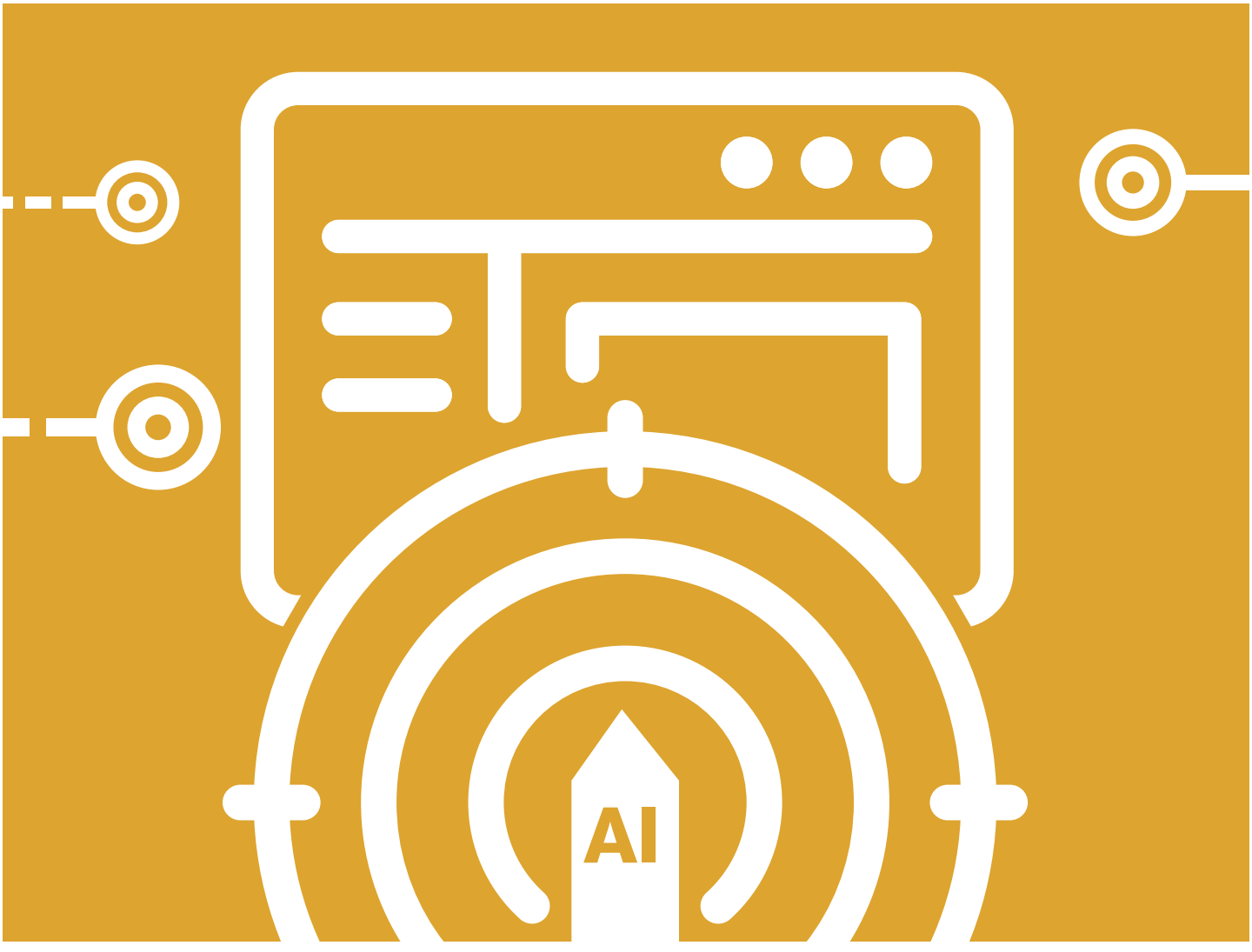
MIT Technology Review
Insights

Produced in association with

DARKTRACE

Artificial intelligence in the hands of cybercriminals poses an existential threat to organizations—IT security teams need "defensive AI" to fight back.

Preparing for AI-enabled cyberattacks



Cyberattacks continue to grow in prevalence and sophistication. With the ability to disrupt business operations, wipe out critical data, and cause reputational damage, they pose an existential threat to businesses, critical services, and infrastructure. Today's new wave of attacks is outsmarting and outpacing humans, and even starting to incorporate artificial intelligence (AI). What's known as "offensive AI" will enable cybercriminals to direct targeted attacks at unprecedented speed and scale while flying under the radar of traditional, rule-based detection tools.

Some of the world's largest and most trusted organizations have already fallen victim to damaging cyberattacks, undermining their ability to safeguard critical data. With offensive AI on the horizon, organizations need to adopt new defenses to fight back: the battle of algorithms has begun.

MIT Technology Review Insights, in association with AI cybersecurity company Darktrace, surveyed more than 300 C-level executives, directors, and managers worldwide to understand how they're addressing the cyberthreats they're up against – and how to use AI to help fight against them.

Key takeaways

1. Cybercriminals are turning to artificial intelligence (AI) to scale up their attacks and evade detection.

2. According to a global survey, more than half of business leaders say security strategies based on human-led responses to fast-moving attacks are failing. Nearly all have begun to bolster their defenses in preparation for AI-enabled attacks.

3. To keep up with evolving cybercriminal innovation, "defensive AI" uses self-learning algorithms to understand normal patterns of user, device, and system behavior in an organization and detect unusual activity without relying on historical attack data.

As it is, 60% of respondents report that human-driven responses to cyberattacks are failing to keep up with automated attacks, and as organizations gear up for a greater challenge, more sophisticated technologies are critical. In fact, an overwhelming majority of respondents – 96% – report they've already begun to guard against AI-powered attacks, with some enabling AI defenses.

About this report

Based on a combination of survey-based market research and in-depth executive interviews, this report explores organizations' biggest cybersecurity concerns and how they are adopting artificial intelligence (AI) in preparation to find and repel AI-enabled cyberattacks. It is sponsored by Darktrace, and the views expressed within are those of MIT Technology Review Insights, which is editorially independent.

- From December 2020 to January 2021, MIT Technology Review Insights surveyed 309 senior global business leaders – 60% of whom are C-level executives or directors.
- Survey respondents are global, with 48% from North America; 36% from Europe, the Middle East, and Africa; 11% from Asia-Pacific; and 5% from Latin America.
- Respondents work in more than a dozen industries: IT and telecommunications, at 31%, represented the largest response group, followed by manufacturing (13%) and financial services (12%).
- Respondents were asked to evaluate how they are preparing for AI-powered cyberattacks.

60% of respondents report that human-driven responses are failing to keep up with automated attacks.

Offensive AI cyberattacks are daunting, and the technology is fast and smart. Consider **deepfakes, one type of weaponized AI tool,** which are fabricated images or videos depicting scenes or people that were never present, or even existed.

In January 2020, the FBI warned that deepfake technology had already reached the point where artificial personas could be created that **could pass biometric tests.** At the rate that AI neural networks are evolving, an FBI official said at the time, national security could be undermined by high-definition, fake videos created to **mimic public figures** so that they appear to be saying whatever words the video creators put in their manipulated mouths.

This is just one example of the technology being used for nefarious purposes. AI could, at some point, conduct cyberattacks autonomously, disguising their operations and blending in with regular activity. The technology is out there for anyone to use, including threat actors.

Offensive AI risks and developments in the cyberthreat landscape are redefining enterprise security, as humans already struggle to keep pace with advanced attacks. In particular, survey respondents reported that email and phishing attacks cause them the most angst, with nearly three quarters reporting that email threats are the most worrisome (see Figure 1). That breaks down to 40% of respondents who report finding email and phishing attacks "very concerning," while 34% call them "somewhat concerning." It's not surprising, as **94% of detected malware** is still delivered by email. The traditional methods of stopping email-delivered threats rely on historical indicators – namely, previously seen attacks – as well as the ability of the recipient to spot the signs, both of which can be bypassed by sophisticated phishing incursions.

Figure 1: Most concerning cyberattacks
While all potential cyberattacks raise alarms, email and ransomware cause executives the greatest worry.

Source: MIT Technology Review Insights survey of 309 business leaders worldwide, January 2021. Respondents were asked to choose all that apply.

"The bad guys know that everybody relies on remote work. If you get hit now, and you can't provide remote access to your employees anymore, it's game over."

Max Heinemeyer, Director of Threat Hunting, Darktrace

When offensive AI is thrown into the mix, "fake email" will be almost indistinguishable from genuine communications from trusted contacts.

How attackers exploit the headlines

The coronavirus pandemic presented a lucrative opportunity for cybercriminals. Email attackers in particular followed a long-established pattern: take advantage of the headlines of the day – along with the fear, uncertainty, greed, and curiosity they incite – to lure victims in what has become known as "fearware" attacks. With employees working remotely, without the security protocols of the office in place, organizations saw successful phishing attempts skyrocket. Max Heinemeyer, director of threat hunting for Darktrace, notes that when the pandemic hit, his team saw an immediate evolution of phishing emails. "We saw a lot of emails saying things like, 'Click here to see which people in your area are infected,'" he says. When **offices and universities** started reopening last year, new scams emerged in lockstep, with emails offering "cheap or free covid-19 cleaning programs and tests," says Heinemeyer.

There has also been an increase in ransomware, which has coincided with the surge in remote and hybrid work environments. "The bad guys know that now that everybody relies on remote work. If you get hit now, and you can't provide remote access to your employees anymore, it's game over," he says. "Whereas maybe a year ago, people could still come into work, could work offline more, but it hurts much more now. And we see that the criminals have started to exploit that."

Figure 2: Challenges of responding to AI attacks

Surveyed business leaders indicate that automated cyberattacks threaten to overwhelm the ability of managerial response and current tools.

Source: MIT Technology Review Insights survey of 309 business leaders worldwide, January 2021. Respondents were asked to choose all that apply.

Cybercriminals are lightning-fast in their attacks, and their dwell time — the length of time they have free reign before their missions are complete — is shrinking to hours rather than days.

What's the common theme? Change, rapid change, and — in the case of the global shift to working from home — complexity. And that illustrates the problem with traditional cybersecurity, which relies on traditional, signature-based approaches: static defenses aren't very good at adapting to change. Those approaches extrapolate from yesterday's attacks to determine what tomorrow's will look like. "How could you anticipate tomorrow's phishing wave? It just doesn't work," Heinemeyer says.

Offensive AI: Not a human-scale problem

Already, cyberattacks are proving to be too fast and too furious for humans and first-generation tools to keep up with, as they struggle to protect data and other assets. The limitations of traditional security tools were made clear once again in December 2020, when a campaign attributed to Russian intelligence groups **infiltrated some of the world's most prominent organizations** — including branches of the United States government and Fortune 500 companies — through their software supply chains. Public health and safety are also at risk — hackers recently attempted to disrupt the **supply of coronavirus vaccines.** And in February 2021, hackers infiltrated the systems of a water facility in Oldsmar, Florida, trying to change the levels of chemicals in the **water supply to poisonous extremes.**

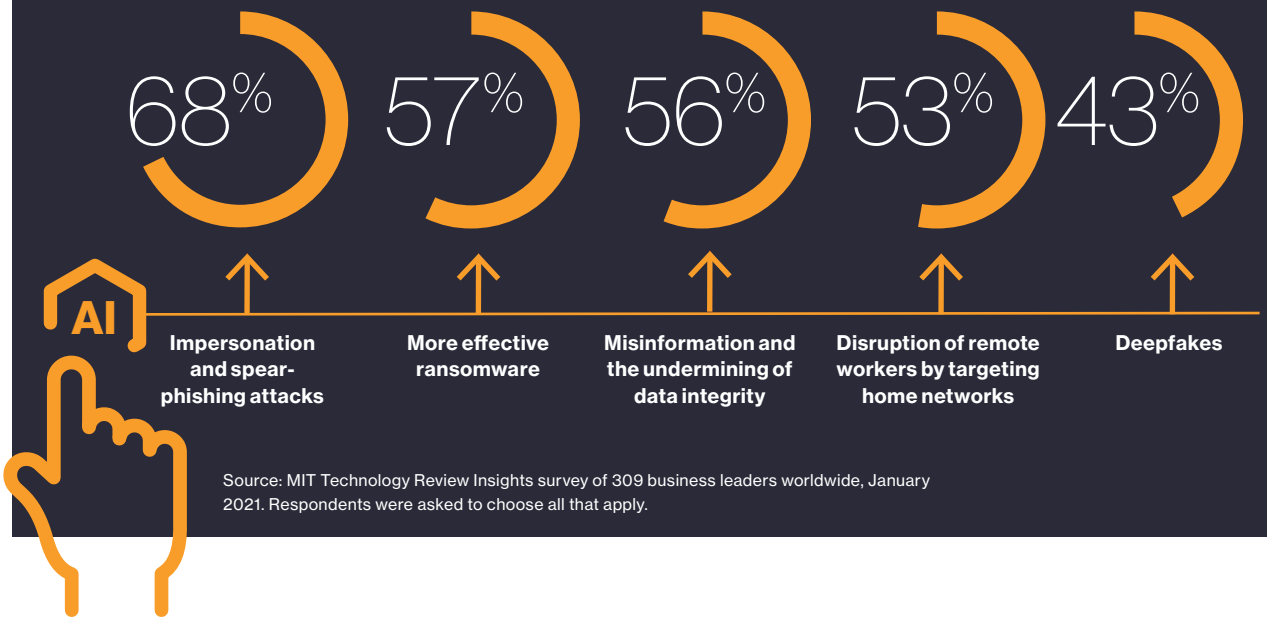
According to survey respondents, companies worry that they have inadequate resources to quell threats. This was, in fact, respondents' biggest challenge: 60% reported that human-driven responses can't keep up with automated attacks (see Figure 2).

The IT skills gap is aggravated by increasing digital complexity, Heinemeyer says. It's not just that things are changing; it's that they're changing in an "increasingly rapid and vicious threat landscape." Cybercriminals are lightning-fast in their attacks, and their dwell time — the length of time in which attackers have free reign in an environment before their missions are complete — is **shrinking to hours** rather than days. Heinemeyer says cybersecurity teams are increasingly relying on AI to stop threats from escalating at the earliest signs of compromise, containing attacks even when they strike at night or on the weekend.

One organization that's embracing the use of AI in cybersecurity is McLaren Racing. In the world of professional car racing, speed isn't just important on the racetrack; it's also crucial when it comes to responding to fast-moving cyber threats. McLaren Racing's principal digital architect Edward Green gives an example: one Saturday afternoon on a race weekend, under intense pressure, the team simply did not have the time to assess whether every email might be a threat. "Everyone was moving very, very quickly," because "you've got a limited amount of time" to read and respond to data and then make adjustments. The quicker the team can access the data flowing from the race cars, the faster it might find an advantage over another team. The data flow on race days is at its peak, the perfect time for an impersonation attack — an email that attempts to impersonate a trusted sender and gain access to data or finances.

On this particular race weekend, McLaren Racing had recently deployed Darktrace's defensive-AI platform, and the technology was already learning what the data flow should look like on racing days. It spotted an email that was unusual in the normal patterns of activity for the sender, recipient, and wider organization — and locked the suspicious link inside the email, so anyone who tried to open it wouldn't have been able to click through to the link.

Figure 3: How AI will be used against companies
AI can be used to impersonate friendly correspondents and launch searing ransomware attacks, execs say.



Source: MIT Technology Review Insights survey of 309 business leaders worldwide, January 2021. Respondents were asked to choose all that apply.

"Increasingly sophisticated social engineering attempts meant that our employees continued to engage with these phishing and spoofing emails."

Edward Green, Principal Digital Architect, McLaren Racing

Regaining the upper hand with defensive AI

Many organizations are turning to defensive AI to fight fire with fire. Rather than relying on historical attacks to find new ones, defensive AI learns what's normal for an organization and can detect abnormal, potentially malicious activity as soon as it appears – even if it has never been seen before.

A year ago, before the pandemic and the issues of a remote workforce complicated the company's security operation, the technology team at McLaren Racing would encounter crude, brute-force password attacks that Green likened to "machine-gunning" of credentials sprayed across Microsoft 365 accounts. In such attacks – known as "spray-and-pray" – hackers employ bots to automatically try to log in to their targets' systems by using lists of user credentials stolen in other breaches.

But in the past year, these attacks have been tailored to focus on individuals, roles, or teams, Green says. "They're far more targeted," he says. "Attackers are impersonating employees, or they're going really smart, and embedding themselves inside these new, transformative digital processes," such as signing a convincingly forged document or joining conference calls.

Spear phishing – or sending emails to specific targets – is getting more refined and is the big challenge for his team, Green says. Email attacks targeting users have sought to solicit fraudulent payments or access intellectual property. "Increasingly sophisticated social

Thousands of organizations rely on AI to react to a fast-developing cybersecurity incident, whether or not their security teams are in the office.

engineering attempts meant that our employees continued to engage with these phishing and spoofing emails, despite having an array of tools and procedures in place to avoid such an eventuality," Green says.

These are huge challenges for organizations. When survey respondents were asked how worried they are that future cyberattacks against their companies will use AI, 97% cited future AI-enhanced attacks as troubling, with 58% of respondents saying such cyberattacks are very concerning. When asked which attacks, in particular, are worrisome, the most respondents, 68%, reported that impersonation and spear-phishing was their biggest fear (see Figure 3).

With humans unable to keep up with the pace of AI innovation, let alone respond fast enough, new technological answers are needed. Thousands of organizations rely on AI to react to a fast-developing cybersecurity incident, whether or not their security teams are in the office.

Known as an autonomous response, and enabled through self-learning AI, the technology can surgically interrupt an in-progress attack without interrupting day-to-day business. Here's an example of an autonomous response in action: an **electronics manufacturer was hit by ransomware** that rapidly spread, encrypting files. The strain of ransomware had never been encountered before, so it wasn't associated with publicly known compromise indicators, such as blacklisted **command-and-control domains** or **malware file hashes**. But the autonomous response AI identified the novel and abnormal patterns of behavior and stopped the ransomware in seconds. The security team then had enough time to catch up and perform other incident response work.

Figure 4: Gearing up for AI attacks
Nearly all organizations plan to protect themselves against evolving threats: just 4% report they're doing nothing at all.

Source: MIT Technology Review Insights survey of 309 business leaders worldwide, January 2021. Respondents were asked to choose all that apply.

Organizations need to reform their strategies quickly, be prepared to defend their digital assets with AI, and regain the advantage over this new wave of sophisticated attacks.

Defensive AI is a force multiplier, Heinemeyer says. By automating the process of threat detection, investigation, and response, AI augments human IT security teams by stopping threats as soon as they emerge, so people have the time to focus on more strategic tasks at hand.

Preparing for offensive AI

The vast majority of respondents are actively gearing up for AI-powered cyberattacks. Security teams are increasingly relying on autonomous technologies that can respond at machine speed when a cyberattack occurs. When asked how they're preparing, survey respondents said that outside of allocating more budget to IT security and security audits, their organizations have also prioritized several defensive AI projects (see Figure 4).

With the onset of AI-powered attacks, organizations need to reform their strategies quickly, be prepared to defend their digital assets with AI, and regain the advantage over this new wave of sophisticated attacks.

Fortunately, it's easier to flip the switch than some may realize. McLaren Racing's Green can speak to that: when the pandemic forced lockdowns, he and his infrastructure team installed Darktrace's AI-powered technology to defend their email environment within two days. With the help of AI, Green's team can now focus on strategic priorities instead of stamping out the small fires of constant, low-level alerts. "Our cybersecurity team can then work with us on all of our weird and wonderful sensors in the cars and make sure those are nice and secure."

It's a simple formula: IT teams need to be duly prepared, Green says, because cybercriminal minds are turning to AI, too. "Much in the same way that lots of organizations look to use AI and machine learning to be more competitive, more efficient, to solve those big challenges they've got as companies, then you would start to expect that the same tools you're using to be more efficient and effective – other people will use those to try to attack you."

To learn more about how AI responds to sophisticated cyberattacks, visit **darktrace.com/en/supercharged-ai/.**

"Preparing for AI-enabled cyberattacks" is an executive briefing paper by MIT Technology Review Insights. We would like to thank all participants as well as the sponsor, Darktrace. MIT Technology Review Insights has collected and reported on all findings contained in this paper independently, regardless of participation or sponsorship. Jason Sparapani and Laurel Ruma were the editors 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 Panel, Insights has unparalleled access to senior-level executives, innovators, and thought leaders worldwide for surveys and in-depth interviews.

From the sponsor

Darktrace is the world's leading cyber AI company and the creator of Autonomous Response technology. It provides comprehensive, enterprise-wide cyber defense to over 4,500 organizations worldwide, protecting the cloud, email, IoT, traditional networks, endpoints, and industrial systems.

A self-learning technology, Darktrace AI autonomously detects, investigates and responds to advanced cyber-threats, including insider threat, remote working risks, ransomware, data loss and supply chain vulnerabilities.

The company has 1,500 employees and 44 office locations, and is headquartered in Cambridge, UK. Every 3 seconds, Darktrace AI fights back against a cyber-threat, preventing it from causing damage.

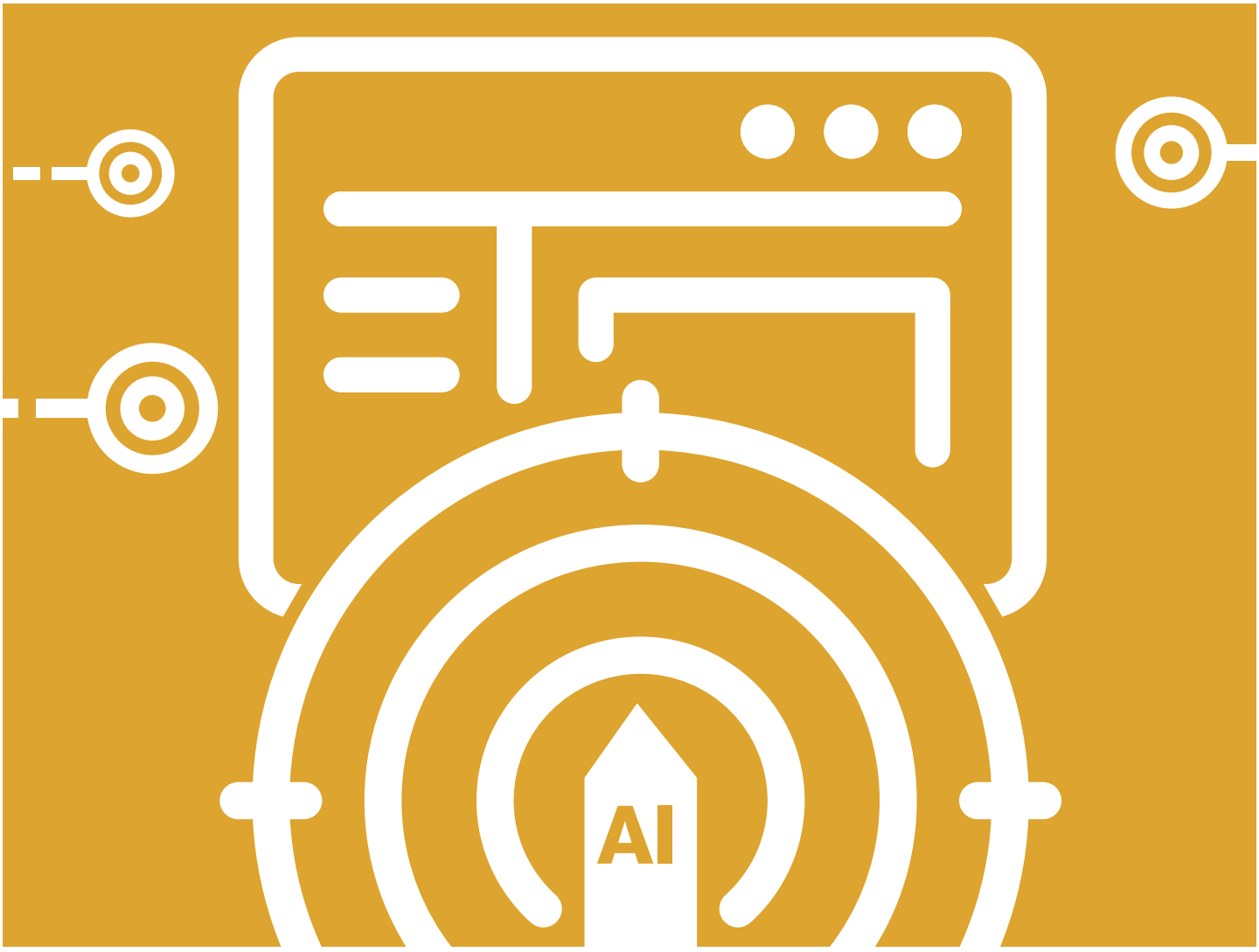
DARKTRACE

Illustrations

All Illustrations assembled by Scott Shultz Design. Covers, pages 3 and 7 illustrations by Motorama, Shutterstock. Page 4 illustration by Real Bot, Shutterstock; pages 5 and 8 illustrations by MuPlus, Shutterstock.

While every effort has been taken to verify the accuracy of this information, MIT Technology Review Insights cannot accept any responsibility or liability for reliance on any person in this report or any of the information, opinions, or conclusions set out in this report.

© Copyright MIT Technology Review Insights, 2021. All rights reserved.



MIT Technology Review Insights

www.technologyreview.com

@techreview @mit_insights

insights@technologyreview.com