



**V**IDEO TOOLS NOW ON THE MARKET ALLOW one person's face or voice to be replaced by another's, creating what are known as "deepfakes" that can fool most viewers.

Audio tools can replicate a person's voice from samples. Anyone can be made to say anything.

Such hoaxes could be used to destabilize delicate situations like trade negotiations or criminal investigations, delegitimize reputable sources or slander celebrities or political candidates. They could also wreak global havoc: Imagine a convincing, but fake announcement by the President of the United States of a nuclear missile strike against North Korea.

In the spring of 2016, before the tsunami of "fake news" roiled the 2016 US presidential race, only a few saw the impending danger. A young MIT alumnus and Silicon Valley consultant named Aviv Ovadya was one.

"It became clear we were at an inflection point," says Mr. Ovadya, now Chief Technologist for the University of Michigan's new Center for Social Media Responsibility and a Knight News

# PUPPET

Innovation Fellow at Columbia University's Tow Center for Digital Journalism. While noting that much good can come from these innovations, Mr. Ovadya compares the growing threat to that of nuclear weapons, and sees society's awareness as myopic – "a one-inch view of the outside through the windshield" of a car careening out of control.

Over the past two years, his warning of a looming "infocalypse" has drawn attention, and Facebook, Twitter, Google and other platforms have put more resources into preventing malicious use of their products. The next step, Mr. Ovadya

says, is a commitment to massive investment to develop countermeasures, and to allocate “nimble money” – talent pools and shared resources across organizations that can be deployed quickly as the fast-moving technology creates new threats.

**Can you tell me a little about your background?**

At MIT, I studied computer science. But a big part of the conversation in the community around me was about the impact of technology on society. During that time, I came to terms with the idea that maybe technology isn't an unqualified good. It can change the way the world works; it can put you into a better world and it can put you into a worse world.

That was pretty formative – realizing that there's a trade-off between the efficiency that comes from technology and resilience, which is often lost as a result. Technology can make the world much more fragile. I went on to get my Master's at MIT and then spent a bunch of time in Silicon Valley, as a software engineer and product design consultant. But on the side, I was working on understanding some of these systems around technology and society.

About two years ago now it became clear that we were at an inflection point. The means of distribution of information was being manipulated, co-opted and optimized in a way that was really harmful for democracy, for public discourse, for health – for all these things that we clearly care about in society.

The truth is under threat from “extreme reality manipulation.” **AVIV OVADYA**, prophet of a looming “infocalypse,” speaks to Brunswick's **CARLTON WILKINSON** about how business can fight back

# MASTERS

Not only was it very bad already, but it was going to get much worse very quickly. And there was nothing being done that would make it not continue to get worse. That was what triggered me into action. That isn't acceptable. That isn't a world I want to live in. So, I decided to focus my energies to see what I can do about that.

**What kind of reception did you get when you started to spread the word about this in 2016?**

Probably the most common response was, “That's not actually a problem. Prove to me that it's a

problem.” You still hear some of that: “This has always been true. Nothing's new.” But there's a lot of evidence to the contrary at this point.

It's sort of like saying, “Nukes aren't really a problem because there was always war.” Well, they actually are. They changed the game in a way that wasn't possible before and as a result, you need to change the entire face of diplomacy, among other things. It's true, nukes don't do anything new – you could use a spear to kill someone. But at some level, it's definitely new – in terms of the scale and scope, for instance.

**Do you think the impact of “fake news” in the election helped prove your point?**

Yes, there's a lot more interest – whether or not there's actually been effective investment. But that's starting to happen and it's good to see. It's still too little, too slow. It's a big ship, but when you've decided you want to move it, it can be moved quick.

There are organizations that have invested single-digit millions of dollars, where tens of millions actually need to be invested by many different organizations across the board – and billions across the ecosystem – to address these threats as they continue to spiral.

Likewise, it's good to see some of the platforms taking this seriously. Even people at the very top in some cases are owning up, saying, “Hey, we didn't do a great job.” The more that happens, the more likely it is that there will be significant progress.

**Are there specific technologies you're most concerned about or is it a pool of technologies?**

The overall threat is really in two components. One is the ability to make it look like anything has happened – this extreme reality manipulation. The other is being able to persuade people because you build a model that fits what that person would like to believe. Those go hand in hand and can be extremely powerful.

These threats are worrisome from a cybersecurity perspective, from a diplomacy perspective, from a policy perspective, from an electoral integrity perspective, from an education perspective. Just so many interesting ways that this could be applied that aren't particularly positive.

In the US, the issue is seen to involve a bias of far right against the far left. But if you have people known for manipulating video, trying to manipulate narratives, the bias isn't left and right. The bias is people who are willing to manipulate reality versus people who aren't.

**AVIV OVADYA**

Chief Technologist at the Center for Social Media Responsibility at the University of Michigan and a Knight News Innovation Fellow at the Tow Center for Digital Journalism at Columbia University, Aviv Ovadya is focused on identifying, measuring and mitigating indirect harms of social media and related technologies that affect public discourse.

**CENTER FOR SOCIAL MEDIA RESPONSIBILITY**  
As part of the University of Michigan School of Information, the Center for Social Media Responsibility opened in 2018 to foster dialogue between media makers, consumers, platform companies, researchers and academics about social media in society.

**Are you worried that technologies might emerge that you and others aren't predicting?**

I don't profess to know all the horrible things that might happen – and also all the good things. Probably the worst and the best things that will happen we can't quite predict. But that doesn't absolve us from doing our best to predict them. Otherwise you're going to be reactive. And maybe you're reactive two years too late because that's how long it takes for the funding timeline to work. That's a recipe for disaster.

Just having a body of experts who understand what is already happening, doing scenario models, that's crucial.

**Do you have recommendations for boards or investors?**

The investment that should be happening is not just within the social media part of the tech industry but in the entire supply chain. How a camera or phone gets made – there are things there that are relevant to talk about. We need an authenticity infrastructure. There can be a very long delay before it starts being created. They have to start now.

To prevent the kind of abuses we know are coming, we need investment now beyond just the Big Five – Apple, Microsoft, Amazon, Google and



In April, "Get Out" director Jordan Peele, working with BuzzFeed, used President Barack Obama's face and voice to call President Donald Trump "a total and complete dipshit." Peele reveals the deepfake ruse in the YouTube video and warns viewers to "stay woke" about growing threats to truth.

Facebook. It also requires lower-level or different parts of the ecosystem. And perhaps even a new kind of corporate social responsibility.

These problems are evolving very rapidly and threats are going to emerge very quickly, so we'll need nimble money. Being able to address new threats as they come up – not having a six-month, one-year, two-year cycle before that happens – that's absolutely critical. That means a talent pipeline, to ensure that people put into these roles can actually

work effectively. You need to be able to create an emergency task force with amazing people very quickly – these are people who would have gone to Google or Facebook, or been a partner at McKinsey or something. Have them really working to understand and address new threats, in combination with all the types of stakeholders that are relevant – social scientists, diplomats, journalists, whoever they may be.

These are all challenging organizational problems. But if we don't address them, it's unlikely we'll be able to handle what gets thrown at us, whatever that ends up being. We're going to have repeats of information ecosystem failures, another step function in the de-legitimization of institutions that ensure that our society actually works.

**So this requires board level conversations and new organizational functions?**

Probably the most realistic way is for each company to execute on this independently, given the way companies work. But we should also have cross-company organizations that are focused on this, not just for one company's benefit, but for the benefit of all – for the benefit of these other "brands," like democracy.

These are broad recommendations. Executives, venture capitalists, board members, technology officers – these people individually are going to have very specific questions about various aspects of these issues, how to proceed, how to measure a particular threat, how to coordinate with one another, where best to invest. I'm here to help – to answer many of those questions – to create the infrastructure we need to take on these challenges.

**Looking ahead 20 years, do you think we're going to have found the right solutions? Are you optimistic?**

It is possible we can make it to 20 years from now. My goal is to make sure that we make it that long – while still having this level of democracy and a functioning society.

People are only now waking up to the coming threats. I'm trying to go beyond that – to actually build the necessary institutions that can take on these challenges. If we make it to 20 years, we will have figured it out. So, *if* we make it, then yes, I'm optimistic.

**CARLTON WILKINSON** is a Brunswick Director and Managing Editor of the Brunswick Review, based in New York.