Synthetic Media & Content
**KEY INSIGHT**

Synthetic media consists of algorithmically generated digital content, including audio, video, deepfakes, virtual characters and environments, and more. The technology will become an integral aspect of future XR experiences.

**EXAMPLES**

Synthetic media is created using artificial intelligence. Algorithms use an initial set of data to learn—people, voices, photos, objects, motions, videos, text, and other types of media. The end result is realistic-looking and realistic-sounding artificial digital content. Voice clones, voice skins, unique gestures, photos, and interactive bots are all part of the ecosystem. Synthetic media can be used for practical reasons, such as generating characters in animated movies or acting as a stand-in for live action movies. Synthetic media can automate dubbing in foreign languages on video chats and fill in the blanks when video call frames are dropped because of low bandwidth issues. Imagine an entirely new genre of soap opera, where AI systems learn from your digital behavior, biometrics, and personal data, and use details from your personal life for the storylines of synthetic characters. In an ultimate expression of a “reality show,” synthetic characters would play to an audience of exactly one: you.

**DISRUPTIVE IMPACT**

Watch for synthetic media to appear more frequently in 2021, representing new opportunities and risks for businesses, and reshaping the entertainment, service, and communications landscape.

**EMERGING PLAYERS**

- Samsung Next
- Loudly
- Endel
- Replica Studios
- Lovo
- Modulate
- Rephrase.ai
- Synthesia
- Alethea.ai
- Carv3d
- Animatico
- Narrativa
- DeepNatural
- Baidu Research

Chinese virtual influencer Ling debuted in May 2020.
**Synthetic Media Technologies**

---

**Speech Synthesis**

Also known as “synthetic speech” or “text-to-speech technology,” speech synthesis mimics real human voices and deploys them to various interfaces. With enough data and training, a speech synthesis system can learn the spectral frequency of anyone’s voice and produce someone’s digital voiceprint. One company, Synthesia, uses this technology to dub people through automated facial reanimation. This will be especially useful for movies with wide, international releases. Actors’ facial expressions and mouths can be reformatted to ensure local languages are synchronized correctly.

---

**Modulating Custom Voices**

Generative algorithms are creating synthetic voices that sound just like the original, and those voices can be modulated to the exact pitch and tone desired. The AI learns over time to recognize not only intonation, but also emotional cadences. Replica Studios, Lovo, Voicemod, Resemble AI, and DeepZen synthesize voices for a host of purposes. For example, you can fake a conversation between yourself and your favorite celebrity, provided there are enough publicly available audio files of that celebrity to build a dataset. Soon, the technology will be able to match and rapidly deploy synthetic voices personalized for every consumer. If you loved “Daria” as a kid, you might hear Daria Morgendorffer’s voice (or Jane Lane’s) during a car commercial, while your spouse might instead hear Phil and Lil’s mom from “Rugrats.” We should remember that in this era of misinformation, synthetic voices might also be used to trick unwitting consumers.

---

**Deep Behaviors and Predictive Machine Vision**

Researchers at MIT’s Computer Science and Artificial Intelligence Laboratory (CSAIL) trained computers to not only recognize what’s in a video but to predict what humans will do next. Trained on YouTube videos and TV shows such as “The Office” and “Desperate Housewives,” a computer system can now predict whether two people are likely to hug, kiss, shake hands, or slap a high five. This research will someday enable robots to more easily navigate human environments—and to interact with humans by taking cues from our body language. It could also be used in retail environments, while we’re operating machinery or while we’re in classrooms learning.

---

**Generative Algorithms for Voice, Sound, and Video**

A team at University of California, Berkeley, created software that can transfer the movements of one person in a video to someone in another video automatically using a generative adversarial network. For some time, we’ve been training computers to watch videos and predict corresponding sounds in our physical world. For example, researchers at MIT’s CSAIL experimented to learn whether a computer could accurately predict what sound is generated when a wooden drumstick taps a couch, a pile of leaves, or a glass windowpane. The focus of this research is to help systems understand how objects interact with each other in the physical realm.

---

**Mapped Synthetic Environments**

Companies are now mapping the real world to generate synthetic digital twins. Amazon has been studying Snohomish County in Washington, building realistic simulations of the region’s roads, buildings, and traffic. Its maps are reported to be accurate down to the centimeter, precisely tracking subtle gradients in pavement and noting unique markings on sidewalks. Amazon fused maps and 3D data to build synthetic versions of the county to test delivery drones. These kinds of virtual environments will be necessary as the company moves drones from research labs into the mainstream. Amazon tested its Scout delivery robot in the real world, having trained it in the synthetic environment.
Live Portraits

In a new application of synthetic media, Israeli startup D-ID used its face recognition technology to generate live portraits from old photos. Members of MyHeritage, a genealogy platform, could upload old family photos or animate those already in the platform’s database. The result: live portraits, reminiscent of “Harry Potter.” The technology maps a digital picture to an AI system that renders smiling, blinking, and head movements in a short video.

Synthetic Voice Fraud

Synthesized media has a known problem area: It can be used by malicious actors to mislead people, to trick voice authentication systems, and to create forged audio recordings. Voice fraud costs U.S. businesses with call centers $14 billion last year alone, according to a study by call center software maker Pindrop. Google has been working on a synthetic speech dataset as part of the ASVspoof Challenge, which is an open-source, global initiative designed to help develop countermeasures to fight spoofed speech. Researchers hope that the challenge will lead to more secure synthetic voice content.

Synthetic media will give rise to an entirely new kind of celebrity in the 21st century: synthetic pop stars. It also affords a host of opportunities to make and save money. Already there are a number of synthetic pop stars with very large fan bases. Lil Miquela is a sort of Beyonce of synthetic stars, with 3 million followers on Instagram as of the start of this year. She is a model for brands like Prada and Calvin Klein, a musician with popular tracks on Spotify, and a paid brand ambassador for enormous, global companies like Samsung. And she has friends: Bermuda, a rule-breaking bad girl model/influencer and Blawko, a Los Angeles-based guy who likes fast cars and Absolut vodka, and is never without his trademark face scarf covering his nose and mouth. In many ways, these stars are the antidote to teen stars like Lindsay Lohan and Shia LeBeouf who, for one reason or another, stray from their carefully crafted public images and cause headaches for their agents, managers, and the brands or projects they represent. Synthetic stars don’t sleep. They don’t eat. They never get tired, even if they’re pushed 24 hours a day. They don’t drink alcohol or use drugs, would never say anything off-messaging, and their mug shots would never go viral. (Unless it was planned, of course. Over the summer, Bermuda posted her own mugshot on Instagram to “get ahead” of the press.) While Bermuda and Blawko aren’t programmable yet, China’s AI news anchors are. China’s state news agency Xinhua employs AI news anchors Xin Xiaomeng, Qui Hao, and Xin Xiaohao, who appear in videos and also write stories for the agency.

Simulating Human Experiences

What if you could interact with a simulated person to learn from them or practice management techniques? Would you invite a synth to a dinner party? Samsung’s Technology and Advanced Research Labs (STAR Labs) thinks the answer is yes. It developed Neon, “a computationally created virtual being that looks and behaves like a real human, with the ability to show emotions and intelligence.” Neons aren’t intended as a stand-in for the internet. They were built to hang out with you. U.S.-based startup Talespin built synths in virtual reality to teach people “soft” management skills, including how to encourage team members or how to fire someone with empathy and compassion. Canadian startup TwentyBN built a synthetic sales associate to cheerfully interact with customers—and convince them to spend more money.

Synthetic Media and Society
Synthetic Media and Society

Synthetic Sex Tapes

Natalie Portman, Emma Watson, Taylor Swift, and Daisy Ridley—smart, talented artists—began “appearing” in adult videos in late 2018. Convincing short clips were made using deepfake techniques and soon went viral on Reddit. Not too long after, another Reddit user published a mobile application allowing anyone to make their own porn deepfakes. This poses a particular problem for public figures, because right now there isn’t an easy way for the average person to tell what’s real and what’s fake. Photos and videos can spread through social networks and online without much protection for those victimized. In the absence of digital tools to spot fakes, we’re left relying on critical thinking and common sense.

Synthetic Property Rights and Legal Protections

The video game “Call of Duty: Modern Warfare” was designed with brutal realism. Players enter lifelike combat situations and must decide whether to shoot synthetic civilians. Where do we draw the lines between disclosure and pure fantasy? Parody for laughs and deepfakes for harm? What happens when synthetic content seems so real that the psychological implications are intense and profound? What if someone generates synthetic environments that mirror real-world situations and real people? No existing laws or regulations govern synthetic content, although some people suggest adapting current laws, such as those covering libel, defamation, identity fraud, or impersonating a government official.

“Under Pressure” (you know the base line hook: da-da-da---da-da-da-dum) but didn’t get permission first. He tried to get around copyright law saying that he added a beat between notes (Ice’s version: da-da-da...da-da-da-dum, DA---da-da-da-dum) and made it a distinctly different song. The case settled out of court, but it shined a light on how U.S. copyright laws were created to protect the financial—not creative—interests of artists. What if someone created a slightly altered copy of you for promotional commercials? For example, if your likeness was edited to include facial hair and a pair of glasses you don’t have in real life, and then used without your permission—would those details eliminate the legal requirement for consent?

Synthetic Media Marketplaces

We already got a taste of what our future synth media marketplaces will look like. In 2018, a subreddit dedicated to publishing deepfakes morphed into a makeshift marketplace. Users were volunteering to create deepfake videos of celebrities, co-workers, family members, neighbors, and enemies in exchange for crypto-currency. In the near future, marketplaces to commission, buy, and sell synthetic media, as well as their attributes, will be visible on the dark social web.

Truth Decay in an Era of Synthetic Media

In 2021, we expect to see synthetic media technologies further commercialized and made widely available. But without the infrastructure in place to help consumers distinguish between synths and humans, the likelihood of misinformation campaigns remains high. Synthetic media could be weaponized by governments, activist groups, and individuals, and could be treated the same as all other internet content, showing up in search results, on our smart speakers as audio content, on our connected TVs, in our inboxes, and throughout social media. Synthetic media can be particularly inflammatory in the political realm, where both sides of a polarized public are motivated by social media algorithms to share and engage with sensationalized content without necessarily verifying its authenticity.