

CHAPTER 5

Virtual Style Assistants

Assistance is the new black.

– Aparna Chennapragada, director of product, Google

The idea of creating an artificially intelligent personal stylist has been frequently revisited in popular entertainment. The first time I can remember being exposed to this idea was watching the film *Clueless* (circa 1995). Cher is choosing her outfit for school by using a computer system that tells her “Mis-Match” for outfits that don’t style well together and shows her what the outfit will look like on her.

According to the McKinsey and Business of Fashion report for 2018, “75% of retailers plan to invest in AI over the next two years.”

The AI personal stylist concept is a culmination of the technologies discussed in this book, and a look into the future of specialized assistants. The virtual style assistant will be one of the more intimate in its class, knowing more anthropomorphic data about users than other virtual assistants and comparable products. It also provides a personalized future for e-commerce.

The virtual style assistant brings together several areas of artificial intelligence, including natural language processing, natural language understanding, computer vision, neural networks, and other types of machine learning.

Virtual Style Assistants

The only real elegance is in the mind; if you've got that, the rest really comes from it.

—Diana Vreeland, former editor-in-chief, *Vogue* magazine

The virtual style assistant is useful when it comes to fashion sales. Bringing personal stylists to the retail and e-commerce environment can improve a brand's ability to match consumers with desired products and guide context-based decision-making. It can also be brought into the home of the consumer, making better use of existing wardrobes. An AI stylist can help consumers discover apparel items that meet a wide variety of expectations: flatter their figure, work well as an outfit, align with current trends and values, and provide a personalized experience.

To understand the impact and context of creating an artificially intelligent style assistant, it helps to be familiar with both the role of a personal stylist and recent developments in technology.

Personal Stylists

Personal stylists help individuals to look their best by curating clothing, outfits, makeup, and other aspects of personal style. This service is often available to only people who can afford it. Hiring a personal stylist would be considered out of reach for average citizens across the United States because of economic or geographic constraints.

Several solutions on the market use technology to address a desire for personal styling. Platforms that act as social networks between stylists and consumers bridge this gap. However, because they continue to rely on services by humans, these services will not scale at the rate that an AI style assistant can. It is because of its scalability and potential to capture value in this field that the AI stylist is a compelling use case.

AI stylists just aren't very good compared to humans right now. Style encompasses many things machines still don't understand. The introduction of an AI stylist doesn't mark the end of the personal stylist as an employment opportunity. Stylists curate aesthetics for their clients, often interpreting between the lines as the clients describe their personal style. They help guide people through an experience that might cause them to feel fear, uncertainty, insecurity, embarrassment, or confusion. Stylists provide a personal experience of shepherding individuals into a journey of self-expression. Humans trust humans more than they trust businesses or software. When it comes to something as deeply personal as our appearances, a real-life human recommendation will always override an AI system recommendation.

Virtual Assistants

Virtual assistants provide the basis for the virtual style assistant and have become increasingly prevalent in consumer electronics. These assistants generically refer to a software agent that provides services to individuals. Today, this is often carried out through voice command prompts. By brand name, virtual assistants refer to Apple's Siri, Google's Google Home and Google Assistant, Amazon's Alexa, and other similar AI-based assistants.

The first step of the process for these systems is interpreting human input. For today's virtual assistants, that usually means converting speech to text. The AI records the human voice and creates text from those recordings in real time. This process is generally referred to as **speech to text**, or **automatic speech recognition (ASR)**.

Voice Interfaces

The ability to purchase products is also available in many virtual assistants. Voice interfaces, sometimes referred to as **voice user interfaces (VUIs)**, are changing the way that consumers behave. According to Google/

Peerless reports from August 2017, 58% of people who use a voice-activated speaker are now creating and managing shopping lists at least once per week, and 62% say they are likely to purchase something through their voice-activated device in the next month.

The hardware that is being manufactured today is able to support **far-field voice input processing (FFVIP)**. This has opened up a wider variety of use cases for voice interface-based devices by making it possible to speak to them from further away. This has been enabled in part by using multiple microphones; the iPhone 5 has three, and the Amazon Echo has seven. The devices use delay between picking up the same sound in different microphones to identify where the sound is coming from and cancel the sound coming into the other speakers.

Features of the Virtual Style Assistant

The virtual style assistant is unique compared to other virtual assistants in that it emphasizes the use of images more than any other use case. Images are critical to giving style advice.

The design of a virtual style assistant requires a few key components: the ability to take a photo of oneself and the ability to store those photos in the application. It also requires underlying technologies such as computer vision capabilities for image recognition and visual search, recommendation engines, analytics, and access to fashion products.

Existing Examples

Images collected by a virtual style assistant can be put to use in several ways. In this and other existing virtual style assistants, images are used to catalog a wardrobe, make product recommendations, and provide insights into the user's style preferences.

Amazon's Echo Look is the most public example of a virtual style assistant. In addition to Amazon, various startups have begun to develop technologies for this purpose. From the behind-the-scenes software created by some of the subscription box companies to unique apps (for example, Lookastic, which makes recommendations based on your response to their "What's in your closet?" interface). Companies such as MemoMi are building smart mirrors that can assist customers in retail locations.

Defining personal style is difficult. In some ways, other platforms have been used to meet the need for a virtual style assistant. Pinterest, for example, recommends images on their web site that are like those you collect on boards. Fashion is one of their most popular types of content. Once images are recommended or discovered by the user, some of them are even shoppable.

Amazon's Echo Look

The Amazon Echo Look was released by Amazon in 2017. The device is intended to operate as a personal stylist, most notably giving AI-generated advice about how two outfits look on you when compared. The Echo Look app provides these software features:

- It allows customers to take better selfies. A one-time setup results in a series of consistent selfies. It provides flash and some image correction to create a focus on the person and the outfit. Better selfies provide more consistent data for Amazon to use when providing analysis and recommendations.
- It allows the customer to create a history of outfits they've worn.

- It recommends garments that you can buy on Amazon based on what you're wearing in a given photograph.
- There are insights about the colors in a customer's wardrobe and their prevalence.

As the history of images of outfits grows, that library is more likely to be useful to the user. The Amazon Echo Look is shown in Figure 5-1.

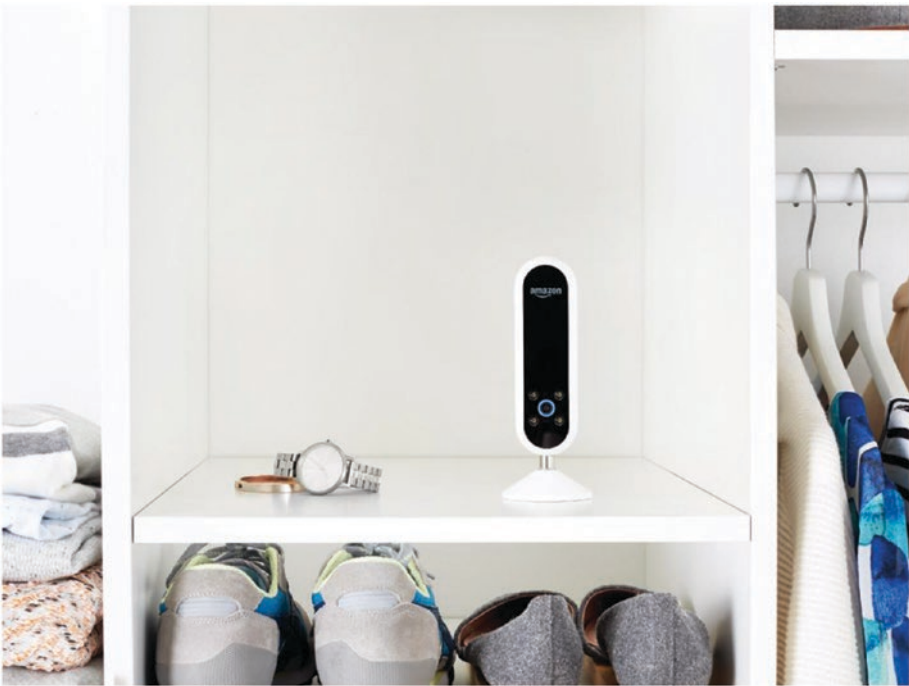


Figure 5-1. The Amazon Echo Look (image courtesy of Amazon)

The Hardware

The hardware in this device is packed with features. It is capable of doing far more than the Echo Look application requires of it. The app itself could perform the functions it provides using hardware from your phone. It has a 5-megapixel camera as well as an infrared camera that gives it the ability to

sense depth. In theory, it could accurately identify points of measure and measurement data from the bodies it photographs. The device can also store up to 8 GB of data locally. We can only speculate about what Amazon is doing with the data from this extravagant piece of hardware.

Infrared cameras detect heat and produce a thermal image to represent that heat. The Echo Look will blur out the background surrounding the person in the image. It is using infrared to find the human in the frame and blur everything around its figure.

Mobile-device technology is quickly outpacing this approach. Google's introduction of Portrait Mode on its Pixel 2 device released in 2017 features similar technical capabilities, such as putting the subject into focus and blurring the background. In fact, Google's introduction of Google Lens also indirectly competes with Amazon shopping recommendations in the app. Figure 5-2 shows an image of Google Lens finding an exact match for the maroon shirt pictured in the background. In the future, the most likely virtual style assistant will be your phone.

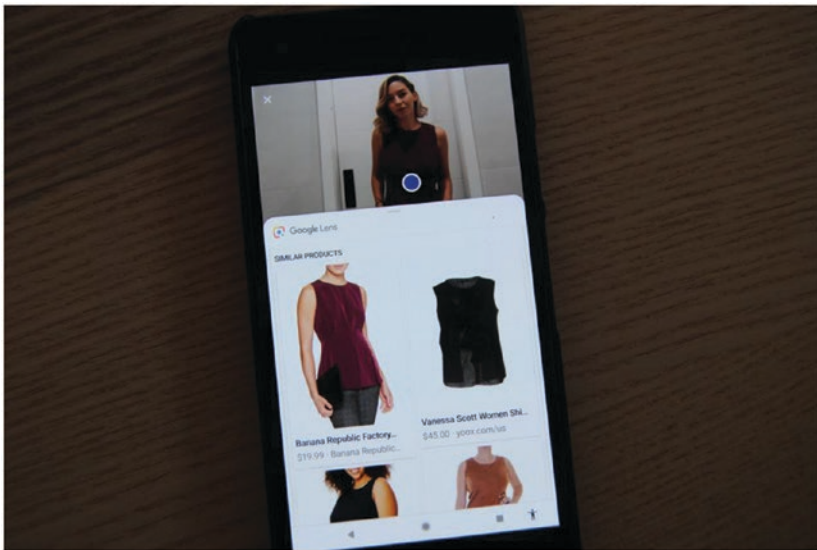


Figure 5-2. An example search result from Google Lens

Image-Based Reviews

Image-based reviews provide a huge amount of information to consumers who are making a decision about purchasing a fashion product—including insights to fit, fabrics, and details that might not be obvious in the studio photos found on product pages.

Despite the value of image-based reviews, none of the top ten fashion retailers include them on their e-commerce sites. In fact, of that top ten, half of them don't have customer reviews on their web sites at all. It's rare that people today will go anywhere, do anything, or buy anything without reading the reviews.

In other e-commerce categories, image reviews are commonplace. Of the top three home furnishing retailers—West Elm, Pottery Barn, and Williams-Sonoma—all three feature image reviews prevalently on their site's product pages.

Why in the fashion industry have image-based reviews been so neglected? Today, it's inconvenient to take a picture, find the product page, and upload a review of a garment. The process is messy and time-consuming, too much of an ask for a busy customer.

The Future of Image-Based Reviews

On our smartphones, when we are inside a restaurant, our map app knows we are inside that restaurant. We are more likely to leave reviews as we are prompted to in that moment, or later when we're reminded that we had been there. We might even include an image of the dish we ate.

Over time, it will become easier and more prevalent for our mobile phones to have information about what we're wearing. In the future, they will find the product pages for us and prompt us to leave garment reviews while we still have that garment on. Our mobile accounts will have a history of all the things we've photographed ourselves wearing and may even recommend outfits curated from what's already in our closet, featuring our favorite bloggers, lookbooks, and other users who have shared their outfits.

Artificial General Intelligence

Like, yes—in particular areas, machines have superhuman performance, but in terms of general intelligence, we're not even close to a rat.

—Yann LeCun, head of AI research, Facebook

The virtual style assistant is in some ways aligned with the goal of **artificial general intelligence (AGI)** to represent the full spectrum of human intelligence.

AGI is also referred to as **strong AI, full AI** or **AI complete**. The specialized intelligence discussed in the rest of this book is often referred to as **weak AI, narrow AI**, or **applied AI** because its use is usually limited to a specific application.

Strong AI is rooted in the premise that artificial intelligence should achieve the same levels of intelligence as humans. Currently, this concept is overshadowed by applied AI methods, not because it isn't a desired outcome, but because there is lower-hanging fruit. Strong AI is a really hard problem that hasn't been solved.

Applied AI is focused on specific tasks or areas of problem solving. Even the virtual assistants we are interacting with today embody narrow AI systems. When a user is interacting with a system like Siri, or Google

Assistant asks a question that is out of reach for the system, it returns an Internet search query rather than smartly connecting the user with the app or activity that they are trying to accomplish.

Hybrid Intelligence

To create the illusion of artificial general intelligence, some companies, such as Fin, have created hybrid virtual assistant services. These hybrid systems capitalize on the strengths of current artificial intelligence and use humans to assist, covering the gap between what machines can do and what humans need. These systems are a stepping stone that enables us to learn about the type of assistance humans are looking for and to close the gap between what is possible and what is useful and expected.

Pitfalls of Artificial General Intelligence

Unfortunately, with the rise of more powerful and accessible tools, a lot of misinformation and misunderstanding about AI has emerged in the public media. For some people, oversimplifications have led them to believe that we are further ahead than we actually are in developing AI tools. Some may be disappointed to find that AI falls short of their expectations.

Dangers of AI

Just as we use law to create a system of checks and balances that guide human behavior, we will do the same for AI systems and again for the humans that guide them.

There is a lot of media discussing the potential dangers of implementing artificially intelligent systems. For most experts, this caution isn't about an AI machine becoming a sentient being and taking over the world. The warning comes in really practical problems that can arise when relying on AI alone.

AI systems, for example, currently do not understand cultural, social, or ethical norms very well. They generally cannot exercise good judgment in the face of complex scenarios that require a lot of context. In terms of NLP and NLU, we have already seen major disruptions to the US economy and other economies based on real-world reactions of these technologies in financial systems.

In 2008, a news article was accidentally published declaring the bankruptcy of United Airlines. Because high-frequency traders were relying on analysis by NLU, automated trading reacted in a matter of seconds. The market value plummeted, losing \$1 billion in a matter of 12 minutes. Examples like these, of which there are several, don't necessarily emphasize the danger of AI. The humans building these systems can't always predict disasters like these but have become more aware of adding safety checks.

Summary

The virtual style assistant is an idea that has provoked our imaginations for decades. Finally, this concept is coming into fruition with the implementation of automatic speech recognition, natural language processing, computer vision, and other technological advances in the last decade.

The virtual style assistant is a specialized concept stemming from the emergence of more-general virtual assistants. This application of AI emphasizes voice interfaces to complete tasks for users at their spoken request. With only a few examples of the virtual style assistant, many of the features for this application remain to be determined, leaving a large opportunity for development in the space.

Terminology from This Chapter

AI complete—The mindset that AI should carry out all human cognitive capabilities. See also *artificial general intelligence (AGI)*, *full AI*, and *strong AI*.

Applied AI—The application of AI to real-world, specific problems, often outperforming humans at these specialized tasks. This is the most prevalent form of AI today.

Artificial general intelligence (AGI)—The goal of creating “thinking machines” that serve as general-purpose systems that are as intelligent as a human. See also *AI complete*.

Automatic speech recognition (ASR)—Turns spoken human language into text in real time. This is a prerequisite task to voice interface systems.

Far-field voice input processing (FFVIP)—The processing of voice commands that are dictated at a distance from the microphone of a smart device.

Full AI—See *AI complete* and *artificial general intelligence (AGI)*.

Narrow AI—See *applied AI*.

Personal stylist—A professional who consults individuals on their personal style, including hair and makeup, fashion and accessories, and other aspects.

Speech to text—Also described as speech recognition, speech-to-text takes spoken language and translates it to text. See also *automatic speech recognition (ASR)*.

Strong AI—See *AI complete* and *artificial general intelligence (AGI)*.

Virtual assistants—Specialized AI-based assistants including Google Assistant, Siri, and Amazon Alexa. See also *AI assistant*.

Voice user interfaces (VUIs)—Allow people to use voice-based interactions to interface with machines rather than using screens.

Weak AI—See *applied AI*.