

# 4,320 hours in the metaverse: How the metaverse shapes our view of reality

Michael Bjorn

*Ericsson AB, Torshamnsgatan 21, 164 83 Stockholm, Sweden*

## Abstract

This paper argues that there are specific roles for AI to take in the metaverse. The paper adapts the descriptive, predictive, and prescriptive roles for AI presented in Shoshana Zuboff's book, *The Age of Surveillance Capitalism*, using augmented reality (AR) as the example. The primary descriptive role for AI is to identify surfaces in the physical world where digital objects can be placed. The predictive role is used to foresee the unfolding of events, such as a human walking towards a digital representation of a creature. Finally, prescriptive AI is used to maintain the illusion of reality also when there is insufficient sensor data, such as adding digital branches to a physical tree in the place where the digital creature went to avoid the approaching human. Whereas AI in the first step simply maps the world, we end with an AI that decides how we perceive reality.

## Keywords

AI, AR, reality

## 1. Introduction

Predictions of a future virtual metaverse are many. A parallel and dystopian out-of-body existence is often described, but there is much to be said for a more spatially anchored hybrid world where digital and physical are intertwined.

Join us on a journey into the innermost of this metaverse, where AI is increasingly taking on the role of creator. With their constantly connected sensors, all-seeing algorithms not only build a universal foundation, but also actively engage in all the details with the help of machine learning; from predicting what's happening in every moment to proactively shaping people's behaviors. A world where AI increasingly directs the local environment in your everyday life.

When the company Facebook changed its name to Meta, many people asked themselves what the metaverse really is. The word meta refers to something at a higher level of abstraction, and just like multiple physical worlds make up the universe, the metaverse could simply be taken to mean the sum of all virtual worlds. These multiple worlds don't necessarily have to be interconnected to be part of the metaverse, much like different apps

---

In: Kiemute Oyibo, Wenzhen Xu, Elena Vlahu-Gjorgievska (eds.): *The Adjunct Proceedings of the 19<sup>th</sup> International Conference on Persuasive Technology*, April 10, 2024, Wollongong, Australia  
EMAIL: [fitterstoke@gmail.com](mailto:fitterstoke@gmail.com) (Michael Bjorn)



© 2024 Copyright for this paper by its authors.  
Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).  
CEUR Workshop Proceedings (CEUR-WS.org)

don't need to be able to exchange data to be part of the internet. AI can exist without the metaverse, but while single virtual worlds can exist without AI, the metaverse in a broader perspective hardly can. The metaverse probably can't exist without there being people who want to use virtual worlds either, so let's start there. In my experience, many people imagine something quite dystopian when they hear about the metaverse. The typical human reaction is a slightly frightened look combined with a dismissive gesture and a few words about the SF movie *The Matrix* or possibly about *Ready Player One*. Both of these films depict a 3D virtual world that exists separately from a physical world that is in decline and where the participants walk around as avatars. In this respect, they are reminiscent of the scenario in Neal Stephenson's novel *Snow Crash* where the term metaverse was first used. But the history of virtual 3D worlds is so far really more about video games.

## 2. Brave new 3D game world

Virtual 3D worlds have been around since the early 80s, thanks to groundbreaking video games such as *Battlezone* and *Monster Maze*. Moreover, with environments like *Minecraft* and later *Roblox*, such 3D worlds have become quite mainstream. The fact that the ambition in these game environments is actually to create something so complex that it is a world in itself can be exemplified by *Minecraft*, where there are now universal Turing machines (machines that can be programmed like computers, you could, for example, build a *Minecraft* game within the *Minecraft* game) [1], and also ambitious attempts to build a model of our entire universe [2].

Another 3D virtual world with ambitions to be universal is, of course, Linden Labs' *Second Life* [3], which also tried to launch a virtual economy with its own currency, the Linden dollar, so that you could buy and sell virtual buildings and land properties. Sweden opened an official embassy there in 2007, but after initially being widely publicized, interest in *Second Life* died out quite quickly. Not very strange, you might think, because how fun is it to sit in front of a screen and control an avatar while trying to pretend that it's something more than a video game?

However, interest took off again when Luckey Palmer and his company *Oculus* launched the VR headset *Rift* on *Kickstarter* in 2012, promising a simple and high-quality way to explore these 3D worlds from the inside with the help of virtual reality, VR. Before the headset had hit the market, Facebook had bought the product for what was considered an astronomical sum. In 2016, Valve, which operates the world's largest computer gaming platform *Steam*, together with HTC, launched another VR headset, *Vive*, with improved technology for controllers that could be tracked freely in 3D space and thus the new wave of advanced consumer-oriented VR was here.

## 3. VR or AR?

Although VR technology gained quite a few admirers, many questioned the market potential. At the time of writing, I have spent exactly 4,320 hours in *SteamVR* and an estimated couple of thousand hours in various Facebook/*Meta* headsets, a total of maybe three working years.

Even so, I understand the criticism leveled at VR (and by extension at the metaverse). Allowing yourself to be completely surrounded by and immersed in a virtual world can be a claustrophobic experience. In addition, many people experience side effects similar to motion sickness because movements in the virtual world do not correspond to the movements that the body physically perceives. Even though I think VR is a lot of fun and don't mind describing myself as a VR nerd, I find it very difficult to see that VR will be as common a way to use the internet tomorrow as smartphones are today. On the other hand, I believe that augmented reality (AR) will become as commonplace. This is because you can see the physical world in AR all the time and can move freely, while VR restricts mobility to a minimum. Smartphones have made us accustomed to using the internet wherever we are, and AR offers the same freedom.

AR and VR relate to each other much like the old idiom of Muhammad and the mountain. That is, in AR, the mountain comes to Muhammad, but in VR, Muhammad must go to the mountain. You can clearly hear that it's easier for Muhammad to go to the mountain instead of the other way around, and VR is also technically less complicated. More concretely, in VR, we humans move into the virtual world while AR tries to move the virtual world to us, where we are in the physical space.

The idea of AR is that we should be able to experience, and ultimately also handle, digital objects in the same way as all the physical objects we have around us. One example is the mobile game Pokémon GO, which tries to make Pikachu and other digital creatures behave as if they were here in our midst. In a world where digital objects behave in exactly the same way as physical objects, we do not need to learn anything new to use the internet. Instead, the abilities that we learn by growing up in a physical world can also be applied to digital objects; we get a natural user interface (NUI). This is different from today's graphical user interface (GUI) which is based on metaphors where we talk about windows, menus, trash cans, and so on. Unfortunately, these items don't behave like physical windows, menus, or trash cans, so it's actually quite difficult to learn how to use today's tech gadgets. Moreover, many of the items we model GUIs belong to the previous century, and do not always have any meaning for people who were born in the current century. We sometimes use the term digital natives to describe those that are young enough to have grown up with the internet, but as long as we have a GUI based on broken metaphors that they have to make conscious effort to learn, they are in fact digital immigrants like all the rest of us. Only when there is an interface that is completely natural will there be true digital natives in this world.

Therefore, since both VR and AR are based on natural user interfaces, all people would potentially be digital natives regardless of the type of technology generation they were born into, if the technology were made widely available. However, since VR technology also has challenges with motion sickness and claustrophobia, a future metaverse is more likely to be based on AR instead. Today, when the metaverse is still in its infancy, we're building it on VR, but as the technical challenges of AR are overcome, we'll gradually move there. And the technology that, more than any other, will enable this shift is AI.

#### **4. Descriptive AI: The AR Scene floor layer**

It should be mentioned that today it is difficult to talk about any digital technology at all that does not use AI in one way or another, and not least such technology that has to do with image processing. Additionally, video games make extensive use of AI for avatars, and those interested in algorithmically generated 3D environments should study movies like *The Lord of the Rings*. Here, I therefore focus on AI that is more or less unique to AR. This is not a difficult limitation because AI is actually required for there to be any AR at all.

The very foundation of AR, the literal floor layer, we could call descriptive AI. In very simple terms, AR is about placing three-dimensional digital objects in the physical world, and for this to happen, at least one flat surface needs to be identified on which these digital objects can be placed. This is done using data from sensors (which can be cameras, LIDARs, or something else) which is then analyzed with the help of AI to create a digital map of the outside world, a kind of digital twin that is three-dimensionally mapped and positioned by the sensor.

If the digital twin is then constantly updated as the user moves (and the sensor thus moves), digital objects can be placed on the physical surface that the user can interact with from all different angles as they move around the object. You can easily try this yourself by, for example, playing Angry Birds AR on your smartphone. When you move your smartphone across the floor in a circle, an AI algorithm in your phone's sensor puts the digital stage floor and then you can walk around the stage and play.

Unfortunately, a lot of calculations are needed if we want to be able to walk around freely and have our digital objects follow us in real time, so it takes even more AI support to make fully mobile AR applications. One way to solve this could be for several users to walk around with sensors and collaborate by sharing information and jointly building up the digital maps of the surroundings. Here, too, AI will be needed to stitch together all users' map parts into a unified digital twin over a surface that covers all users.

#### **5. Predictive AI: The AR stage prompter**

So far, we have mostly talked about flat surfaces, even though the physical reality is much more complicated. Let's say we have a Pikachu (from the game Pokémon GO) that can now run with us when we walk across a flat surface. But what if Pikachu suddenly gets scared when another person approaches and therefore jumps up a tree? Now, the Pokémon GO game needs to understand to some extent what is happening. The digital map becomes a stage where we use AI to recognize objects and find out that it is actually a person who is approaching. However, it is not enough to just recognize objects, such as the approaching human and maybe a tree to the left of that human: AI is also needed to understand that there are branches in the tree that a Pikachu could perch on, even if the sensors have not actually identified these branches (for example, because the sensors are mostly pointed at the ground to detect flat surfaces).

As we start trying to manage people and trees in our AR app, descriptive AI is no longer enough. Our descriptive AI can only conclude that one is a human and the other is a tree. Instead, now we also need predictive AI to understand the dynamics of how objects interact

and what's about to happen next, much like a prompter on a stage who whispers in the actor's ear if he gets lost. In order for Pikachu to be able to become afraid of humans in a believable way, the game needs to understand that humans are heading towards us, maybe they will even step on Pikachu, ouch! This, in turn, leads to a host of predictive domino effects; perhaps Pikachu wants to know where the approaching human's hands and gaze are, and here computer models and AI come will need to predict what human body movements look like.

## **6. Prescriptive AI: The AR playwright**

The tree, on the other hand, gives rise to a different kind of reasoning, which can be described, at least to some extent, as prescriptive (or normative) and not just predictive. Pikachu has fled up into what our object-recognizing computer models have determined to be a tree. The model knows that a tree is full of branches, but on the other hand, there may not be sensors that have actually scanned the tree, because the people wearing sensors in their AR glasses have not yet looked up at the tree. Therefore, a prescriptive AI like a playwright in a digital theater performance now creates a completely new image of the tree's branches so that Pikachu can sit on one of these. When people look up for Pikachu, the sensors are very likely to detect that the branches were not in exactly the way the AI model made them out. But instead of suddenly letting Pikachu hang in thin air, the AI model allows the imaginary branch to remain among the real ones. The problem is solved and the person playing Pokémon GO is content and happy because it all feels real. One or two extra branches pretending to belong to reality feels quite innocent, but the more philosophically inclined may feel a certain discomfort. And if so, with some justification. We have gone from using AI to find out what the world around us is like and where we can place digital objects in it, to instead letting AI decide what reality looks like!

This may seem like a hypothetical case, but there are already commercial mass-market products that use prescriptive AI of a related nature. NVIDIA's RTX 4000 series of graphics cards have a feature called DLSS 3 that is used to increase frame rates. Previous versions of DLSS used AI to predict where the next pixel in a frame would go, but DLSS 3 takes a massive leap forward, prescriptively simply plotting entirely new frames that are inserted between actual frames, thereby significantly increasing the frame rate. DLSS 3 is not yet used to manipulate the reality experience in AR, but since its predecessor, DLSS, is already being used for VR [4], the step doesn't feel like that far.

A different type of prescriptive AI, on the other hand, is already being used in a completely different way with AR. Let's go back to Pokémon GO, whose owner Niantic sells something called sponsored locations [5]. Through these sponsored locations, Niantic offers its corporate customers to purchase access to its user base. This takes the form of various location-specific mini-games such as battles between different Pokémon characters or perhaps the ability to catch an unusual Pokémon. All of this is placed in close proximity to the physical location that the buyer wants to attract visitors to. If we in traditional advertising talk about "eyeballs", this is now instead all about "footfalls", i.e. that consumers actually visit the store or whatever it may be. Getting visitors delivered in this way costs significantly more than getting someone to only see an ad on a web page, perhaps as much

as 50 cents per person, and therefore it becomes important for the buyer to optimize the number of visitors. Let's say a McDonald's restaurant is short of customers at three o'clock in the afternoon and wants to get 100 extra visitors. With the help of AI that has been trained on the user database, it is then calculated what kind of Pokemon needs to be placed at the restaurant to attract the right number of players, because exactly the right number of them have gone so far that they feel thirsty or hungry enough to go into the McDonald's restaurant in question and order something. So there is a picture of reality where the restaurant has 100 customers too few, but Pokémon GO players are now being used as pixels to redraw that reality so that it contains 100 new restaurant visitors, no more, no fewer.

However, just because prescriptive methods are being used does not mean that descriptive and predictive methods will disappear. On the contrary. Sensors and AI will not only map and act as prompters for the world around us, but also the users themselves will be mapped and their behaviors will feed predictive data models. The Meta Quest Pro is one of the first widely available headsets to be equipped with inward-facing sensors to analyze eye movements as well as facial expressions, something that is likely to become more common in the coming years, not least since Apple is employing such technology in their Vision Pro headset as well. In other words, in the future, the combination of AR and AI will be used in many ways to determine what we humans experience, and what we do.

## **7. Is augmented reality, AR, reasonable?**

In light of this, the natural question is how much augmentation of reality is reasonable, or if reality reinforcement should be allowed at all. One kind of augmentation I definitely think is needed is the move of the AR experience itself from smartphones to glasses. Although, as mentioned earlier, I've spent thousands of hours in VR, the number of hours I've spent with AR on smartphones is almost negligible. Already today, I've spent more time with AR experiences on Microsoft HoloLens and similar AR headsets, despite the fact that these have both limited viewpoints and limited application offerings. But anyone who's tried to hold a phone in front of them to experience the world of AR will quickly notice two things. Firstly, the phone becomes heavy as lead after just a few minutes in an outstretched hand, and secondly, you can easily see past the screen and then discover with disappointment that the digital objects are not there and are just fake. Despite the fact that Tim Cook, for example, claims that Apple has over 14,000 AR apps in its App Store today [6], the usability is limited and amounts to more like 14,000 cries for an Apple gadget to put on the nose. Again, with Apple Vision Pro, one indeed in the works, although a couple of iterations may be required before it's ready for the mass market.

The question of how much subjective augmentation we can imagine when it is AI that is the subject, and the human being is the object of such augmentation is of course more difficult to answer unambiguously. The division of AI into descriptive, predictive, and prescriptive components is borrowed from Shoshana Zuboff's well-known book on surveillance capitalism [7]. One of her main points is that prescriptive AI is difficult for us humans to detect because it manipulates us in a way that is often beyond our horizon of understanding. In addition, according to her, the economic driver of prescriptive AI is very

strong because it aims not to just reduce the risk behind commercial ventures, but to simply eliminate it completely. The future profit in terms of dollars and cents thus changes from a forecast to an established fact even before the activity that generates the profit has occurred. However, Zuboff does not see an easy way to prevent the development of prescriptive AI, or any simple regulatory framework to limit any negative effects on society. Instead, she hopes for a broad commitment to these issues, from the general public as well as from journalists and politicians.

Thus, manipulations using prescriptive AI are by no means limited to the metaverse. On the contrary, developments are essentially taking place on other fronts today, as the metaverse is still in its infancy. Prescriptive processing of digital images will probably continue to be about game development and different types of simulations for professional use before the step to the metaverse is taken, and the focus of prescriptive AI for the consumer market may well continue to be the more traditional advertising for the flat-screen version of the internet we already have today for many more years.

But it is precisely in the metaverse that the question becomes literally existential: If the metaverse cannot exist without AI, how should we deal with a digitized reality that is increasingly staged by a dramaturge that is not human?

## References

- [1] Universal Turing Machine implemented in Minecraft redstone logic, 2011. URL: <https://www.youtube.com/watch?v=1X21HQphy6I>
- [2] I Built the Entire Universe in Minecraft, 2022. URL: [https://www.reddit.com/r/Minecraftbuilds/comments/xviuyi/i\\_built\\_the\\_entire\\_universe\\_in\\_minecraft/](https://www.reddit.com/r/Minecraftbuilds/comments/xviuyi/i_built_the_entire_universe_in_minecraft/)
- [3] Second Life, 2003. URL: [https://en.wikipedia.org/wiki/Second\\_Life](https://en.wikipedia.org/wiki/Second_Life)
- [4] Nvidia DLSS improves the frame rate of Microsoft Flight Simulator VR, 2022. URL: <https://mixed-news.com/en/nvidia-dlss-improves-the-frame-rate-of-microsoft-flight-simulator-vr/>
- [5] Sponsored Locations for Business, 2018. URL: <https://nianticlabs.com/en/sponsoredlocations/>
- [6] Tim Cook once again teases Apple's AR tech, 2022. URL: <https://9to5mac.com/2022/06/22/tim-cook-ar-headset-interview/>
- [7] S. Zuboff, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*, 1<sup>st</sup>. ed., p. 704, PublicAffairs. p. 704, New York, NY, 2019.