aws

Werner Vogels' Tech Predictions for **2022** and **Beyond**

Dr. Werner Vogels,
VP & CTO at Amazon.com

Contents

We have reached an inflection point. After AWS pioneered cloud technology more than 15 years ago, cloud infrastructure has evolved to a place where we are seeing all parts of the cloud reach practically anywhere on the planet—and even into space. The cloud has allowed what was once science fiction to become science fact. Models and techniques in the artificial intelligence (AI) and machine learning (ML) realm have gotten better and better—so much so that we see glimpses of new kinds of use cases emerging that we previously only imagined in movies and comics. We are entering a phase where data is abundant, access to it is almost instantaneous, and our ability to make sense of it in new and subtle ways is practically automatic. But this technology is not replacing humans; it is augmenting how we engage with the world. 2022 will be an exciting year for technology, with it pushing all of us, and our planet, forward in the process.

AI-supported software development **takes hold**

Software development is a creative process, but one that has many repetitive tasks. In 2022, ML will begin to play a major role in augmenting software developers' workstreams, helping them create more secure and reliable code.

Since the advent of the cloud, we've seen companies across the world bring new ideas to their customers at scale faster than ever. However, even with this acceleration in product delivery, people still spend a disproportional amount of time in one area of technology: software development.

While there have absolutely been improvements in programming languages, software development toolkits (SDKs), and other tools that enable developers to be more efficient, these have all been minor evolutions. There haven't been the major leaps forward that we've seen in other areas of technology … yet. Over the past few years, we've started to see the foundations laid for what will become the future of software development.
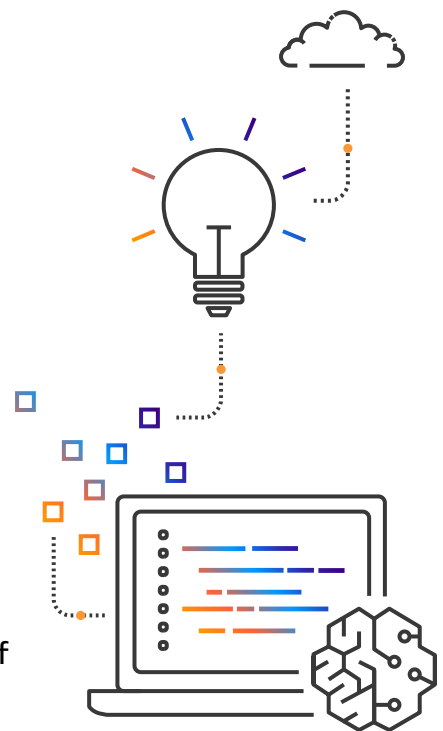
Tools like Amazon DevOps Guru, Amazon CodeGuru, GitHub Copilot, and GPT-3 are the first steps in what I see as the future of development, where ML is used in code development and software operations workstreams to help developers become more effective. In the coming years, I believe that we are going to see an explosion of capabilities in this area.

ML will free developers from the mundane parts of their jobs, such as code reviews and bug fixes —the undifferentiated heavy lifting of their world—and allow them to focus more on creating. The same technology will help us write sophisticated systems faster than ever and in ways that open the door to a new class of developers. Imagine a scenario where a builder describes how they want an app to operate, and then the tools interpret the request through natural language processing and deliver back the fully functional code. On the backend, ML techniques will also check for software bugs and continuously verify that the software is doing what it is supposed to do. This kind of ML-supported software development will be a game-changer by allowing more people across an organization to help define and build software and software-driven products. Longer term, this same approach of ML freeing us up to create more will be used in all kinds of other areas, including media creation. We will see generative AI techniques increasingly create movies, music, and literature. Just as importantly, in a similar way, this technology will also start to play a role in detecting fake content, scams, and fraud. 2022 is the year where AI/ML takes on the heavy lifting in the lives of developers.
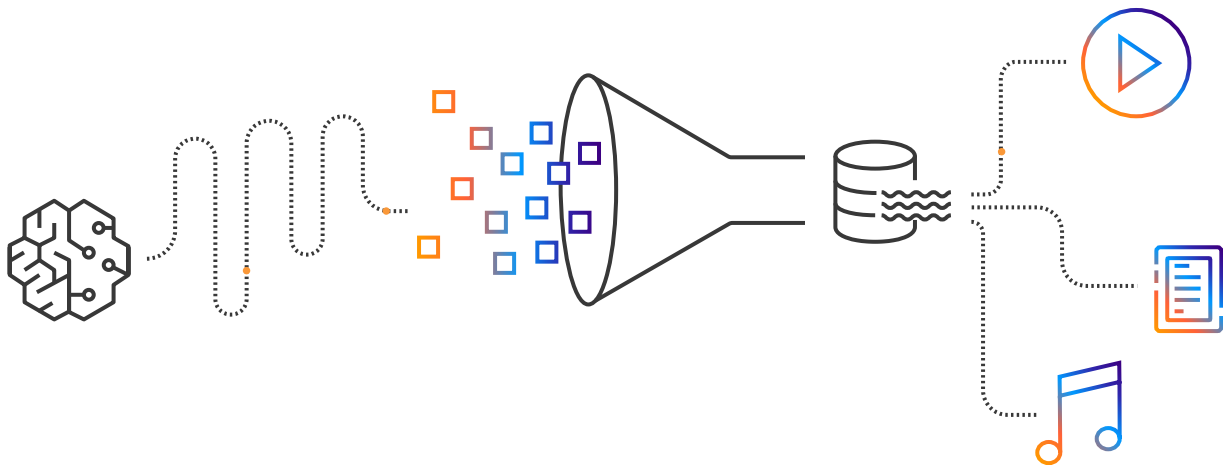
The everywhere **cloud has an edge**

The cloud will extend into every locale via purpose-built devices and specialized solutions. In 2022, we'll see these solutions bring all the muscle of the cloud to transform warehouses, restaurants, retail stores, farms, and more.

When talking about supply chain and transportation, we refer to the last leg of a journey as the "last mile." It's that final trip to your door. In ecommerce, it is a challenging part of the journey as there are many variables depending on the location. For instance, think of the difference between delivering a package to a customer on a crowded street in Tokyo and delivering it on a rural road in the United States. At Amazon, we're developing specialized solutions for this last mile delivery, with innovations like Amazon Scout, a fully electric autonomous delivery device. The cloud has its own challenges in its "last mile," and I predict a number of innovations are coming to help address it.

Over the last 15 years of AWS, we've built out an impressive global infrastructure that spans over 25 geographic regions and 81 Availability Zones across the globe, with more than 310 points of presence to serve over 245 countries and territories. But now, we are seeing cloud services that are pushing beyond the bounds of our traditional AWS Regions and out to the edges of the network—or, like in ecommerce, the last mile.

We've already seen the cloud go practically everywhere. The shift we'll witness in 2022 is the cloud becoming highly specialized at the edges of the network. To fully realize the benefits of the cloud in workshops and warehouses, in restaurants and retail stores, or out in remote locations, there must be tailored solutions at the edge. The parallels to Amazon Scout in the cloud are devices like Amazon Monitron and AWS Panorama, purpose-built devices that bring cloud capabilities to the edges of the network to do a specific job. They bring all of the high security, advanced features, and speed of delivery of the cloud, but they can be placed almost anywhere in the world. Yet, rather than isolated boxes sitting someplace, these devices become true extensions of the cloud with a link back to all of its core capabilities.

What we will see in 2022, and even more so in the years to come, is the cloud accelerating beyond the traditional centralized infrastructure model and into unexpected environments where specialized technology is needed. The cloud will be in your car, your tea kettle, and your TV. The cloud will be in everything from trucks driving down the road, to the ships and planes that transport goods. The cloud will be globally distributed, and connected to almost any digital device or system on Earth, and even in space.

The rise of **smart spaces**, especially in senior care

In 2022, our homes and buildings will become better assistants and more attentive companions to truly help with our most human needs. The greatest impact in the next few years will be with the elderly.

One of the things that inspires me the most about technology is its role in solving truly hard human problems. After years of meeting with customers across the globe that are using the cloud to do just that, we had an idea to begin documenting their stories through a show called Now Go Build. One of the inspiring customers we featured on the show was a company in Japan named Z-Works, which focuses on improving elder care and helping scale it through technology.
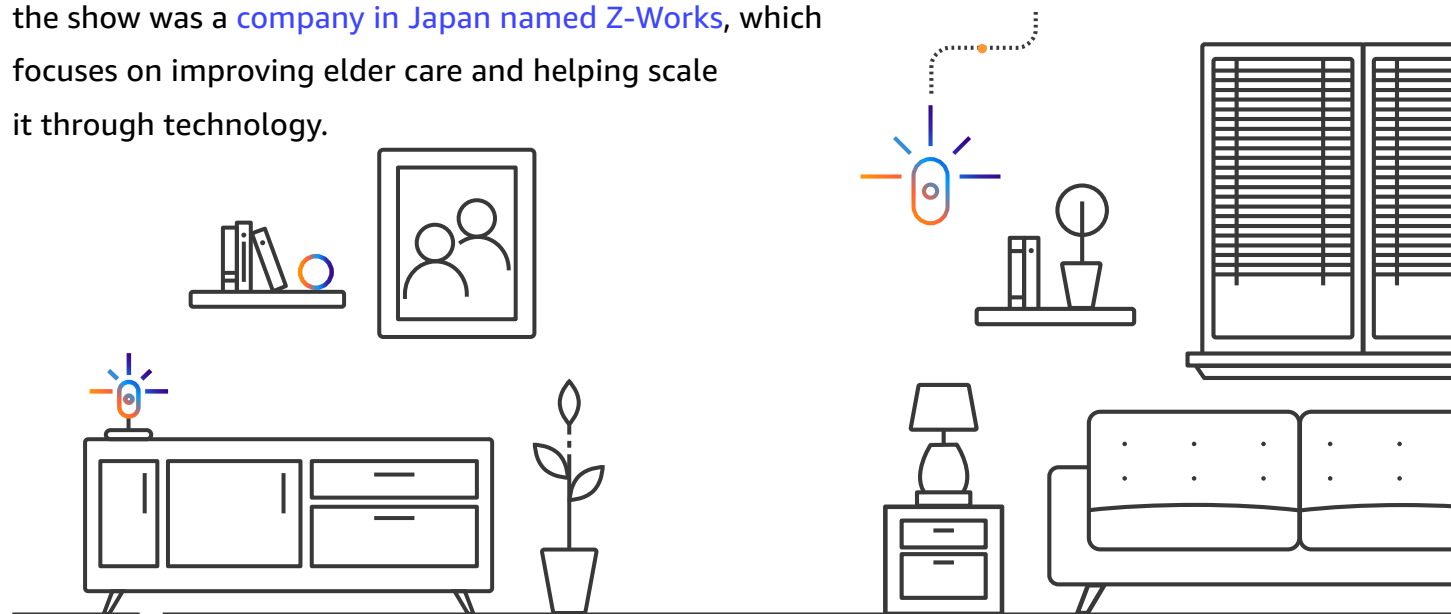
The problem Z-Works faced was how to offer smart and attentive care for seniors in Japan when there are fewer and fewer people available to do the job. The solution this company arrived at was to arrange sensors in beds and throughout rooms in senior homes and connect all of them to the cloud for continuous data analysis. The sensor arrays don't just monitor vital signs. Because they run machine learning models trained in the cloud, the sensors can also tell if a person goes to use the bathroom and simply doesn't return. In that case, the system can notify someone on duty to check on the resident's well-being. In essence, it is a very human response made possible by a very smart space. People would do that intrinsically, if they were aware of the need. In this case, ML models, fed by a private and secure stream of data, arrive at a similarly intelligent action to signal people. We are approaching a point where concepts like ambient computing, collections of IoT sensors, remote/mobile data collection and processing at the edge, and smart devices like Amazon Alexa will have the positive impact we always knew they would.
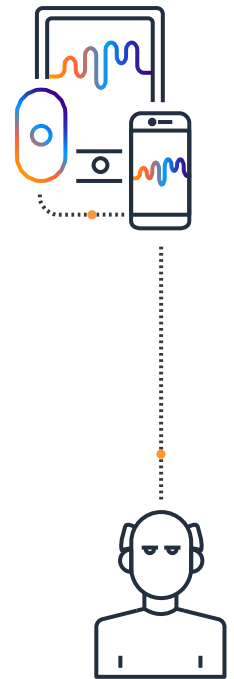
Over the next several years, we'll see smart spaces come to life in a number of settings, but none with higher impact than elder care. It will be a combination of the simple tasks you would expect—from dimming lights, locking doors, and switching off the oven if someone forgets—to the more contextual and proactive things that technology can do: asking questions when normal living patterns diverge and enacting common sense solutions when necessary. It will result in taking better care of people, and in the case of an aging population, it means that we will create a new class of homes so people can actually stay at home.

Sustainability gets its **own architecture**

In 2022, developers will begin to make sustainability-conscious decisions about the systems and applications they are building. They will seek new approaches to cloud architectures that optimize for the needs of the planet as well as the needs of end users.

As developers, we are trained to think about how to optimize our architectures for factors like security, performance, reliability, and cost. In 2022, you can add sustainability to that list. What we will begin to see in the coming years is developers taking an active role in building sustainability-conscious architectures that take into account not just the problems they are solving, but the planet as well.
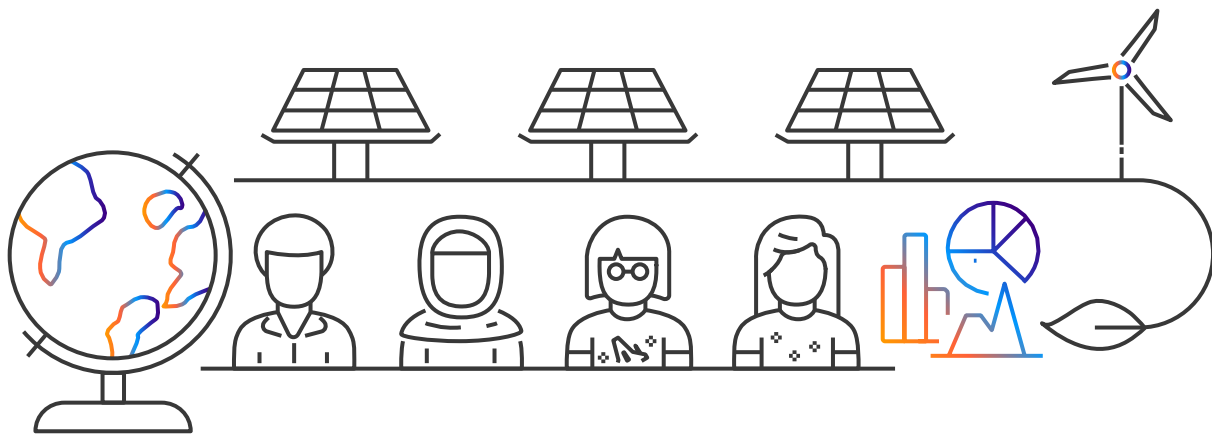
As consumers, we have long wanted our videos and music streamed in higher quality, our webpages to load faster, and the ability to store more and more photos, but people are also beginning to realize the negative impacts this convenience can have on the environment. Do consumers truly need a download to happen as fast as absolutely possible or can we help them make informed decisions about storing an asset in a more carbon-friendly way with a minor penalty in download speed? Likewise, can we provide the ability to stream a video in slightly lower quality than 4K to reduce our carbon footprint and, by doing this at scale, have a positive impact on the planet? Sustainable architectures will take decisions like these into account.

Developers will take an active role in reducing the carbon footprints of their applications. This will happen in a variety of areas, like taking into account where in the world they choose to run their applications to take advantage of green energy in the grid, considering the time needed to process a task, or even specifying the chipset they use. When operating at web scale, small savings can scale out to have a large impact. We will also start to see developers further consider the power of switching off resources. "Always on" is a mantra that many build to, but it comes with a cost. When considering the idle resources needed for an "always on" architecture, we may start to see a new mantra emerge that "the greenest energy is the energy we don't use." This doesn't mean we don't architect for high availability; it just means being more sustainability-conscious in our architectural decisions.

Let's be clear: technology uses energy, and at AWS, we are committed to making smart infrastructure choices. We're on track to run on 100% renewables by 2025— five years ahead of our original goal of 2030. But making the cloud run on wind, sun, or hydropower is only part of the sustainability responsibility that we developers, and really anyone running a company, must shoulder. It is a shared responsibility and a genuine effort that customers, employees, and potential employees will increasingly demand to see.
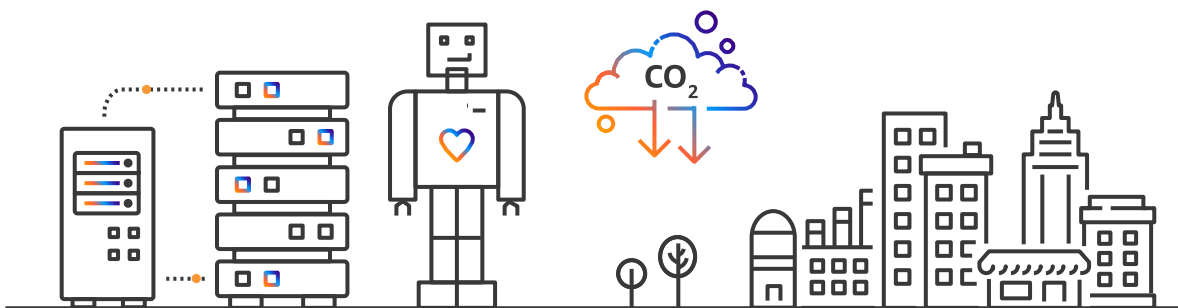
A **new wave** of connectivity will bring about a new class of applications

Low Earth orbit (LEO) satellites are set to bring affordable broadband to every corner of the planet. This will change the lives of billions of people as teachers, students, small businesses, and virtually anyone gets online.

Over the next five years, more than 20,000 satellites are going to fan out in LEO above our planet. Among them will be about 1,500 from Amazon's Project Kuiper, a network of satellites with the goal of delivering fast, affordable broadband to unserved and underserved communities around the world. (The first are planned to go into orbit in the fall of 2022.) What I see coming along with this planet-scale broadband are a whole new class of applications that will benefit from it.
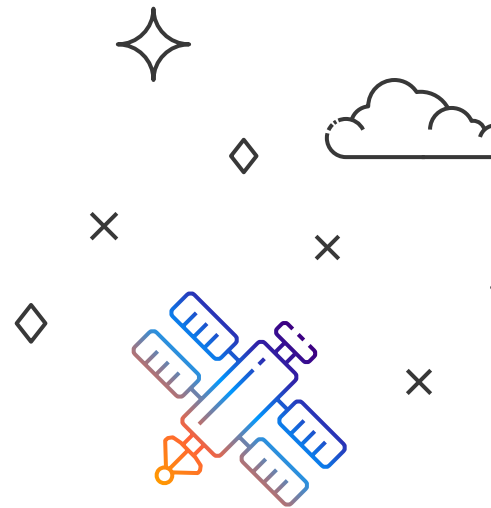
Today, most digital applications are constrained by the existing network—designed for low bitrates or intermittent connectivity. In some cases, we have digital applications and systems that are designed to be operated offline, but these typically go out of date quickly or have limited functionality compared to their connected counterparts. Think about a traditional GPS navigator as compared to using a mobile app on your phone. But what happens when you are no longer constrained by connectivity, bandwidth, or high latency? A world of untapped possibilities will become reality when affordable connectivity reaches these places. Enter LEO satellites.

With ubiquitous connectivity, we start to unlock use cases that simply aren't possible today. Try to imagine what happens in schools when every kid can use the same learning tools, or when small and medium-sized businesses get hold of digital tools they need to win more customers, grow their businesses, and create jobs in rural and remote communities around the world. We can more readily monitor reforestation efforts in remote locations, and we can better track and take quicker action on disaster situations such as fires and floods.

Large enterprises with remote assets—such as solar installations, heavy equipment, or far-flung buildings—will be able to better optimize the use and maintenance of those assets. Transportation companies with vehicles, planes, and vessels on the move will have access to continuous data streams uploaded to the cloud, and regular updates downloaded to vehicles and vessels on the ground, in the air, and on the water. Ubiquitous connectivity will take us from intelligent spaces to intelligent cities, intelligent countries, and finally, toward an intelligent world.