MIT Technology Review
Insights

Produced in association with

UBS

Internet-connected consumer goods are poised for renewed growth, thanks to the advent of generative AI alongside hardware and networking technology innovations.

Housetraining robot dogs: How generative AI might change consumer IoT

Key takeaways

1. The convergence of advances in generative AI, IT hardware, and networking technology means that new and powerful Internet of Things (IoT) consumer devices will increasingly emerge.
2. This technological capacity will dramatically increase the utility and functionality of existing smart devices and traditional appliances. It will also allow innovative new products.
3. Rapid integration of generative AI and consumer IoT is unlikely due to cost, problematic attributes of generative AI, a lack of common implementation standards, and concerns over safety and data privacy in a context of highly limited regulation.

As technology goes, the Internet of Things (IoT) is old: Internet-connected devices outnumbered people on earth around 2008 or 2009, according to a contemporary **Cisco report**. Since then, IoT has grown rapidly. Researchers say that by the early 2020s, **estimates** of the number of devices ranged anywhere from the low tens of billions to **over 50 billion**.

Currently, though, IoT is seeing unusually intense new interest for a long-established technology, even one still experiencing market growth. A sure sign of this buzz is the appearance of acronyms, such as **AIoT and GenAIoT**, or "artificial intelligence of things" and "generative artificial intelligence of things." What is going on? Why now?

Examining potential changes to consumer IoT could provide some answers. Specifically, the vast range of areas where the technology finds home and personal uses, from smart home controls through smart watches and other wearables to VR gaming – to name just a handful. The underlying technological changes sparking interest in this specific area mirror those in IoT as a whole.

Rapid advances converging at the edge

IoT is much more than a huge collection of "things," such as automated sensing devices and attached actuators to take limited actions. These devices, of course, play a key role. A recent **IDC report** estimated that all edge devices – many of them IoT ones – account for 20% of the world's current data generation.

IoT, however, is much more. It is a huge technological ecosystem that encompasses and empowers these devices. This ecosystem is **multi-layered**, although no single agreed taxonomy exists. Most analyses will include among the strata the physical devices themselves (sensors, actuators, and other machines with which these immediately interact); the data generated by these devices; the networking and communication technology used to gather and send the generated data to, and to receive information from, other devices or central data stores; and the software applications that draw on such information and other possible inputs, often to suggest or make decisions. The inherent value from IoT is not the data itself, but the capacity to use it in order to understand what is happening in and around the devices and, in turn, to use these insights, where necessary, to recommend that humans take action or to direct connected devices to do so.

A recent IDC report estimated that all edge devices – many of them IoT ones – account for 20% of the world's current data generation.

"Very small language models now have the sophistication to make these machines more intelligent. With hundreds of words versus just a single wake word, you can start doing more to control the objects around you."

Randy Abrams, Asia Pacific Regional Technology Analyst, UBS

Today's intense interest in IoT arises from various mutually reinforcing innovations within several of its layers. Collectively, these advances will drive more storage, computing, and potentially decision-making capacity to the devices themselves – or the IoT's physical edge. This, in turn, could dramatically increase the utility of IoT technologies.

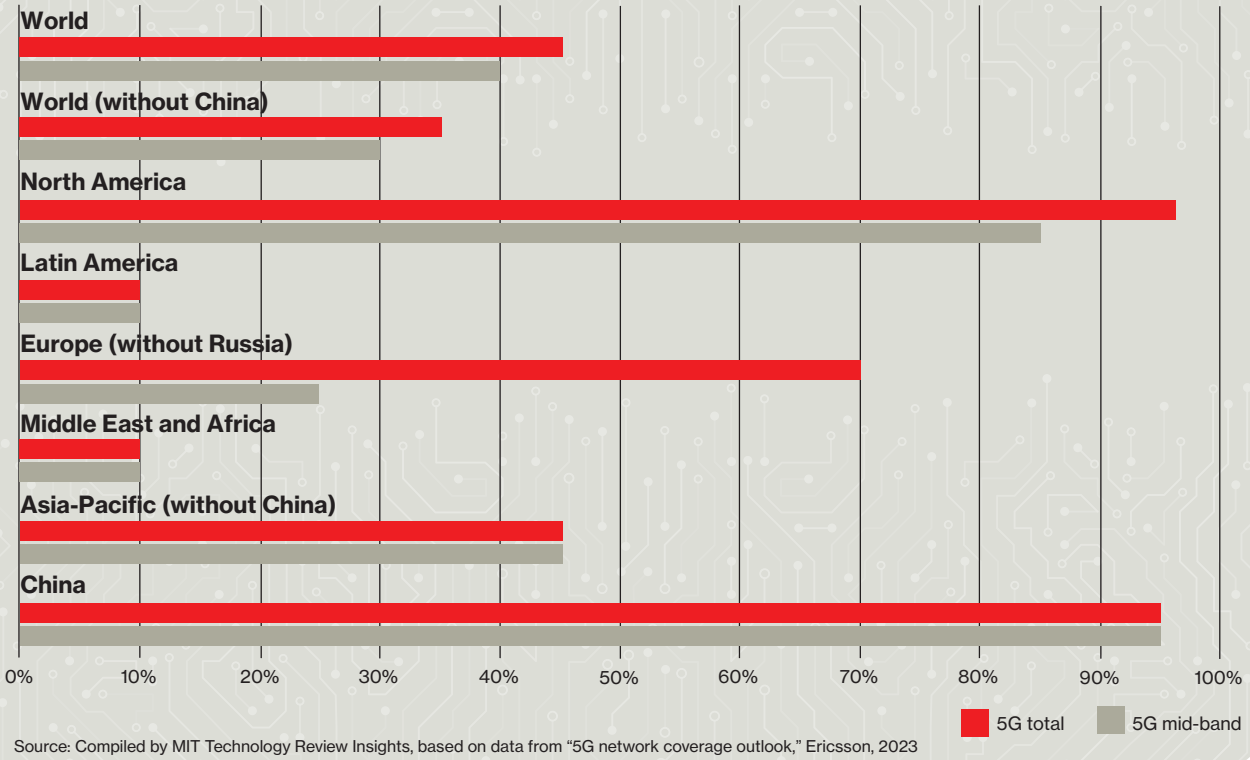
The highest-profile software development of recent years has been the mainstream advent of generative AI, which has also brought renewed interest in older, predictive AI. Both rely on extensive data for training or to populate their learning models. The volume of this data, however, can impede the ability of smaller devices – including IoT – to use such software.

Accordingly, a whole technological subfield, sometimes called tiny machine learning (TinyML), has been emerging to allow such devices to use machine learning and, increasingly, generative AI (hence the acronyms AIoT and GenAIoT). **Trade-offs** inevitably exist: the software on the edge may require extensive parameter pruning or approximations and, as a result, be less accurate or have fewer capabilities. Nevertheless, the increase in user capacity from earlier generations of devices can be huge. Randy Abrams, Asia Pacific Regional Technology analyst at UBS, explains that "very small language models now have the sophistication to make them [these machines] more intelligent. With hundreds of words versus just a single wake word, you can start doing more to control the objects around you."

The highest-profile software development of recent years has been the mainstream advent of generative AI, which has also brought renewed interest in older, predictive AI.

Worldwide 5G mid-band coverage

5G mid-band population coverage was estimated to have reached approximately 40% worldwide by the end of 2023.



Source: Compiled by MIT Technology Review Insights, based on data from "5G network coverage outlook," Ericsson, 2023

Meanwhile, the hardware to run this software is improving rapidly. Microprocessors, in particular, continue their decades-long increase in computational power. One result is a growing market for such chips specifically designed to support AI on IoT and other edge devices. Competing companies include big names such as Qualcomm, Nvidia, and ARM. Already, Abrams notes, it costs only "a few dollars per device to add small blocks of processing memory, and a bit of software development." More advanced capabilities would increase the relative expense, but given historical trends with information technology products, experts believe these costs could drop over time.

Although the above developments allow more processing and decision-making by IoT edge devices, certain more complex tasks will inevitably continue to require interaction between these assets and centralized computing hubs. An important drawback is the communication time required – or latency. The ongoing rollout of 5G mobile telephony, which includes new standards to support IoT connectivity, is meant, in part, to reduce this problem. So too will the greater adoption of **WiFi 6 and 6E**, which are designed to allow faster local hub communication with a high number of low-power IoT devices. New specifications for **WiFi 7**, released in January 2024, take these advances further. Such innovations, says Danny Goh, founder and CEO of Nexus Frontier Tech, an AI research firm, are "breaking the barriers to deploy smart home or other consumer devices freely anywhere you like."

These mutually supportive developments will undergird the increasing integration of generative AI with IoT. Gartner forecasts that, between 2023 and 2027, the global IoT market will grow at an average annual compound rate of 14% and reach $124 billion. During that time, **UBS estimates** that the value of AI devices will grow from 4% of the market to 12%. Meanwhile, in the same period, consultancy GlobalData predicts that IoT software will surpass IoT hardware in market size, consistent with greater computing on these distributed devices.

GlobalData also projects that 27% of the total market will be for consumer IoT. It is therefore worthwhile to examine the likely effects of generative AI on consumer goods at the IoT's edge and the key challenges to implementing the technology.

Innovation of two sorts

The impacts of generative AI on consumer goods fall into two broad categories: transforming already familiar products and allowing the creation of a whole range of new ones.

Abrams explains that it will be relatively straightforward to upgrade any number of commonly owned household and personal smart tools, ranging from smart watches and wearables, through vacuum cleaners and lawn mowers, to smart thermostats and speakers. Goh agrees. While the introduction of generative AI–enabled devices into smart homes "is still very new at the moment," he adds, "things are starting to come in."

The addition of AI will give these devices substantial new capabilities. Goh notes that even the restricted versions of AI on edge devices will enable more complex interaction and commands. The ability to process and store data locally also presents new possibilities. Goh believes that the resultant reduced latency will be crucial for everything from making home vacuum cleaners respond to their environments more quickly to allowing use of facial recognition for home security functions, such as unlocking doors. Moreover, AI-based power management will extend battery power and increase the physical range in which these machines can be deployed.

Nor will evolution be restricted to traditional smart devices. Kitchen appliances are already also gaining new features. For example, Whirlpool, Bosch, and Samsung have introduced generative AI–enabled smart ovens, which can recognize a dish and then cook it until they sense that the food is ready.

As Abrams noted above, the cost of adding appropriate microprocessors to hardware upgrades as consumers replace old ones will be small compared to the overall price of these tools. Goh adds that, in many cases, "right now the sensors and all those smart devices are already very advanced." The holdup is "more on the software side," he explains. However, improvements could conceptually be distributed online.

Generative AI will also allow development of a wide variety of new devices. Already, various existing or soon-to-be-launched products combine in different ways the newly available capacity of edge devices to see, and act on, the world around them.

AI-enabled smart glasses with built-in cameras, such as the Ray-Ban Meta Wayfarer and its competitors, are already available. The wearer of these devices can request descriptions of images captured by the camera, translate any text it sees, or ask for enriched information, such as where to buy a product seen via the glasses.

"Right now the sensors and all those smart devices are already very advanced. The holdup is more on the software side."

Randy Abrams
Asia Pacific Regional
Technology Analyst, UBS

These are only the beginning. Abrams reports that consumer robotics – currently a niche area – could grow quickly depending on what the market provides. Companies are certainly interested. Amazon has begun to sell a range of Enabot robots, which are marketed as family and pet companions as well as security robots. Samsung will soon release its Ballie robot. Both can act as smart device hubs and mobile voice assistants to carry out various tasks. For example, they can launch other IoT-connected equipment, monitor activity in the house for security and health issues, project videos, check emails, or even, when told, feed and play with the family dog.

As ever with a new technology, the extent of innovation will reflect a combination of human imagination and consumer interest. Goh says that "things we only imagined just two or three years ago now can be a lot more real." Indeed, aspects of fairy tales may be possible. A modern Sleeping Beauty cannot yet ask a mirror who is fairest of them all, but Baracoda's smart mirror, BMind, will give mental health coaching to those looking into it.

Barriers to, and risks arising from, generative AI–enabled IoT

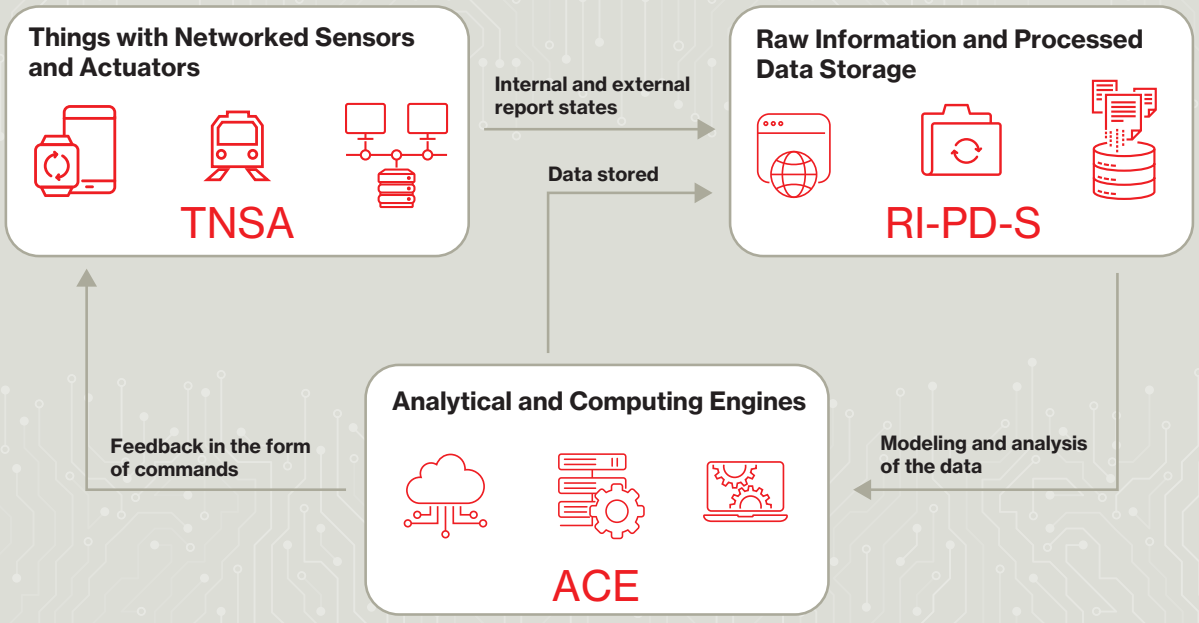
While holding out great promise for consumers, a variety of challenges could dramatically slow adoption or create unintended negative consequences from IoT devices using generative AI. The following four are key issues.

Cost and an exacerbated digital divide. Expenses related to integrating generative AI with IoT are as multifaceted as IoT itself. As mentioned earlier, additional microprocessors for specific devices will add relatively little to the price of already smart devices. New machines will set buyers back more: smart glasses are currently in the low hundreds of US dollars; some high-end robots can cost over a thousand. The market will decide what is viable.

Changes to other IoT layers can be more expensive. At the consumer level, Abrams explains that many will need to upgrade "to a much better home gateway that can deal with interference from a lot more signals and a lot more devices within your home."

IoT component interactions

The three components that make up IoT interactions are things with networked sensors and actuators (TNSA), raw information and processed data storage (RI-PD-S), and analytical and computing engines (ACE).



Source: Compiled by MIT Technological Review Insights, "The Internet of Things: Insights into the building blocks, component interactions, and architecture layers," ScienceDirect, 2018

The biggest expense, however, needs to go into underlying infrastructure. The rollout of 5G telephony, let alone any eventual adoption of 6G, is expected to provide substantial **economic benefits**, according to consultancy PwC. It will, however, require huge investments. In the U.K. alone, the upgrade to 5G – excluding rural coverage – is estimated to require **£34 billion pounds by 2030**, most of which will need to come from state investment. This helps explain why, while in the US and China over 90% of the population have **access to some version of 5G telephony**, in Latin America, the Middle East, and Africa, the average is 10%, according to Ericsson.

On an even more basic need, access to electricity remains a challenge in many developing countries. Formal electrification data does not include reliability – this is clearly illustrated by data from health care, which should presumably have privileged access to power. According to a **World Health Organization report**, a billion people depend on health facilities in locations where the electricity supply is unreliable. They are unlikely to have what they need for extensive IoT use.

As Mark Esposito, a faculty affiliate at Harvard University's Center for International Development, puts it, globally the "infrastructure CAPEX requirements [for generative AI–enabled IoT] are gigantic. I am skeptical that the pace of development will be on par with what private firms are proposing."

Still immature technology. Generative AI is a challenging technology to work with. As Abrams points out, with its propensity for hallucinations, "it can get things wrong and now it's in your house helping you." Moreover, installations of the technology on IoT devices will rely on stripped-down versions of learning models. Esposito points out that using these "is still largely in a testing ground. To what degree will a [a smaller model] be functional to a specific use on the IoT?" This novelty brings more than technical challenges. Abrams adds that companies will need to find application developers to create "a lot of different things that are used in the home." Given that this kind of generative AI is so new, that talent may be difficult, and expensive, to secure, at least initially.

Lack of common standards. Consumers will want their IoT devices, and supporting layers, to work together seamlessly. This could take some time to achieve. Currently, notes Goh, "the adoption of [communication and other] standards in IoT is very different from provider to provider. Everyone is trying to introduce their own." Esposito adds that geopolitics will further slow standards development, citing the conflict between the US and China in establishing technical norms for 5G. Companies, he adds, "cannot simply innovate with the sake of scaling this out right. There will always be some form of a geopolitical implication," including whether any advances raise perceived national security threats.

Safety, data privacy, and regulation. Deploying a technology that can at times be unpredictable in homes, and giving it access to substantial amounts of personal information, brings risks unfamiliar with traditional consumer goods.

Among these is the potential malfunctioning of automated devices. Emerging best practice with generative AI tools in corporate use is to keep a so-called "human in the loop," who can stop processes and equipment when they behave incorrectly. Abrams foresees that consumer IoT devices will need to have not just automated shutdown but controls where people can do so both on site and via cloud-based applications. Esposito agrees that full autonomy for consumer IoT devices will not take place, if for no other reason than it remains unclear who might be legally liable for mistakes.

"Globally the infrastructure CAPEX requirements for generative AI–enabled IoT are gigantic. I am skeptical that the pace of development will be on par with what private firms are proposing."

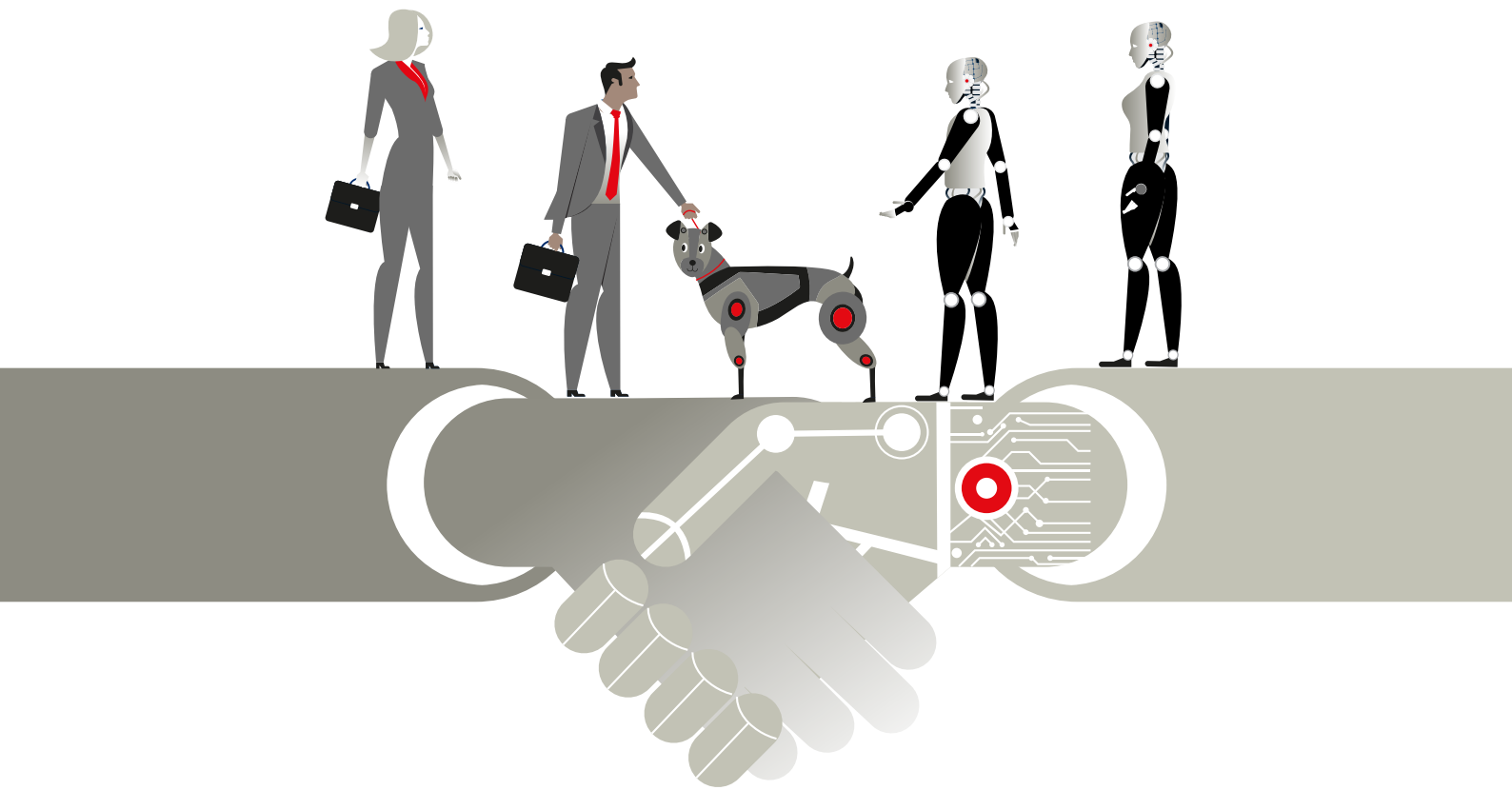
Mark Esposito, Faculty Affiliate, Harvard University Center for International Development

> “We now have more options to allow us to set up the security barrier in each of the devices rather than transmit the data to a central server.”
>
> **Danny Goh**, Founder and CEO, Nexus Frontier Tech

Similarly, personal data leaks could occur in generative AI-enabled IoT devices. Here, the technology itself can help. Goh points out that the capacity to use more data on the device means that “we now have more options to allow us to set up the security barrier in each of the devices rather than transmit the data to a central server.” That said, much will depend on implementation, especially as generative AI may bring with it new threat vectors for hackers to exploit.

Ultimately, regulation will decide how these issues are addressed. Unfortunately, this remains nascent. **Regulation of generative AI** as a whole was put in place in the European Union and China in 2023, but in many other jurisdictions, such as the US and Japan, it is either at the draft stage or non-binding. As for regulatory specifics on the use of generative AI with IoT devices, “simply put, we don’t have this yet,” says Esposito. “Many governments are still trying to figure out the governance.” Goh agrees that key questions remain unanswered. For example, should legal liability lie in the user, manufacturer, or software developer if something goes wrong? Moreover, Esposito notes, case law precedent on which companies could potentially act, or officials base regulation, are also lacking. This will hold back progress more than technological readiness. “I really hope that regulators can give clearer guidance,” concludes Goh.

In short, while technology is moving rapidly to revolutionize consumer technology, any number of factors mean that, in Esposito’s words, “this technology likely is not going to go as fast as the potential of what it could do.”



"Housetraining robot dogs: How generative AI might change consumer IoT" is an executive briefing paper by MIT Technology Review Insights. We would like to thank all participants as well as the sponsor, UBS. MIT Technology Review Insights has collected and reported on all findings contained in this paper independently, regardless of participation or sponsorship. KweeChuan Yeo was the editor of this report, and Nicola Crepaldi was the publisher.

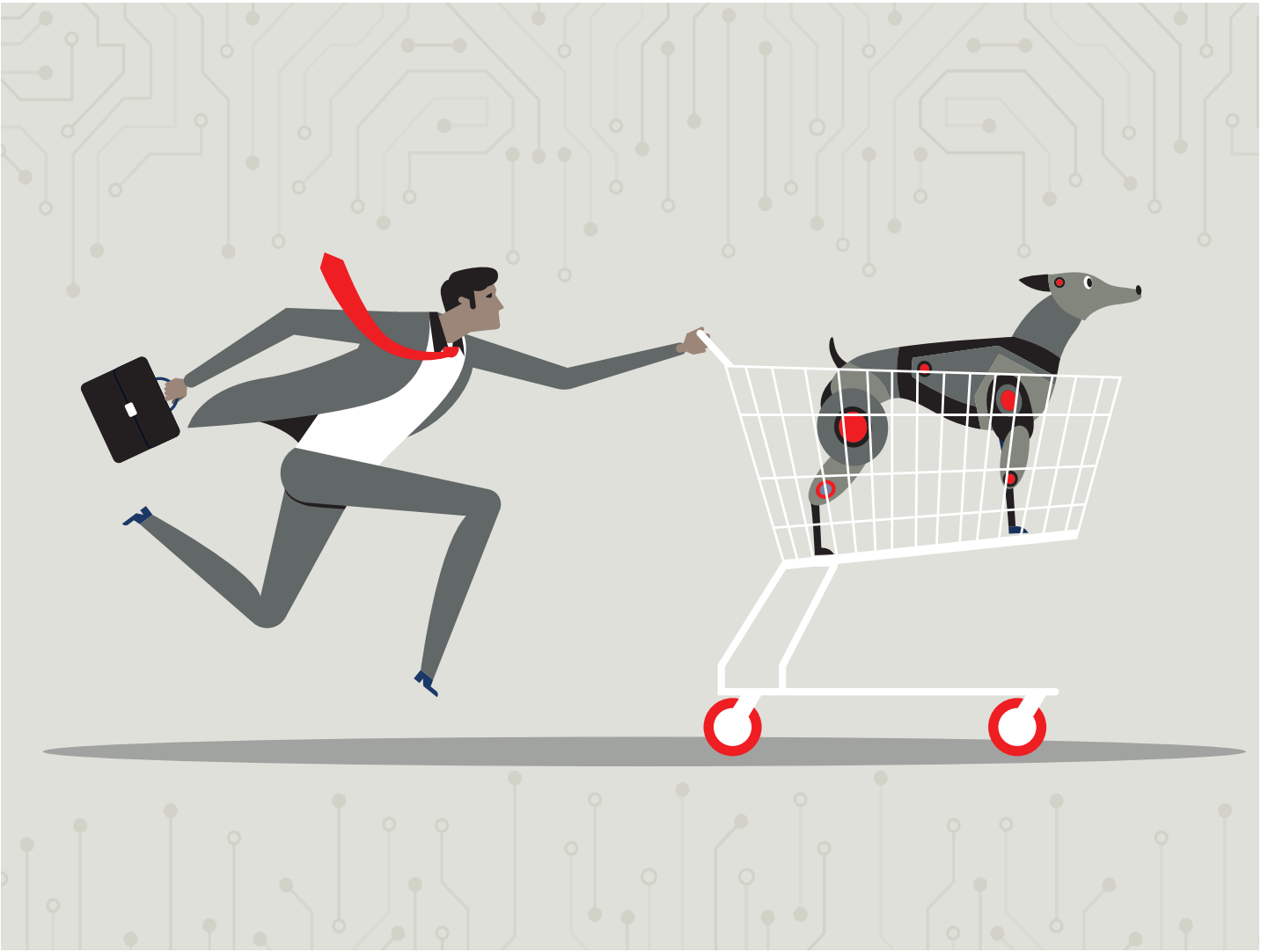
About MIT Technology Review Insights

MIT Technology Review Insights is the custom publishing division of *MIT Technology Review,* the world's longest-running technology magazine, backed by the world's foremost technology institution—producing live events and research on the leading technology and business challenges of the day. Insights conducts qualitative and quantitative research and analysis in the US and abroad and publishes a wide variety of content, including articles, reports, infographics, videos, and podcasts. And through its growing MIT Technology Review **Global Insights Panel**, Insights has unparalleled access to senior-level executives, innovators, and entrepreneurs worldwide for surveys and in-depth interviews.

From the sponsor

UBS is a leading and truly global wealth manager and the leading universal bank in Switzerland. It also provides diversified asset management solutions and focused investment banking capabilities. With the acquisition of Credit Suisse, UBS manages 5.5 trillion dollars of invested assets as per second quarter 2023. UBS helps clients achieve their financial goals through personalized advice, solutions and products. Headquartered in Zurich, Switzerland, the firm is operating in more than 50 countries around the globe. UBS Group shares are listed on the SIX Swiss Exchange and the New York Stock Exchange (NYSE).

Illustrations
Illustrations assembled by Tim Huxford with elements from Shutterstock and Adobe Stock.



MIT Technology Review Insights

www.technologyreview.com

insights@technologyreview.com