whatever it is that you do, it needs to be part of the metaverse. Will that transformation be finished by 10 years? Probably not. But will it be really on its way to the point where your ability to function in society is almost requiring the ability to enter the metaverse at various times to just get things done? And I don't think that's very far away. The point when everybody is putting on VR devices and everybody's using these special display technologies, that may be further along. But I think we're we're entering a real inflection point. And you're going to see the foundations for the 3D representations of everything in a connected way, coming very, very soon and it's going to adoptions can be very, very quick, I believe.

Yongtaek HONG: Thank you. I feel like I should play more games, should watch more movies and should pay more attention to what's happening around me after today's webinar. To conclude today's webinar, I'd like to thank our speakers again for their engagement and candor. And thank you to our audience tuning in online. I hope that today's webinar clarified the curiosities you have about the emerging metaverse platform. And we hope to see you again in our upcoming Scientific Innovation Series. Thank you, everyone.